EXPLORING THE ROLE OF TRANSPOSABLE ELEMENTS IN REPRODUCTIVE ISOLATION

Gerardo Antonio Serrato-Capuchina

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology in the Graduate School.

Chapel Hill
2019

Approved by:

Daniel R. Matute

Karin S. Pfennig

Brian Langerhans

Christopher Martin

Corbin Jones

ABSTRACT

Antonio Serrato-Capuchina: Exploring the Role of Transposable Elements in Reproductive Isolation
(Under the direction of Daniel R. Matute)

Understanding the phenotypic and molecular mechanisms that contribute to genetic diversity between and within species is fundamental in studying the evolution of species. In particular, identifying the interspecific differences that lead to the reduction or even cessation of gene flow between nascent species is one of the main goals of speciation genetic research. Genetic novelty is often implicated in the origin of new molecular functions and in some cases, new phenotypes. As a result, the mechanisms that produce genetic changes are integral in understanding the evolution of species as well as speciation itself. Here, I explore the reproductive consequences of novel mobile genetic elements across the *D. simulans* complex and its influence on hybridization between Drosophila species along distinct points in the speciation continuum. A precise quantification of the frequency of P-elements (PEs), and transposable elements in general, across different species is still in its infancy. Even though the phenomenon of hybrid dysgenesis has been rigorously characterized in *D. melanogaster*, the discovery of PEs in other *Drosophila* species allows us to understand how these elements behave in different genetic backgrounds and the role PEs, and TEs generally, play in affecting gene flow between closely related species. My work suggests that the invasion of TEs in a genome should not be quantified as the singular phenotype of hybrid dysgenesis, but rather multiple distinct effects that vary in intensity based off PE copy-number. I also explore the how PEs affect gene

flow between naturally hybridizing species, the first such case in understanding PEs and reproductive isolation between species. This work suggests that 1) the deleterious effects of PEs within a species is dependent on copy-number and 2) PEs lead to a reduction in hybrid fitness within a species-complex, increasing reproductive isolation and potentially facilitating the speciation process.

To my mother, who taught me the value of hard work and always believed in me.

ACKNOWLEDGEMENTS

First, I'd like to thank my immediate family for raising me in a manner that emphasized the virtue of hard work and the value of family. My parents Guadalupe (Lupita) and Miguel "El Tigre" Serrato who provided me with everything I needed in order to reach my dreams, sacrificing everything they could without a moment's hesitation so that I could have a shot at the American Dream. To my siblings Mikey, Yuri, and Vianey whom despite distance and time, I can always count on to love me unconditionally as we take our own independent paths in life and expand our families. Additionally, my godparents Teresa and Rogelio Gonzalez who provided me an example of independence and success while simultaneously remaining nurturing and supportive, it eased my mind greatly knowing that if all other avenues failed, I could always count on them.

To my friends throughout all stages of life, I am sincerely thankful to have had you in my life. Growing up would have been a dull affair for it not been for my brother-from-another-mother David Johnson and my cousin Angel L. Gonzalez, it has been a great pleasure making mistakes together and I'm very happy we were able to mature into fine outstanding citizens and look forward to seeing where life takes you both. I'm particularly thankful for my life partner, Sofia Maria de la Serna, for having willfully joining her life to mine and providing me an avenue towards my own emotional growth, through her selfless dedication to me and our growing family. To Matias Serrato, whom I learned the meaning of unconditional love from, may you remain a joy filled soul as you grow up, you fill my heart with a rush of love simply with a giggle.

I could not have completed my work without the great support of my advisor Daniel R. Matute. Support is not limited to my own dissertation work, but also the comradery and genuine concern you felt towards my development. I look forward to continuing our interactions throughout the years, as peers but most importantly friends. Although our personalities differ, our life philosophies overlap greatly, it was incredibly helpful to no longer feel like a cultural outsider. Also, thank you Corbin Jones for always having time for helpful advice, both personal and academic, while maintaining a high level of positivity despite your busy schedule. I hope to integrate lessons I have learned from both of you into my professional goals and wish you all the best. Thank you for making me feel like you genuinely cared, I would not have completed my degree without your vital support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

.

Chapter 1: The Role of Transposable Elements in Speciation

Aim 1: Assess the importance of TEs in reproductive isolation.

My goal is to compile all the available literature to understand what is the relative importance of TEs in generating phenotypes that can either set the speciation process in motion or keep potentially hybridizing species apart. The compiled information will reveal whether there is a consensus on the role of these mutations in speciation or, on the contrary, if there are no particular known trends.

Introduction

Speciation is the evolutionary process by which one lineage splits into two reproductively isolated groups of organisms [1]. One of the central goals of speciation research is to understand the processes that drive the evolution of reproductive isolation (RI) between species [2–6]. Significant strides have been made towards identifying barriers that generate RI between species [1,7], the processes underlying their evolution [2,3,8–11], and the rate at which they evolve during speciation [4,5,12–15]. Even though some progress has been made in identifying genes and loci associated with RI, few studies have explored the evolutionary processes that produced these barriers. Because of this, there is not yet a consensus as to what types of mutations or which mechanisms are typically involved in speciation or RI.

There are two broad approaches to identify the genetic underpinnings of RI. First, if crosses can be made, one can genetically map the loci underlying RI between organisms. Such studies can establish the genetic changes that maintain species identity and, if divergence is recent, potentially

reveal the molecular changes that were initially involved in speciation. This approach is particularly informative when coupled with closely related organisms at different stages of reduced gene exchange [1,16,17]. An alternative approach is to assess whether a particular type of molecular change is commonly associated with isolation between genotypes. If a barrier to gene flow is commonly caused by a certain type of molecular change, then one can argue that that molecular change is important in either the origin of new species or the persistence of them when they face the possibility of collapse through gene flow. This approach has, for example, revealed that chromosomal inversions are commonly associated with the suppression of recombination and frequently harbor gene combinations involved in isolation between species [18,19] (reviewed in [20]). However, this approach has rarely been used to understand the impact of other molecular changes on RI. Here we highlight transposable elements as recurring agents that underlie a variety of manifestations of RI, which suggests they should be explored across various taxa in order to better understand their mechanistic and evolutionary contributions.

Transposable elements (TEs) are DNA sequences able to copy and insert themselves throughout the genome. TEs represent up to 80% of nuclear DNA in plants, 3–20% in fungi, and 3–52% in metazoans [21–23]. TEs are classified according to the mechanism they use to transpose. Class I elements require an RNA intermediate in order to integrate/duplicate themselves within a genome, while Class II elements act without an intermediate through a cut-and-paste mechanism that replicates its DNA directly to DNA as it mobilizes (Figure 1). A full classification of TEs is shown in Table 1. Interestingly, the predominant class of TEs can vary greatly between taxa [24–28] and species, and their genomic frequency, location, and activity levels can vary greatly even at the population level. TEs were described for the first time in maize by Barbara McClintock in 1950 [29] where they lead to somatic mutations affecting various phenotypes/genes depending on

their chromosomal location and transposition time. The insertion of a TE can disrupt the coding or regulatory sequences of genes, which can cause deleterious effects by the modifying or eliminating a gene's expression [30–34]. TEs are ubiquitous throughout nature [35–37] and their effect on their hosts' fitness is generally considered to be deleterious; TEs are commonly considered selfish elements. However, gene disruptions are not the only consequence of TEs as they transpose throughout the genome. TEs can also cause regulatory changes, genomic expansions, and generate new chromosomal variants through the generation of inversions. Moreover, TEs can produce all of these changes rapidly [38–40] and in response to abiotic stressors—a hypothesis first advanced by McClintock [29]. These changes can provide genetic and phenotypic novelties upon which selection can act [41,42]. Due to their potential to generate novelty when it is needed, some have hypothesized that TEs are maintained in genomes through multilevel selection [43–45].

Thus, TEs are diverse and pervasive components of eukaryotic genomes that have the potential to impact rates of diversification and adaptation. TEs have also long been known to cause RI between genotypes (e.g., [46]). However, the role of TEs as a molecular mechanism capable of directly mediating the origin of new species remains underexplored experimentally.

The idea of selfish genetic elements and their involvement in the formation of new species has been latent in speciation genetics for years [47,48]. RI due to intragenomic conflict (i.e., conflictual speciation, reviewed in [49]) seems to be common but until recently was thought to be rare. Meiotic drive, endosymbionts, and maternal effects have all been implicated as potential sources of RI [1], and theoretical models have examined what role they may play in speciation [50,51]. Yet, the role of TEs in the initiation of the speciation process and in maintaining species has only rarely been experimentally studied. In this review, we highlight research that emphasizes

TEs as important agents involved in the origin and persistence of species, with a focus the evidence for how TEs contribute to contemporary RI. We also propose future directions and questions that need to be addressed in order to understand whether transposable elements are involved in speciation, in the maintenance of species by generating reproductive isolation, and whether they cause distinct macroevolutionary dynamics.

2. Transposable Elements and Reproductive Isolation

Traits involved in keeping species apart can be classified depending on when they occur in the reproductive cycle. Premating barriers include ecological and behavioral traits that reduce the likelihood that two individuals will mate and include habitat isolation and mating choice. Post-mating-prezygotic barriers involve interactions between gametes and include sperm/pollen-egg incompatibility. Finally, postzygotic barriers arise after fertilization has occurred, and include various forms of fitness reductions in hybrids [1,3,52]. The genetic basis of prezygotic and postzygotic reproductive isolating mechanisms has been studied in varying degrees (reviewed in [1,52]), and a few studies have examined their connection to TE transposition (Table 2). In the following sections, we compile the cases for which TEs have been found to affect a trait potentially involved in RI, in an effort to emphasize their potential role as agents involved in various forms of reproductive isolation.

2.1. Premating Isolation I: Transposable Elements and Ecological Isolation

TEs have been hypothesized to promote local adaptation and enable the invasion of new habitats [53]. The initial colonization of a new environment is often accompanied by a reduction in genetic diversity as a result of genomic bottlenecks or founder effects. This hypothesis posits that by rapidly creating new genetic diversity, the transposition of TEs might help populations

adapt to their new environment. Encountering a new environment is frequently stressful, and since TEs can be induced by stress TEs could facilitate an increase in genetic diversity exactly when it is needed [54–56]. The genomic shock model proposed by McClintock [29] that TEs mobilize in response to environmental challenges has been supported by many studies across multiple taxonomic groups [51,57,58]. New environments can select for different traits, and if these traits are associated with assortative mating (i.e., dual traits due to pleiotropy; [59,60]), then RI can evolve through divergent selection [61]. Therefore, we hypothesize that TEs could frequently underlie ecological adaptation and perhaps ultimately, speciation. A roadmap to assess whether local adaptation is commonly caused by TEs has been proposed elsewhere [53]. Notably, methods to detect TEs have evolved over the last five years and a fine scale dissection of the identity of the TEs in a genome and their copy-number throughout the genome is now feasible (Table 3), facilitating population level analysis. To examine evidence for our hypothesis, we focus on phenotypes that might lead to premating isolation and for which TEs have been shown to cause phenotypic differences.

Flowering time: Differences in flowering time are a common barrier to gene flow in angiosperms [62]. The mode of action is simple: differences in flowering time lead to RI between genotypes as the gametes of the two genotypes show a reduced probability of encountering each other. Additionally, changes in flowering time have several downstream effects that can further reduce the possibility of gene flow [63]. Besides the lack of contact of gametes due to the temporal differences, differing flowering time might also lead to differences in pollinators and thus fosters even stronger isolation than that caused by temporal differences alone.

In at least two cases, genetic mapping has revealed TEs underlying the disruption of genes involved in the pathways involved in flowering time and photoperiod.

The vegetative to generative transition 1 (*Vgt*, *ZmRap2.7*.1) locus in maize is an upstream (70 kb) noncoding regulatory element of a repressor of flowering. At *Vgt1*, a miniature inverted repeat transposable element (MITE) insertion into a conserved noncoding sequence was previously found to be highly associated with early flowering in independent studies [64]. The insertion of a *CACTA*-like transposon into the promoter of a second locus, *ZmCCT*, can suppress its expression through methylation and reduces maize sensitivity to photoperiod [65].

Similarly, in *Arabidopsis*, a recessive allele at the locus flowering locus C (*FLC*), is a result of disruptions of the gene by non-autonomous *Mutator*-like transposons, which ultimately leads to a delay in flower time. This transposon renders *FLC* subject to repressive chromatin modifications mediated by short interfering RNAs generated from homologous transposable elements in the genome [66].

TEs might play a role in floral induction and development in the rice shoot apex as a portion of them are silenced during floral induction [67]. The exact role these TEs play in floral induction is unknown, but the recurring activation and silencing of particular TEs, in particular *Gypsy* elements, at specific developmental stages suggests a regulatory overlap in reproductive development and TE produced small interfering RNA (siRNA). The downregulation of some retrotransposons stops them from repressing genes related to their transition into the reproductive phase, essentially activating genes required for flowering.

These examples show that TEs can modify the mean flowering time through a variety of mechanisms (e.g., differential methylation in maize, repression of an intron via siRNA in *Arabidopsis*). However, to our knowledge, all flowering time mapping cases have been done within species, thus far no case of between species difference in flowering time has been ascribed to TEs. It is worth noting that different molecular mechanisms resulting from TE insertion can

produce the same phenotypic outcome. In maize, a TE insertion results in differential methylation in the regulatory region while in *Arabidopsis* there is repression at an intron through siRNAs. Interspecific differences in flowering time caused by TEs remain unidentified but it seems like a possible cause of isolation.

Habitat isolation: Abiotic factors such as light and water availability can greatly influence the range over which a plant is able to spread, as well as influence the conspecific mates it will encounter. In the extreme, adaptation to a new environment can completely prevent contact with other members of a species, initiating the process of allopatric speciation.

The *CACTA*-like TE insertion (in the promoter for *ZmCCT*; [65]) implicated in photoperiod sensitivity in maize, has also facilitated local adaptation to temperate long-day environments. Additionally, variation in drought tolerance has been linked to a TE inserted in the promoter region of *ZmNAC111* [65]. This *MITE* insertion results in histone hypermethylation, which represses the expression of NAC resulting in a higher drought tolerance.

Selection acting across a continuous distribution of habitat preference can lead to RI as a byproduct of local adaptation to changing environmental factors. TEs may have generated the alleles selected during adaptation to temperate climates in *Drosophila melanogaster*. A study comparing temperature/latitudinal clines along Australia and North America found 10 TEs that show signs of positive selection at their insertion points, resulting in local adaptation [53,58,68]. By causing mutations in genes associated with a suite of traits, including circadian rhythm regulation and starvation resistance, several types of TEs (Long terminal repeats (*LTRs*), Long interspersed nuclear elements-like (*LINE*-like), and Terminal inverted repeats (*TIR*)) are thought to underlie the phenotypic differences along the cline. Suggestively, the TEs were more likely to be adaptive in temperate populations compared to tropical populations where they were likely to

be neutral [68]. Taken together, these results strongly suggest that alleles generated by TEs were favored during local adaptation.

Host specificity in oomycetes: One of the main mechanisms of RI in plant pathogens is host specificity, which is regulated by the repertoire of effector genes within each pathogen. Effector proteins alter host physiology and allow colonization by individual pathogens [69]. In oomycetes, genomic distribution of TEs is frequently predictive of host specificity [70–72]. For example, the genome of *Phytophthora*, a major pest of commercial crops, harbors multiple families of retrotransposons (*copia*, *Gypsy/Ty*) [72–76]. In *Phytophthora infestans*—the potato blight pathogen—host specificity is regulated in part by RXLR class effectors that enable *P. infestans* to utilize a host [77]. As in other systems, TE insertion in *P. infestans* causes epigenetic silencing of both the transposon and nearby genes, resulting in regulatory differences. Notably, synthetic chimeras of a short interspersed element (*SINE*) to an effector gene in *P. infestans* leads to the silencing of both the introduced fusion and endogenous homologous sequences [77]. This silencing is also likely to occur naturally in the genome of *P. infestans*, as transcriptional inactivation of effectors is known to occur and over half of RXLR effectors are located within 2 kb of transposon sequences in the *P. infestans* genome. Thus, it is possible that host range in *P. infestans* was shaped by TEs inserted near these genes. Since mating in oomycetes occurs on host plants, it is plausible that TE insertions that modify host specificity have led to reproductive isolation in *P. infestans*. However, a systematic exploration of effector genes and their interactions with TEs would be needed to test this hypothesis.

2.2. Premating Isolation II: Transposable Elements and Sexual Isolation

Self-incompatibility: Fungi engage in diverse reproductive strategies, which often vary between closely related species [78]. Fungi often employ a mating system whereby the mating

type—which is analogous to the sex—of the individual is determined by alternative alleles at one or several loci. In homothallic strains, which can mate with themselves (i.e., are self-compatible (SC)), additional loci generate allelic diversity at the mating type loci by a copy-paste mechanism (e.g., the homothallic switching (HO) endonuclease in *Saccharomyces cerevisiae*) [79,80]. Since single loci can effectively determine whether two individuals can interbreed or not, TEs can mediate transitions from homothallism (SC) to heterothallism (self-incompatible (SI)) in fungi by disrupting these loci. Transitions from self-incompatibility to self-compatibility are associated with speciation events (e.g., [81–83]) because selfing species are effectively isolated from other individuals and species (with the possible exception of somatic fusion; [84,85]).

Retroelements have contributed to the shifts from heterothallic ancestors to homothallic species in the *Neurospora* genus through mediating translocations at the mating-type (MAT) loci [86]. Retrotransposon insertions in the MAT locus also occur in *Blastomyces* and might be involved in decreasing the likelihood of recombination between mating types [87]. In other fungi, transposons have been found within or flanking MAT loci (e.g., *Neosartorya fischeri* [88], *Cryptococcus neoformans* [89], *Paracoccidioides brasiliensis* [87,90]), thus potentially providing an avenue for mating type to evolve independently through a rapid TE-induced mechanism. Specifically, in *Neurospora*, the transposition of *nsubGypsy* has facilitated the movement of genes neighboring the MAT loci to a different chromosome [84]. Transposition of *npanLTR* facilitates unequal crossovers between unrelated intergenic regions of opposite mating types, which in turn facilitates the transition into self-crossing species. Phylogenetic studies in *Neurospora* and *Kluyveromyces lactis* show multiple transitions from SI to SC species [86,91]. Taken together, these studies demonstrate that TEs may frequently be inserted at MAT loci, but it remains to be seen whether these patterns can be extended to other species.

Besides these effects on mating compatibility, genomic rearrangements mediated by transposition can also lead to viability issues in hybrids [92,93]. Barley rusts, *Ustilago hordei*, show a large increase in TE activity not observed in other closely related species (*Ustilago maydis* or *Sporisorium reilianum*), which has also led to both the reorganization of the MAT loci in the former species as well as large chromosomal rearrangements [94]. Few reports have explored a potential causal connection between TE activity and genome reorganization. A systematic assessment of how often TEs are involved in gene movement across chromosomes is sorely needed.

Transposons may also play an important role in transitions to self-compatibility in plants. *Solanum*, a flowering plant genus that contains tomatoes, consists of SC and (SI) taxa, with multiple transitions from self-incompatibility to self-compatibility [95]. SC taxa are characterized by low levels or no expression of stylar RNase (S-RNAse). The seven SC and the three SI taxa differ in the 5′ coding region of S-RNAse by several point mutations. Additionally, in one of the SI taxa, the source of low S-RNAse levels stems from an insertion of a transposon-like repetitive element. These results show how single-base mutations and the insertion of TEs can result in similar evolutionary outcomes [95].

These results suggest that transitions to self-fertility mediated by TEs might be common in fungi and plants. We hypothesize that since transitions to self-incompatibility have been associated with lower speciation rates and higher extinction rates in plants [96–98], TEs might be associated with differential diversification rates (i.e., species selection [99–101] in fungi. A formal test of this hypothesis remains to be performed.

Mating behavior in *Drosophila*: Behavioral isolation in *Drosophila* is mediated through a multimodal signaling system that involves cuticular hydrocarbons (CHCs), visual cues, and

auditory signals [102–104]. CHCs are waxy compounds that are involved in desiccation protection (e.g., the species pair *D. serrata*/*D. birchii*; [105,106]) in the abdominal cuticle and are often necessary for mate discrimination and in some cases species discrimination [107–110]. Marcillac et al. [111] studied the effects of an insertion of a TE in the *desat1* locus and measured two different traits: the expression of CHCs and the ability of males to discriminate between the sexes. Even though no naturally occurring TEs have been found in the *desat1* locus, over 30 TEs have been found ~20–50 kb upstream of the gene [112,113]. *desat1* mutants (i.e., with a TE insertion) had lower CHC abundance (reducing the natural sex dimorphism) than lines without the TE. Moreover, mutant males showed poorer discrimination between control males and females suggesting that the TE insertion changed not only the emitted sexual signal but also how that signal is recognized. It remains to be seen if there are naturally occurring transposon-induced mutants in *desat1* or any other allele involved in the production of CHCs.

TEs have been conclusively shown to lead to interspecific differences in mating song in some *Drosophila*. Male flies in the *D. melanogaster* species subgroup produce a courtship song with two components: trains of continuous sinusoidal sound, called sine song, and pulses separated by an interval, called pulse song [27,114]. In the case of the sister species *Drosophila simulans* and *Drosophila mauritiana*, two species that diverged within the last 240,000 years [115–117], *D. mauritiana* males have a higher song frequency than *D. simulans* males, which in turn affects mating behavior and is a trait used by females to distinguish between conspecific and heterospecific males [118,119]. A retrotransposon, *Shellder*, has caused the disruption of the slowpoke (*slo*) locus in *D. simulans* [120]. The *slo* gene is expressed broadly in the fly nervous system and influences many locomotor behaviors and the insertion of *Shellder* leads to alternate splicing of the gene. *Shellder* insertions are polymorphic in their insertion sites in wild type strains

of *D. simulans* and *D. mauritiana*, which strongly suggests that *Shellder* is probably propagating actively in *Drosophila* populations. The retrotransposon insertion seems to be polymorphic within *D. simulans*, which then leads to the question of whether this has led to isolation between different genotypes of *D. simulans*.

2.3. Transposable Elements and Postzygotic Isolation

TEs and chromosomal rearrangements: Chromosomal rearrangements are one of the genome features known to affect the likelihood of gene flow between species (extensively reviewed in [19,121–123]). In general terms, theoretical models indicate that chromosomal inversions can preclude gene flow at certain regions of the genome. Multiple empirical examples have shown that chromosome rearrangements can indeed contribute to postzygotic isolating mechanisms [124] and assortative mating [122], particularly when the rearranged regions contain alleles involved in reproduction. An active research program is trying to assess whether TEs can indeed lead to the origination of new chromosomal rearrangements (illustrated in [125] and reviewed in [126,127]). In *Drosophila buzzati*, the breakpoints of the *2j* inversion contain TEs. It has been hypothesized that *2j* might have originated by ectopic recombination of the TE at its breakpoints [128]. Even though this inversion has not been formally associated with RI, *2j* is involved with differences in life history traits among *D. buzzatii* populations [129,130]. The phenotypic effects of *2j* are contingent on genetic background, which suggests epistatic interactions with the rest of the *D. buzzatii* genome [131,132]. If TEs commonly induced inversions and other chromosomal aberrations, then TEs might play a role in maintaining species boundaries.

TE reactivation: In animals, fungi and plants, TEs are often targeted and silenced by siRNAs [133]. In plants, siRNAs involved in heterochromatin formation often target TEs and silence them [134]. Unlike animals, where the germ cells are formed early in development, plant germ cells

differentiate from somatic cells in the adult and the chromatin remodeling ATPase decrease in DNA methylation 1 (*DDM*1) is crucial for this process. In *Arabidopsis, DDM1* is necessary to silence TE activity [133–136]. Even though TE reactivation and accumulation is restricted to the vegetative nucleus and not the sperm cells, TE accumulation in the vegetative nucleus can affect the sperm cells of the pollen and result in heritable changes [133]. In tobacco, just as in *Arabidopsis*, cytoplasmic connections between sperm cells and the pollen vegetative nucleus have previously been observed [133,137] and might provide a channel for siRNA and facilitate TE silencing. As a result, TE misregulation, which in essence is a hybrid specific defect of the TE-repressor system, might be a potential source of hybrid defects in pollen.

DDM1 is also required to produce hybrid vigor (heterosis; [138]). *Arabidopsis* F1s between divergent accessions regularly show hybrid vigor in vegetative biomass throughout their lifecycle [139]. However, crosses involving *DDM1* loss-of-function mutants do not show heterosis; TEs are extensively expressed, which in turn causes abnormal and expression of genes related to salicylic acid metabolism [140]. Since fitness is so drastically affected by TEs, through either heterosis or hybrid incompatibility, these results might indicate that expression of TEs in hybrids changes their epistatic landscape (in a way that does not occur in pure species) with potentially deleterious effects. The role that *DDM1* plays in establishing RI could be tested by mutating *ddm1* across multiple plant lineages. The results from such mutagenesis approach will reveal whether this epigenetic regulator of TEs is involved in reproductive isolation in multiple species pairs.

Hybrid breakdown through deregulation of TEs is another postzygotic barrier between species. Lake whitefish lineages have repeatedly colonized postglacial lakes across North America. During these colonizations, a dwarf limnetic species has evolved from a benthic species

multiple times. This repeated evolution has led to incomplete RI between the limnetic and benthic lineages [141,142]. Although the two lineages can produce viable hybrids, there is significant mortality in all hybrid types and backcrosses regularly show a malformed phenotype. Analysis of the transcriptome of hybrids reveals a 232-fold increase in TE activity in malformed embryos compared to pure crosses. This transcriptome wide deregulation of TEs results in shutdown of vital metabolic pathways drastically reducing the fitness of hybrids [143].

The reactivation of retroelements in hybrids can also lead to changes in chromatin profiles. Interspecific crosses of two Wallaby species, *Wallabia bicolor* and *Macropus eugenii*, produce hybrids with autosomes from *Macropus eugenii* that have a larger centromere [144,145]. The extended centromeres differ from those found in either parental species as hybrid centromeres consist primarily of un-methylated retrotransposons. TEs, then, can also affect chromatin structure and chromosomal composition in hybrids. Transpositions resulting from TEs being released from siRNA, epistatic, or epigenetic suppression mechanisms are pervasive across various eukaryotic groups and drastically change the fitness of hybrids.

Hybrid inviability: An extensively studied case of reproductive isolation is the genetic interaction between *Hmr*, *Lhr*, and *gfzf* in F1 hybrids between *D. melanogaster* females and *D. simulans* males. Alleles from these genes genetically interact to cause hybrid lethality between *D. melanogaster* and *D. simulans* [146–149]. RNA-seq analyses revealed that *Hmr* and *Lhr* are required to repress transcription from satellite DNAs and many families of TEs in their native hosts [149]. One possible cause of aberrant TE expression in hybrids is altered expression of *Piwi*-interacting small RNAs (piRNAs), a class of small RNAs that interacts with the *Piwi* family of Argonaute proteins to control the expression of TEs in the germline [150]. This is because the piRNA population in a host rapidly adapts, within ~6 generations [151], to the TE content through

generation of new piRNA clusters, allowing de novo production of piRNA and other types of small RNAs for silencing of the invading TE [152,153]. Overexpression of TEs is frequently found in F1 hybrids, and is often associated with male sterility [154,155]. Overall these results suggest that the regulation of TEs might be of importance in maintaining contemporary species boundaries.

Hybrid dysgenesis: *Drosophila* is arguably one of the premier systems to understand the spread of TEs in animals. At least three families (*hobo*, *P*-elements, *I*-elements) have been found in *D. melanogaster* [156–158]. Of these families, *P*-elements (PEs) have received the most attention, as a result of a suite of defects in F1 hybrids (i.e., hybrid dysgenesis). Hybrid dysgenesis occurs in F1 hybrids from crosses between an uninfected female and an infected male [159,160], whereas individuals from the reciprocal cross are fertile. In dysgenic individuals, TEs proliferate and lead to a suite of defects such as chromosomal breakage, germ line cell apoptosis, and an increase in point mutations [46,161–163]. Despite drastic consequences PEs have spread throughout *D. melanogaster* [112,158,164] and *D. simulans* worldwide [165]. PEs are thought to have originated in the neotropical *D. willistoni* species group [166–168]. Although mites have been proposed to serve as a vector for PEs, potentially as a byproduct of their syringe-like feeding method [169], the precise mechanisms of this horizontal transfer remain unknown and untested.

The unidirectional development of hybrid dysgenesis between crosses stems from the way that genomes protect themselves the deleterious effects of PE activation. In F1 hybrid females, hybrid dysgenesis is only present in daughters from mothers with no PE and fathers with PEs. Usually the infertility that characterizes hybrid dysgenesis is silenced through piwi-interacting RNA silencing [170–173], which are exclusively maternally inherited. piRNAs seem to be present in all arthropods [174], and in the case of *Drosophila* piRNAs are cytoplasmatically deposited in embryos from females that contain PEs. Recent work shows that piRNAs are not

alone in mitigating PE's effects. PEs in *D. melanogaster* lead to hybrid sterility when the germoplasm does not carry the molecular machinery to regulate the expansion of PEs through minimizing cell apoptosis by co-opting the use of genomic maintenance genes such as *p53* [175].

A similar phenomenon, yet much less studied, occurs in *Drosophila virilis* [176]. The elements *Penelope, Ulysses, Paris* and *Helena* and *Telemac* have rapidly increased in frequency in natural populations. Experimental injection of *Penelope* causes germ line mutations as well as the activation of other TEs [176]. Similar to the hybrid dysgenesis phenomenon observed in *D. melanogaster*, when uninfected females are crossed to infected males, the resulting progeny show a high level of gonadal sterility, chromosomal nondisjunction and rearrangements, male recombination, and the occurrence of multiple visible mutations. There are however, notable differences between these two systems. While in *D. melanogaster* only one family of TEs are activated at once, in the *D. virilis* dysgenesis, all families are activated simultaneously [176–177]. The *Penelope* family seems to be primarily responsible for the hybrid dysgenesis syndrome of *D. virilis* [176].

If hybrid dysgenesis is a mechanism that can generate RI in populations of the same species, then the molecular machinery that regulates TEs might be important to not only maintain species boundaries at present but also facilitate speciation. This includes an assessment of whether TEs and TE-repressor system act as traditional genetic incompatibilities in hybrids [178]. A valuable research avenue will be to evaluate the effects of PEs in interspecific crosses and whether hybrid dysgenesis is a source of selection for speciation via reinforcement.

Genomic imprinting in endosperm: Maturation of the embryo in angiosperms is contingent on normal development of the endosperm, a tissue that feeds the embryo during seed development [179,180]. Allocation of nutrients in the endosperm is consistent with parental conflict theory and

excess dosage of paternal alleles promotes larger seeds while an excess of maternal alleles produces small seeds. This tissue is usually triploid and its normal development depends on the proper balance of gene imprinting [180]. Imbalances between paternally and maternally imprinted genes can lead to changes in gene expression through regulatory changes, a phenotype that is commonly aberrant in heterospecific hybrids (e.g., [181,182]).

Arabidopsis arenosa and *Arabidopsis thaliana* hybrid seeds show an overgrown endosperm and arrested or abnormal embryo development. *A. thaliana* harbors LTR retrotransposons of the *Ty3/Gypsy* family, known as *Athila*. These elements are large, with an internal region up to 10.5 kb long, flanked by an average of 1.8 kb LTRs on either side. This internal region produces two proteins: the *gag* capsid structural protein and *pol*, which carries the protease, reverse transcriptase and integrase domains essential for element duplication [183,184]. Seed inviability is positively correlated with the relative paternal genome dose, suggesting that maternal genomic excess suppresses incompatibilities in hybrids [181]. Moreover, the maternal genomic contribution (and thus seed viability) is inversely correlated with expression of *Athila* retrotransposons, expressed mostly from the pericentromeric regions. The normally silenced *Athila* (but not other TEs) is extensively expressed in hybrids. Only the paternal, and not the maternal, copies are expressed in these interspecific hybrids.

The precise reason why TEs are misregulated in hybrids relative to parentals remains unclear and likely varies across species. The interactions between paternally and maternally imprinted genes might lead to changes in silenced regions, which in turn is a common cause of postzygotic isolation in heterospecific crosses. Imprinting in plants is intimately associated with changes to methylation of TEs [185,186], and TE activity is known to alter DNA methylation patterns and gene imprinting in plant genomes [187–189]. Alternative molecular mechanisms—that might act

in concert with perturbed imprinting—have also been proposed to account for seed failure, such as poor regulation of TEs by siRNAs in hybrids [190].

A systematic exploration of how often TEs promote post-zygotic isolation remains a promising research avenue to understand the link between TEs and speciation.

3. Introgression and Transposable Elements

Introgression, which is defined as the transfer of genetic material between species through the production of fertile interspecific hybrids, has recently been shown to be common across all domains of life [191,192]. Understanding what factors allow for gene exchange is crucial to understanding how species—especially nascent ones—persist in cases where they have the chance to interbreed and fuse into a single lineage. The relationship between transposable elements and introgression is multifaceted and includes (i) TE-aided introgression of non TE-DNA and (ii) interspecific transmission TEs alone.

First, TEs might facilitate or hamper introgression of surrounding DNA. Surprisingly, this hypothesis remains untested even though its prediction is straight forward: if TEs increase the likelihood of introgression, then in hybridizing species regions that are TE-rich should show a larger amount of introgression compared to the rest of the genome. If, on the contrary, TEs hamper introgression through selection against regions containing TEs, then TE-rich regions should be refractory to introgression. These two scenarios are illustrated in Figure 2. Even though no systematic study has addressed whether TEs facilitate introgression, there are some indications TEs might be involved in horizontal gene transfer (HGT) [193,194]. The coffee berry borer beetle, *Hypothenemus hampei* [195], and the mustard leaf beetle, *Phaedon cochleariae*, appear to have acquired the genes necessary for their specialized diet through a HGT from bacteria [196], allowing them to degrade plant cell walls. Interestingly in both cases, the genes acquired by the

beetles are flanked by two transposons. The potential role TEs might have played in this transfer remains suggestive but inconclusive.

Introgression might also lead to the transfer of TEs across species boundaries [197,198]. HGT have been linked to speciation events (or at least specialization events) in bacteria, providing novel gene sets that expand host specificity. Horizontal gene transfer regularly acts as a genetic bridge between vastly diverged species [193,194]. Horizontal transfers of TEs between angiosperm genomes have been documented in nature [199–201] and experimentally [202]. In *Drosophila*, HGT seems to have occurred from the *willistoni* species group to *D. melanogaster*. The two groups diverged over 50 million years ago and there is no possibility of hybridization [203]. Many other cases of HGT between species (with a rapidly growing list) have also been reported but the precise mechanisms of gene exchange remain largely unknown and might differ between taxa and reproductive strategies [179,204,205]. By serving as a pathway to TE acquisition, HGT can result in RI when coupled with the effects of new TEs entering a genome.

The most likely mechanism of transfer of genetic material between closely related species is arguably the production of fertile hybrids with subsequent introgression. Even though it is clear that TEs can be mobilized by HGT, it remains unclear to what extent TE activation can occur through introgression. This question remains largely unexplored both in natural and experimental populations. This scarcity is puzzling because the proposal that introgression mediated by hybridization could lead to transposon introduction and mobilization within the genome of rice is not new (i.e., a genome shock, [206]). Two examples of TE mobilization following introgression stand out. First, recombinant inbred lines produced by hybridizing rice species (cultivar Matsumae and wild rice *Zizania latifolia*) have shown that the miniature-Ping (*mPing*) TE together with its putative transposase-encoding partner, Pong, can be mobilized between species [207,208]. Likely,

the mobilization of mPing and Pong is a result of introgression-induced malfunction of the established cellular control systems in the rice genome, as their transposition is transitory and rapidly repressed.

The second example comes from experimental hybrid swarms between two divergent species of *Drosophila*: *D. melanogaster* and *D. simulans*. Both species harbor the *Bari-I* element, a Class II TE with an open reading frame able to encode a polypeptide with 339 amino acids. (The sequence of the putative protein in *Bari-I* is similar to the transposase of the *Tc-1* element of *Caernorhabditis elegans*, which might in turn suggest HGT across animal orders [209].) In synthetic hybrid swarms using *D. simulans* C167.4, an unusual line that produces fertile hybrid offspring with *D. melanogaster*, *Bari-I* elements, originally from the *D. melanogaster* parent, are maintained in hybrid strains, suggesting that introgression can indeed be a mechanism of transfer of TEs. The element is present across the geographic range of both species and shows such similar sequence that it seems to be transmitted horizontally and not vertically [210].

Introgression of TEs has been hypothesized for *Drosophila bifasciata* and *Drosophila imaii* [211], species of the *simulans* complex [212], species of the groups *willistoni* (reviewed by [213]), *saltans* ([214,215]), and the species pair *Drosophila seriodo* and *D. buzzatii* [216]. The main lines of evidence in these studies have been the ability of these species to produce fertile hybrids and the sequence similarity of TEs across species [211,217–219]. A full and detailed characterization of the rates and nature of introgression awaits for most of these groups and should be coming in the near future as TEs will continue to be a focus of research due to their diverse effects across organisms.

4. Future Directions

The relationship between TEs and RI is an open field of research that will likely increase in prominence over the next few years. Given the broad range of roles TEs have played in affecting gene exchange between various species, further study is required in order to better understand the extent to which TEs influence evolution and speciation. Box 1 lists focal questions that remained unanswered. These questions fall into three broad categories.

4.1. Are Transposable Elements a Common Cause of Reproductive Isolation?

Mapping the precise genetic basis of interspecific differences will reveal what type of mutations and genomic interactions are more likely to cause and maintain interspecific differences and their relative contribution to various forms of RI. This will lead to a better assessment of the relative importance of TEs as a genetic cause of RI. A second line of research will explore the role of TEs in adaptation to the peripheral areas of geographic range of a species. In maize, for example, *Mutator* TEs are reactivated in response to environmental stress [220], which is most likely to occur at the edge of the optimal range of the species. TE reactivation might induce to genomic changes that in turn lead to RI between peripheral populations in extreme environments and the central populations' (akin to peripatric speciation; [1,221]). Moreover, hybrid zones are usually found at the edges of the range of the hybridizing species so the interplay of hybridization and potential activation of TEs due to environmental or competition induced stress should be examined (Questions 1–5 in Box 1).

4.2. Are Transposable Elements Responsible for Differential Rates of Diversification?

The broad range of genome sizes across eukaryotes is partially explained by the quantity of repetitive, non-coding DNA—including TEs—interspersed throughout the genome [22,179,222–

226]. The consequences of genome expansions are significant and have been linked to the duration of meiosis, ecological distribution, speciation rate, and extinction risk (e.g., [227,228] reviewed in [229]). Genome rearrangements and, in particular, genome duplications have been associated with higher rates of diversification in teleosts [230–232] and angiosperms [233,234]. The reasoning behind why genome duplications lead to an increase in diversification rates remains unclear but generally there are two explanations. First, genome duplication allows for gene subfunctionalization and neofunctionalization that would not be possible in a non-duplicated genome [235–238]. Second, large genomes might simply have the chance to accumulate more hybrid incompatibilities. Only one systematic evaluation of the relationship of genome size and cladogenesis has been performed (for angiosperms) and it found evidence of a positive correlation between overall genome size and rates of speciation [239,240]. Since TEs commonly lead to an increase in genome size, this is consistent with the hypothesis that invasion by TEs can increase the rate of speciation.

An evaluation of this hypothesis has been carried out in haplochromine cichlids. A comparative analysis to determine what traits were correlated with successful adaptive radiations in Lakes Malawi and Victoria found that traits like decoupled pharyngeal jaw and maternal mouth brooding—which have been hypothesized to be key innovations enabling diversification in cichlids—could not account for differences in the rate of diversification in this group. In contrast, increased numbers of *SINE* insertions preceded the extensive radiations within each lake [241]. These results are consistent with TEs mediating adaptation through either gene disruption or altered methylation patterns near insertion sites. However, determining whether TEs generally lead to increased speciation rates will require a formal macroevolutionary test in which the sample size (i.e., potential radiations caused by TEs) is larger than one [242].

Conversely, TEs could result in increased extinction rates and therefore lead to decreased diversification rates. This is related to an intriguing hypothesis that posits that asexual groups form reciprocally monophyletic clusters (akin to asexual species) rapidly but also disappear rapidly due to the proliferation of deleterious transposons inherited from their sexual progenitors that cannot be purged by recombination, leading to extinction [42,243]. (Questions 6–9 in Box 1) Additional studies are required to move conclusively test a possible connection between TEs and diversification rates.

4.3. Are Transposable Elements and Hybrid Dysgenesis a Source of Selection for Reinforcement?

Hybrid dysgenesis is a phenomenon that occurs in animals and might also exist in plants. Even though its natural frequency remains currently unknown, it is possible that it might be rather common [174,177]. Similarly, reinforcement, the evolutionary process in which prezygotic isolation is strengthened as a byproduct of the production of unfit hybrids, seems to be pervasive in nature [244]. If F1 interspecific hybrids consistently suffer fitness defects due to hybrid dysgenesis, then natural selection might indirectly penalize individuals that mate with heterospecifics, thus fostering the completion of speciation (i.e., increasing RI until there is cessation of gene flow). This question also remains unanswered and will require the identification of sister species that hybridize in nature and for which hybrid dysgenesis represents a major cost to heterospecific mating (Questions 10–11 in Box 1).

Box 1. Unanswered questions about the connection between TEs and speciation.

1. Are interspecific differences in flowering time disproportionately caused by TEs?
2. Do transposons play a significant role in pathogen adaptation to new hosts?
3. How commonly are TEs involved in antibiotic resistance?
4. How is the likelihood of chromosomal inversions caused by recombination affected by TEs?
5. What is the role of TEs in causing hybrid breakdown?
6. Do TEs regularly mediate the transition from outcrossing to self-crossing in fungi?
7. What is the taxonomical distribution of TEs?
8. Do TEs cause changes in the net rates of diversification across the tree of life?
9. Can TEs be deleterious enough to cause extinction?
10. Is introgression facilitated or hampered by TEs?
11. Does hybrid dysgenesis facilitate speciation by reinforcement?
12. Can TE-repressor systems generate hybrid incompatibilities during speciation?

5. Conclusions

Transposable elements are hypothesized to promote bursts of diversification or biological and genomic differentiation between species (e.g., [51,56]). Yet there is little direct evidence that TEs can indeed facilitate RI and ultimately speciation. That does not mean TEs are not related to the generation of new genetic elements, genetic circuits, and ultimately of phenotypes. On the contrary, TEs are commonly associated with the origin of new genetic and phenotypic diversity through gene regulatory element innovation, genic disruptions, siRNA/epigenetic suppressor mismatches, and chromosomal remodeling. In vertebrates, TEs have regularly contributed to the evolution of regulatory and coding sequences, leading to new lineage-specific gene regulations

and functions. Their role has been pivotal to generate new phenotypic diversity. In primates for example, TEs are the main source of new variants in regulatory sequences [245]. In angiosperms a significant portion of adaptive novelty is due to the activity of TEs (active TE-Thrust), resulting in an extraordinary array of genetic changes, including gene modifications, duplications, altered expression patterns, and exaptation to create novel genes, with occasional gene disruption [179]. Even though it is clear that TEs are involved in generating the genetic material for new traits (some of them involved in adaptations), the question of whether TEs are involved in RI has remained largely understudied. The combination of natural history, genetics and genomics will reveal the prevalence of TEs in nature and to what extent they have played a role in generating and sustaining new organismal diversity.

Figure 1. A graphical classification of transposable elements (TEs). The left panel shows Class 1 retrotransposons, and the right panel shows Class 2 DNA transposons. The upper panels show three examples of the genetic structure of each of these two classes of elements. The lower panels show the mode of movement (transposition mechanism) of each class. *LTR*: Long Terminal Repeats; *LINE*: Long interspersed nuclear elements; *SINE*: Short interspersed elements.

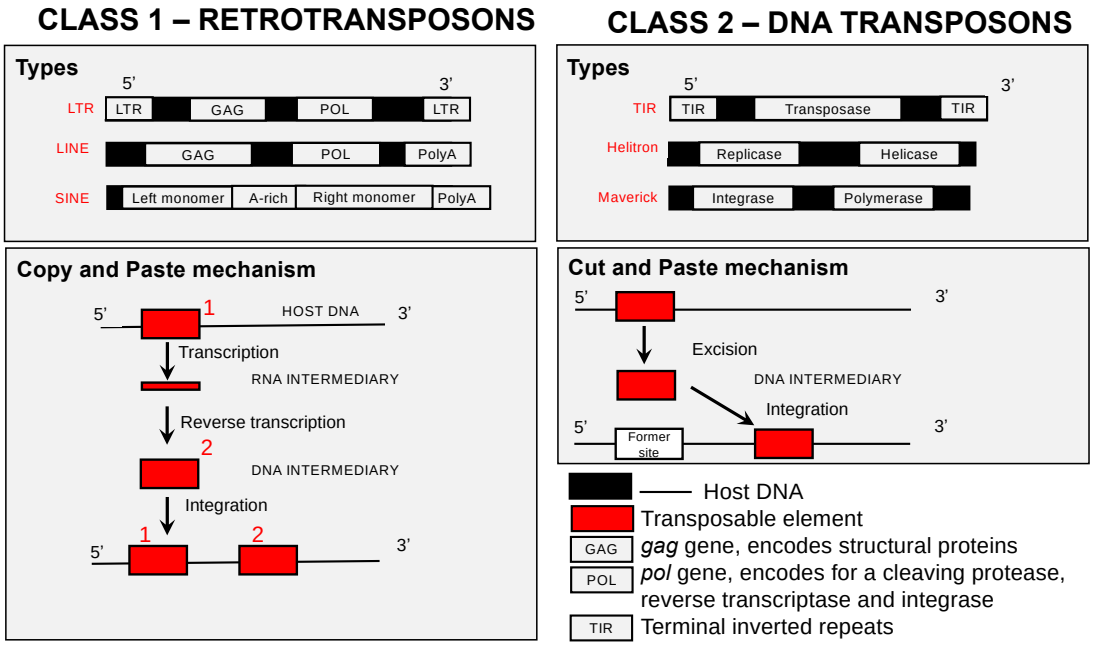


Figure 2. Two possible scenarios that illustrate potential connections between TEs and the likelihood of introgression. Two species are illustrated (blue and red). Stripped bars show chromosomes that contain TEs, while solid bars are chromosomes with no TEs. The left panel (Scenario 1) shows a potential scenario in which TEs facilitate the transfer of a full chromosome. The right panel (Scenario 2) shows a potential scenario in which TEs cannot cross the species boundary and thus chromosomes that harbor them are less likely to be introgressed. For simplicity only one direction of introgression is shown.

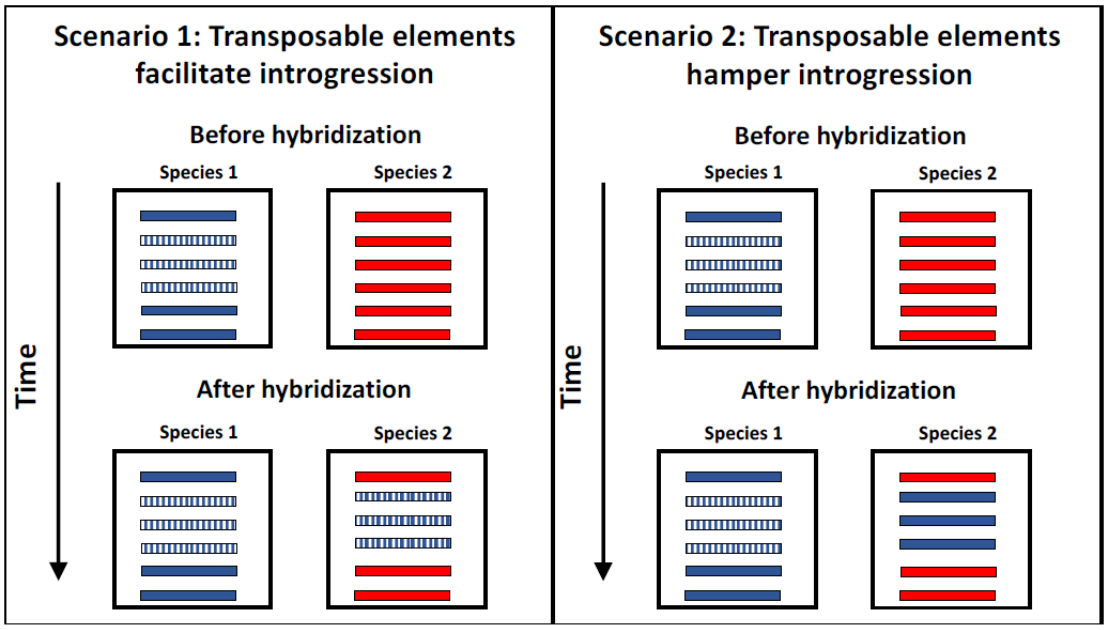


Table 1. A classification of the different types of transposable elements.

Type	Name	Activity	Taxonomic Distribution	Insertion	Function/Pathway Influenced [Reference]
Retro-transposon (class 1)					Replicate through reverse transcription of an mRNA intermediate, the resulting cDNA product integrates
Long-tandem repeats (class 1)					
	BEL/Pao-like elements	non-autonomous	Metazoans	Undescribed	Second most abundant retrotransposon but very little is known. [246,247]
	DIRS1-like retro-transposon	autonomous	Common in decapods, sparse among other Eukaryotes	Preferentially integrates into other DIRS-1 sequences and GTT sequences	Undescribed [248,249]
	Ty1/copia	autonomous	Eukaryotes	Preference towards upstream region of RNA Pol III, near tRNA genes	Mutational agent and can mediate genome rearrangement through recombination [248,249]

	Ty3/gypsy	autonomous	Eukaryotes	Upstream of RNA polymerase III transcription, near tRNA genes	Mutational agent and can mediate genome rearrangement through recombination. [250]
	Ty5	non-autonomous	Fungi	Integrates near areas of silent chromatin at the telomeres and mating loci	An increase in recombination at insertion points [252]
Non-LTR (class 1)					
	Alu	non-autonomous	Primate specific	Fixed at C-terminus of Human HPK1 and throughout genome	Cause insertion mutations, increase recombination, change gene expression through gene conversion [253]
	LINE (long interspersed nuclear elements: *Jockey*, L1, L2, R2)	autonomous	Eukaryotes	R2 inserts into 28S ribosomal DNA genes but has a strong bias against previous R2 insertions.	Encodes proteins responsible for packing of RNA transcript and a polymerase that enables reverse transcription, with an endonuclease subsequently integrating it into the genome. [254]

	Penelope	autonomous	Metazoans, rare in Plants	Insertions of element have been linked to breakpoints in inversions within *D. virilis*	Element that underlies hybrid dysgenesis in *D. virillis*. [176,255]
	RTE (RNA transport element)	non-autonomous	Metazoans	Do not appear to be sequence specific	Upon insertion has been shown to result in target site duplications [256-258]
	SINE (short interspersed nuclear element)	non-autonomous	Plants, metazoans, fungi	Bias against insertion in intronic splice sites and preferentially inserts into the 3′ region of introns	Shown to control mRNA production and repress transcription of protein coding genes [259,260]
	VIPER/Ngaro	autonomous	Metazoans, fungi	Undescribed	Undescribed [266]
Transposon (class 2)					Replicate through a DNA intermediate
	CACTA	autonomous	Plants	Located near centromere	Results in increased methylation and structural changes between genetic orthologs [261]
	Crypton	autonomous	Fungi, arthropods	Unknown	*Crypton*-derived genes function as transcriptional regulators [262]

	Helitron	autonomous	Plants, metazoans, fungi	Preferentially inserts in gene-rich regions	Ability to capture gene sequences, including introns. [263]
	hobo	autonomous	Arthropods	Biased towards areas with high recombination rate	Can mediate recombination and inversions [126,127,156, 175]
	I-element	autonomous	Plants, metazoans	Located near centromere heterochromatin	Transpose in germline at a high rate and are repressed maternally [264,265]
	Mariner/Tc 1	autonomous	All groups	Associated with heterochromatin	Provide a hotspot of recombination in *Drosophila* females [217,266-268]
	Mavericks/ Polinton	autonomous	Eukaryotes , some prokaryotes	Unknown	Retrovirus-like and codes its own DNA polymerase [269,270]
	Mutator	autonomous	Plants	Insertions concentrate in epigenetically marked open chromatin	Insertion sites are correlated with recombination rates [271]
	P-element	autonomous	Plants, metazoans	Insert at random with a preference for 5′ untranslated regions	Underlies hybrid dysgenesis and greatly increases mutation rate [46,167,211]
	PIF-Harbinger	autonomous	Plants	Target site preference for TAA	Insertion into regulatory genes resulted in pigmentation

					changes in maize [272]
	piggyBac	autonomous	Metazoans	Throughout the genome	Acts as an insertional mutagen. [273,274]
	pogo	autonomous	Metazoans	Likely to insert in regions with low denaturation temperature	Often leads to deletions [112,275]
	Rag-like	autonomous	Metazoans	Undescribed	Linked to recombination and affects immune system response [276,277]
	Transib	autonomous	Eukaryotes	Undescribed	May underlie the development of new genes [278,279]

Table 2. A summary of reproductive isolating barriers for which TEs have been invoked as a potential cause. A full description of the involvement of TEs is presented in the text. Stars represent cases that remain suggestive but for which more evidence is required (see text).

Type of Reproductive Isolation	TE-Mediated Phenotype	Examples and References
Premating isolation	Adaptation to new habitats.	Flowering time [67,261] Host specificity [76,77]
	Insertions at loci that control self-compatibility.	Shift of reproductive strategies lead to reproductive isolation [89,280] TE movement can lead to gene movement and aneuploidy in hybrids [85]
	Changes in traits involved in recognition of conspecifics.	Mating song frequency between sibling species [120]
	Changes in genome structure.	TE-induced chromosomal inversions [128,131]
Postzygotic isolation	Hybrid sterility as a result of reactivated transposition.	Hybrid dysgenesis [159,160,175,176]

	Misregulation of TEs leading to hybrid inviability	Overgrown endosperm; abnormal embryo development [179,180,186]

Table 3. Computational methods to detect transposable elements using genomic data.

TE Detection Tool	Year	Language	Reference
MELT	2017	Java	[281]
IT IS	2015	Perl	[282]
Jitterbug	2015	Python	[283]
DD_DETECTION	2015	C++	[284]
TIDAL	2015	Perl, R	[285]
Mobster	2014	Perl	[286]
Tangram	2014	Java	[287]
T-lex2	2014	Perl	[288]
TIF	2014	Perl	[289]
TranspoSeq	2014	Java, R	[290]
TraFiC	2014	Perl	[291]
TIGRA	2014	C++	[292]
TE-Tracker	2014	Perl	[293]
GRIPper	2013	Python	[294]
RelocaTE	2013	Perl	[295]
Tea	2012	R	[296]
ngs_te_mapper	2012	R	[297]
TE-Locate	2012	Java, Perl	[298]
REPET	2011	Python	[40]
VariationHunter	2010	C++, Python	[299]

HYDRA-SV	2010	C++, Python	[300]
MITE-Hunter	2010	Perl	[301]
SeqGrapheR	2010	R	[302]
RISCI	2010	Perl	[303]
MoDIL	2009	Python	[304]
LTRharvest	2008	C	[305]
HelitronFinder	2008	Perl	[306]
TransposonPSI	2008	Perl	[307]

REFERENCES

1. Coyne, J.A.; Orr, H.A. *Speciation*; Oxford University Press, Inc.: Sunderland, MA, USA, 2004.
2. Darwin, C. *On the Origin of Species by Means of Natural Selection*; Murray: London, UK, 1859.
3. Dobzhansky, T. Genetic nature of species differences. *Am. Nat.* **1937**, *71*, 404–420.
4. Coyne, J.A.; Orr, H.A. Patterns of speciation in *Drosophila*. *Evolution (N. Y.)* **1989**, *43*, 362–381.
5. Coyne, J.A.; Orr, H.A. "Patterns of speciation in *Drosophila*" Revisited. *Evolution (N. Y.)* **1997**, *51*, 295–303.
6. Abbott, R.; Albach, D.; Ansell, S.; Arntzen, J.W.; Baird, S.J.; Bierne, N.; Boughman, J.; Brelsford, A.; Buerkle, C.A.; Buggs, R.; et al. Hybridization and speciation. *J. Evol. Biol.* **2013**, *26*, 229–246.
7. Harrison, R.G. The language of speciation. *Evolution* **2012**, *66*, 3643–3657.
8. Liou, L.W.; Price, T.D. Speciation by reinforcement of premating isolation. *Evolution (N. Y.)* **1994**, *48*, 1451–1459.
9. Rundle, H.D.; Nosil, P. Ecological speciation. *Ecol. Lett.* **2005**, *8*, 336–352.
10. Seehausen, O.; Butlin, R.K.; Keller, I.; Wagner, C.E.; Boughman, J.W.; Hohenlohe, P.A.; Peichel, C.L.; Saetre, G.P.; Bank, C.; Brännström, Å.; et al. Genomics and the origin of species. *Nat. Rev. Genet.* **2014**, *15*, 176–192.
11. Maan, M.E.; Seehausen, O. Ecology, sexual selection and speciation. *Ecol. Lett.* **2011**, *14*, 591–602.
12. Bolnick, D.I.; Near, T.J. Tempo of hybrid inviability in centrarchid fishes (Teleostei: Centrarchidae). *Evolution* **2005**, *59*, 1754–1767.
13. Moyle, L.C.; Nakazato, T. Hybrid Incompatibility snowballs between *Solanum* species. *Science* **2010**, *329*, 1521–1523.
14. Matute, D.R.; Butler, I.A.; Turissini, D.A.; Coyne, J.A. A Test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science* **2010**, *329*, 1518–1521.
15. Wang, R.J.; White, M.A.; Payseur, B.A. The pace of hybrid incompatibility evolution in house mice. *Genetics* **2015**, *201*, 229–242.
16. Orr, H.A. The genetic basis of reproductive isolation: Insights from *Drosophila*. *Proc. Natl. Acad. Sci. USA* **2005**, *102* (Suppl. 1), 6522–6526.

17. Widmer, A.; Lexer, C.; Cozzolino, S. Evolution of reproductive isolation in plants. *Heredity (Edinb.)* **2009**, *102*, 31–38.

18. Noor, M.A.F.; Grams, K.L.; Bertucci, L.A.; Reiland, J. Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12084–12088.

19. Kirkpatrick, M.; Barton, N. Chromosome inversions, local adaptation and speciation. *Genetics* **2006**, *173*, 419–434.

20. Hoffmann, A.A.; Rieseberg, L.H. Revisiting the impact of inversions in evolution: From population genetic markers to drivers of adaptive shifts and speciation? *Annu. Rev. Ecol. Evol. Syst.* **2008**, *39*, 21–42.

21. Wicker, T.; Sabot, F.; Hua-Van, A.; Bennetzen, J.L.; Capy, P.; Chalhoub, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O.; et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **2007**, *8*, 973–982.

22. Gregory, T.R.; Nicol, J.A.; Tamm, H.; Kullman, B.; Kullman, K.; Leitch, I.J.; Murray, B.G.; Kapraun, D.F.; Greilhuber, J.; Bennett, M.D. Eukaryotic genome size databases. *Nucleic Acids Res.* **2007**, *35* (Suppl. 1), D332–D338.

23. Bennetzen, J.L.; Wang, H. The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annu. Rev. Plant Biol.* **2014**, *65*, 505–530.

24. Charlesworth, B.; Langley, C.H.; Sniegowski, P.D. Transposable element distributions in *Drosophila*. *Genetics* **1997**, *147*, 1993–1995.

25. Kidwell, M.G. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **2002**, *115*, 49–63.

26. Kidwell, M.G.; Lisch, D. Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 7704–7711.

27. Bennett, E.A.; Coleman, L.E.; Tsui, C.; Pittard, W.S.; Devine, S.E. Natural genetic variation caused by transposable elements in humans. *Genetics* **2004**, *168*, 933–951.

28. Lockton, S.; Gaut, B.S. The evolution of transposable elements in natural populations of self-fertilizing *Arabidopsis thaliana* and its outcrossing relative *Arabidopsis lyrata*. *BMC Evol. Biol.* **2010**, *10*, 10, doi:10.1186/1471-2148-10-10.

29. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA* **1950**, *36*, 344–355.

30. Oliver, K.R.; Greene, W.K. Transposable elements: Powerful facilitators of evolution. *BioEssays* **2009**, *31*, 703–714.

31. Okamoto, H.; Hirochika, H. Silencing of transposable elements in plants. *Trends Plant Sci.* **2001**, *6*, 527–534.

32. Feschotte, C.; Jiang, N.; Wessler, S.R. Plant transposable elements: Where genetics meets genomics. *Nat. Rev. Genet.* **2002**, *3*, 329–341.

33. Woodhouse, M.R.; Freeling, M.; Lisch, D. Initiation, establishment, and maintenance of heritable *MuDR* transposon silencing in maize are mediated by distinct factors. *PLoS Biol.* **2006**, *4*, 1678–1688.

34. Feschotte, C. Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* **2008**, *9*, 397–405.

35. Voytas, D.F.; Cummings, M.P.; Koniczny, A.; Ausubel, F.M.; Rodermel, S.R. *copia*-like retrotransposons are ubiquitous among plants. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7124–7128.

36. Flavell, A.J.; Dunbar, E.; Anderson, R.; Pearce, S.R.; Hartley, R.; Kumar, A. Ty1-copia group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Res.* **1992**, *20*, 3639–3644.

37. Aziz, R.K.; Breitbart, M.; Edwards, R.A. Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Res.* **2010**, *38*, 4207–4217.

38. Van De Lagemaat, L.N.; Landry, J.R.; Mager, D.L.; Medstrand, P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* **2003**, *19*, 530–536.

39. Kapusta, A.; Kronenberg, Z.; Lynch, V.J.; Zhuo, X.; Ramsay, L.; Bourque, G.; Yandell, M.; Feschotte, C. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet.* **2013**, *9*, e1003470.

40. Flutre, T.; Duprat, E.; Feuillet, C.; Quesneville, H. Considering transposable element diversification in de novo annotation approaches. *PLoS ONE* **2011**, *6*, 0016526, doi:10.1371/journal.pone.0016526.

41. Hollister, J.D.; Gaut, B.S. Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Res.* **2009**, *19*, 1419–1428.

42. Arkhipova, I.; Meselson, M. Deleterious transposable elements and the extinction of asexuals. *BioEssays* **2005**, *27*, 76–85.

43. Vinogradov, A.E. Evolution of genome size: Multilevel selection, mutation bias or dynamical chaos? *Curr. Opin. Genet. Dev.* **2004**, *14*, 620–626.

44. Brunet, T.D.; Doolittle, W.F. Multilevel selection theory and the evolutionary functions of transposable elements. *Genome Biol. Evol.* **2015**, *7*, 2445–2457.

45. Fablet, M.; Vieira, C. Evolvability, epigenetics and transposable elements. *Biomol. Concepts* **2011**, *2*, 333–341.

46. Kidwell, M.G.; Kidwell, J.F.; Sved, J.A. Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics* **1977**, *86*, 813–833.

47. Hurst, G.D.D.; Schilthuizen, M. Selfish genetic elements and speciation. *Heredity (Edinb.)* **1998**, *80*, 2–8.

48. Werren, J.H. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (Suppl. 2), 10863–10870.

49. Crespi, B.; Nosil, P. Conflictual speciation: Species formation via genomic conflict. *Trends Ecol. Evol.* **2013**, *28*, 48–57.

50. Ginzburg, L.R.; Bingham, P.M.; Yoo, S. On the theory of speciation induced by transposable elements. *Genetics* **1984**, *107*, 331–341.

51. Zeh, D.W.; Zeh, J.A.; Ishida, Y. Transposable elements and an epigenetic basis for punctuated equilibria. *BioEssays* **2009**, *31*, 715–726.

52. Orr, H.A.; Presgraves, D.C. Speciation by postzygotic isolation: Forces, genes and molecules. *BioEssays* **2000**, *22*, 1085–1094.

53. Stapley, J.; Santure, A.W.; Dennis, S.R. Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Mol. Ecol.* **2015**, *24*, 2241–2252.

54. Casacuberta, E.; González, J. The impact of transposable elements in environmental adaptation. *Mol. Ecol.* **2013**, *22*, 1503–1517.

55. Lisch, D. How important are transposons for plant evolution? *Nat. Rev. Genet.* **2013**, *14*, 49–61.

56. Belyayev, A. Bursts of transposable elements as an evolutionary driving force. *J. Evol. Biol.* **2014**, *27*, 2573–2584.

57. Comai, L.; Madlung, A.; Josefsson, C.; Tyagi, A. Do the different parental "heteromes" cause genomic shock in newly formed allopolyploids? *Philos. Trans. R. Soc. B Biol. Sci.* **2003**, *358*, 1149–1155.

58. Tapia, G.; Verdugo, I.; Yañez, M.; Ahumada, I.; Theoduloz, C.; Cordero, C.; Poblete, F.; González, E.; Ruiz-Lara, S. Involvement of ethylene in stress-induced expression of the, *TLC1.1* retrotransposon from *Lycopersicon chilense* Dun. *Plant Physiol.* **2005**, *138*, 2075–2086.

59. Servedio, M.R.; Van Doorn, G.S.; Kopp, M.; Frame, A.M.; Nosil, P. Magic traits in speciation: "magic" but not rare? *Trends Ecol. Evol.* **2011**, *26*, 389–397.

60. Thibert-Plante, X.; Gavrilets, S. Evolution of mate choice and the so-called magic traits in ecological speciation. *Ecol. Lett.* **2013**, *16*, 1004–1013.

61. Nosil, P.; Vines, T.H.; Funk, D.J. Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* **2005**, *59*, 705–719.

62. Hendry, A.P.; Day, T. Population structure attributable to reproductive time: Isolation by time and adaptation by time. *Mol. Ecol.* **2005**, *14*, 901–916.

63. Ehrlén, J. Selection on flowering time in a life-cycle context. *Oikos* **2015**, *124*, 92–101.

64. Castelletti, S.; Tuberosa, R.; Pindo, M.; Salvi, S. A MITE transposon insertion is associated with differential methylation at the maize flowering time, QTL *Vgt1*. *G3 Genes Genomes Genet.* **2014**, *4*, 805–812.

65. Yang, Q.; Li, Z.; Li, W.; Ku, L.; Wang, C.; Ye, J.; Li, K.; Yang, N.; Li, Y.; Zhong, T.; et al. CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16969–16974.

66. Liu, J.; He, Y.; Amasino, R.; Chen, X. siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in *Arabidopsis*. *Genes Dev.* **2004**, *18*, 2873–2878.

67. Tamaki, S.; Tsuji, H.; Matsumoto, A.; Fujita, A.; Shimatani, Z.; Terada, R.; Sakamoto, T.; Kurata, T.; Shimamoto, K. FT-like proteins induce transposon silencing in the shoot apex during floral induction in rice. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E901–E910.

68. González, J.; Karasov, T.L.; Messer, P.W.; Petrov, D.A. Genome-wide patterns of adaptation to temperate environments associated with transposable elements in *Drosophila*. *PLoS Genet.* **2010**, *6*, e1000905.

69. Kamoun, S. A Catalogue of the effector secretome of plant pathogenic *Oomycetes*. *Annu. Rev. Phytopathol.* **2006**, *44*, 41–60.

70. Schulze-Lefert, P.; Panstruga, R. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends Plant Sci.* **2011**, *16*, 117–125.

71. Restrepo, S.; Tabima, J.F.; Mideros, M.F.; Grünwald, N.J.; Matute, D.R. Speciation in fungal and oomycete plant pathogens. *Annu. Rev. Phytopathol.* **2014**, *52*, 289–316.

72. Judelson, H.S. Genomics of the plant pathogenic oomycete *Phytophthora*: Insights into biology and evolution. *Adv. Genet.* **2007**, *57*, 97–141.

73. Kamoun, S.; Hraber, P.; Sobral, B.; Nuss, D.; Govers, F. Initial assessment of gene diversity for the oomycete pathogen *Phytophthora infestans* based on expressed sequences. *Fungal Genet. Biol.* **1999**, *28*, 94–106.

74. Kamoun, S. Molecular genetics of pathogenic *Oomycetes*. *Eukaryot. Cell* **2003**, *2*, 191–199.

75. Judelson, H.S. Sequence variation and genomic amplification of a family of *Gypsy*-like elements in the oomycete genus *Phytophthora*. *Mol. Biol. Evol.* **2002**, *19*, 1313–1322.

76. Tooley, P.W.; Garfinkel, D.J. Presence of *Ty1-copia* group retrotransposon sequences in the potato late blight pathogen *Phytophthora infestans*. *Mol. Plant Microbe Interact.* **1996**, *9*, 305–309.

77. Whisson, S.C.; Vetukuri, R.R.; Avrova, A.O.; Dixelius, C. Can silencing of transposons contribute to variation in effector gene expression in *Phytophthora infestans* ? *Mob. Genet. Elem.* **2012**, *2*, 110–114.

78. Nielsen, K.; Heitman, J. Sex and virulence of human pathogenic Fungi. *Adv. Genet.* **2007**, *57*, 143–173.

79. Wilson, A.M.; Wilken, P.M.; Van Der Nest, M.A.; Steenkamp, E.T.; Wingfield, M.J.; Wingfield, B.D. Homothallism: An umbrella term for describing diverse sexual behaviours. *IMA Fungus* **2015**, *6*, 207–214.

80. Korf, R.P. The terms homothallism and heterothallism. *Nature* **1952**, *170*, 534–535.

81. Wendt, T.; Canela, M.B.F.; Klein, D.E.; Rios, R.I. Selfing facilitates reproductive isolation among three sympatric species of *Pitcairnia* (Bromeliaceae). *Plant Syst. Evol.* **2002**, *232*, 201–212.

82. Baker, H.G. Reproductive methods as factors in speciation in flowering plants. *Cold Spring Harb. Symp. Quant. Biol.* **1959**, *24*, 177–191.

83. Foxe, J.P.; Slotte, T.; Stahl, E.A.; Neuffer, B.; Hurka, H.; Wright, S.I. Recent speciation associated with the evolution of selfing in *Capsella*. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5241–5245.

84. Gioti, A.; Mushegian, A.A.; Strandberg, R.; Stajich, J.E.; Johannesson, H. Unidirectional evolutionary transitions in fungal mating systems and the role of transposable elements. *Mol. Biol. Evol.* **2012**, *29*, 3215–3226.

85. Laurie, J.D.; Ali, S.; Linning, R.; Mannhaupt, G.; Wong, P.; Güldener, U.; Münsterkötter, M.; Moore, R.; Kahmann, R.; Bakkeren, G.; et al. Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements. *Plant Cell* **2012**, *24*, 1733–1745.

86. Nygren, K.; Strandberg, R.; Wallberg, A.; Nabholz, B.; Gustafsson, T.; García, D.; Cano, J.; Guarro, J.; Johannesson, H. A comprehensive phylogeny of *Neurospora* reveals a link between reproductive mode and molecular evolution in fungi. *Mol. Phylogenet. Evol.* **2011**, *59*, 649–663.

87. Li, W.; Sullivan, T.D.; Walton, E.; Averette, A.F.; Sakthikumar, S.; Cuomo, C.A.; Klein, B.S.; Heitman J Identification of the mating-type (MAT) locus that controls sexual reproduction of *Blastomyces dermatitidis*. *Eukaryot. Cell* **2013**, *12*, 109–117.

88. Rydholm, C.; Dyer, P.S.; Lutzoni, F. DNA sequence characterization and molecular evolution of, *MAT1* and, *MAT2* mating-type loci of the self-compatible ascomycete mold *Neosartorya fischeri*. *Eukaryot. Cell* **2007**, *6*, 868–874.

89. Lengeler, K.B.; Fox, D.S.; Fraser, J.A.; Allen, A.; Forrester, K.; Dietrich, F.S.; Heitman, J. Mating-type locus of *Cryptococcus neoformans*: A step in the evolution of sex chromosomes. *Eukaryot. Cell* **2002**, *1*, 704–718.

90. Desjardins, C.A.; Champion, M.D.; Holder, J.W.; Muszewska, A.; Goldberg, J.; Bailão, A.M.; Brigido, M.M.; da Silva Ferreira, M.E.; Garcia, A.M.; Grynberg, M.; et al. Comparative genomic analysis of human fungal pathogens causing paracoccidioidomycosis. *PLoS Genet.* **2011**, *7*, e1002345.

91. Rusche, L.N.; Rine, J. Switching the mechanism of mating type switching: A domesticated transposase supplants a domesticated homing endonuclease. *Genes Dev.* **2010**, *24*, 10–14.

92. Daboussi, M.J. Fungal transposable elements and genome evolution. *Genetica* **1997**, *100*, 253–260.

93. Moyle, L.C.; Muir, C.D.; Han, M.V.; Hahn, M.W. The contribution of gene movement to the "two rules of speciation." *Evolution* **2010**, *64*, 1541–1557.

94. Laurie, C.C. The weaker sex is heterogametic: 75 years of Haldane's rule. *Genetics* **1997**, *147*, 937–951.

95. Kondo, K.; Yamamoto, M.; Matton, D.P.; Sato, T.; Hirai, M.; Norioka, S.; Hattori, T.; Kowyama, Y. Cultivated tomato has defects in both S-RNase and, HT genes required for stylar function of self-incompatibility. *Plant J.* **2002**, *29*, 627–636.

96. Wright, S.I.; Kalisz, S.; Slotte, T. Evolutionary consequences of self-fertilization in plants. *Proc. R. Soc. B Biol. Sci.* **2013**, *280*, 1–10, doi:10.1098/rspb.2013.0133.

97. Goldberg, E.E.; Kohn, J.R.; Lande, R.; Robertson, K.A.; Smith, S.A.; Igić, B. Species selection maintains self-incompatibility. *Science* **2010**, *330*, 493–495.

98. Pinheiro, F.; Cafasso, D.; Cozzolino, S.; Scopece, G. Transitions between self-compatibility and self-incompatibility and the evolution of reproductive isolation in the large and diverse tropical genus *Dendrobium* (Orchidaceae). *Ann. Bot.* **2015**, *116*, 457–467.

99. Rabosky, D.L.; McCune, A.R. Reinventing species selection with molecular phylogenies. *Trends Ecol. Evol.* **2010**, *25*, 68–74.

100. Rabosky, D.L. Ecological limits and diversification rate: Alternative paradigms to explain the variation in species richness among clades and regions. *Ecol. Lett.* **2009**, *12*, 735–743.

101. Jablonski, D. Species selection: Theory and data. *Annu. Rev. Ecol. Evol. Syst.* **2008**, *39*, 501–524.

102. Griffith, L.C.; Ejima, A. Multimodal sensory integration of courtship stimulating cues in *Drosophila melanogaster*: Contextual effects on chemosensory cues. *Ann. N. Y. Acad. Sci.* **2009**, 394–398, doi:10.1111/j.1749-6632.2009.04367.x.

103. Clowney, E.J.; Iguchi, S.; Bussell, J.J.; Scheer, E.; Ruta, V. Multimodal chemosensory circuits controlling male courtship in *Drosophila*. *Neuron* **2015**, *87*, 1036–1049.

104. Ejima, A.; Griffith, L.C. Courtship initiation is stimulated by acoustic signals in *Drosophila melanogaster*. *PLoS ONE* **2008**, *3*, e0003246.

105. Chung, H.; Loehlin, D.W.; Dufour, H.D.; Vacarro, K.; Millar, J.G.; Carroll, S.B. A single gene affects both ecological divergence and mate choice in *Drosophila*. *Science* **2014**, *343*, 1148–1151.

106. Gershman, S.N.; Toumishey, E.; Rundle, H.D. Time flies: Time of day and social environment affect cuticular hydrocarbon sexual displays in *Drosophila serrata*. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20140821, doi:10.1098/rspb.2014.0821.

107. Dembeck, L.M.; Böröczky, K.; Huang, W.; Schal, C.; Anholt, R.R.; Mackay, T.F. Genetic architecture of natural variation in cuticular hydrocarbon composition in *Drosophila melanogaster*. *eLife* **2015**, *4*, e09861.

108. Jezovit, J.A.; Levine, J.D.; Schneider, J. Phylogeny, environment and sexual communication across the *Drosophila* genus. *J. Exp. Biol.* **2017**, *220*, 42–52.

109. Schwander, T.; Arbuthnott, D.; Gries, R.; Gries, G.; Nosil, P.; Crespi, B.J. Hydrocarbon divergence and reproductive isolation in *Timema* stick insects. *BMC Evol. Biol.* **2013**, *13*, doi:10.1186/1471-2148-13-151.

110. Chung, H.; Carroll, S.B. Wax, sex and the origin of species: Dual roles of insect cuticular hydrocarbons in adaptation and mating. *BioEssays* **2015**, *37*, 822–830.

111. Marcillac, F.; Grosjean, Y.; Ferveur, J.-F. A single mutation alters production and discrimination of *Drosophila* sex pheromones. *Proc. R. Soc. B Biol. Sci.* **2005**, *272*, 303–309.

112. Kaminker, J.S.; Bergman, C.M.; Kronmiller, B.; Carlson, J.; Svirskas, R.; Patel, S.; Frise, E.; Wheeler, D.A.; Lewis, S.E.; Rubin, G.M.; et al. The transposable elements of the *Drosophila melanogaster* euchromatin: A genomics perspective. *Genome Biol.* **2002**, *3*, 1–20.

113. Quesneville, H.; Bergman, C.M.; Andrieu, O.; Autard, D.; Nouaud, D.; Ashburner, M.; Anxolabehere, D. Combined evidence annotation of transposable elements in genome sequences. *PLoS Comput. Biol.* **2005**, *1*, 166–175.

114. von Schilcher, F. The function of pulse song and sine song in the courtship of *Drosophila melanogaster*. *Anim. Behav.* **1976**, *24*, 622–625.

115. Kliman, R.M.; Andolfatto, P.; Coyne, J.A.; Depaulis, F.; Kreitman, M.; Berry, A.J.; McCarter, J.; Wakeley, J.; Hey, J. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* **2000**, *156*, 1913–1931.

116. Garrigan, D.; Kingan, S.B.; Geneva, A.J.; Andolfatto, P.; Clark, A.G.; Thornton, K.R.; Presgraves, D.C. Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. *Genome Res.* **2012**, *22*, 1499–1511.

117. Brand, C.L.; Kingan, S.B.; Wu, L.; Garrigan, D. A selective sweep across species boundaries in *Drosophila*. *Mol. Biol. Evol.* **2013**, *30*, 2177–2186.

118. Gleason, J.M.; Ritchie, M.G. Do Quantitative Trait Loci (QTL) for a courtship song difference between *Drosophila simulans* and *D. sechellia* coincide with candidate genes and intraspecific, QTL? *Genetics* **2004**, *166*, 1303–1311.

119. Peixoto, A.A.; Hall, J.C. Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of *Drosophila*. *Genetics* **1998**, *148*, 827–838.

120. Ding, Y.; Berrocal, A.; Morita, T.; Longden, K.D.; Stern, D.L. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* **2016**, *536*, 329–332.

121. Baack, E.; Melo, M.C.; Rieseberg, L.H.; Ortiz-Barrientos, D. The origins of reproductive isolation in plants. *New Phytol.* **2015**, *207*, 968–984.

122. Ayala, D.; Guerrero, R.F.; Kirkpatrick, M. Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. *Evolution* **2013**, *67*, 946–958.

123. Rieseberg, L.H. Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* **2001**, *16*, 351–358.

124. Lowry, D.B.; Willis, J.H. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **2010**, *8*, e1000500.

125. Li, S.F.; Su, T.; Cheng, G.Q.; Wang, B.X.; Li, X.; Deng, C.L.; Gao, W.J. Chromosome evolution in connection with repetitive sequences and epigenetics in plants. *Genes* **2017**, *8*, 290, doi:10.3390/genes8100290.

126. Ladevèze, V.; Aulard, S.; Chaminade, N.; Périquet, G.; Lemeunier, F. *Hobo* transposons causing chromosomal breakpoints. *Proc. Biol. Sci.* **1998**, *265*, 1157–1159.

127. Lim, J.K. Intrachromosomal rearrangements mediated by hobo transposons in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 9153–9157.

128. Cáceres, M.; Ranz, J.M.; Barbadilla, A.; Long, M.; Ruiz, A. Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **1999**, *285*, 415–418.

129. Soto, I.M.; Soto, E.M.; Carreira, V.P.; Hurtado, J.; Fanara, J.J.; Hasson, E. Geographic patterns of inversion polymorphism in the second chromosome of the cactophilic *Drosophila buzzatii* from Northeastern Argentina. *J. Insect Sci.* **2010**, *10*, 1–11.

130. Fernández Iriarte, P.J.; Norry, F.M.; Hasson, E.R. Chromosomal inversions effect body size and shape in different breeding resources in *Drosophila buzzatii*. *Heredity (Edinb.)* **2003**, *91*, 51–59.

131. Gómez, G.A.; Hasson, E. Transpecific polymorphisms in an inversion linked esterase locus in *Drosophila buzzatii*. *Mol. Biol. Evol.* **2003**, *20*, 410–423.

132. Betran, E.; Santos, M.; Ruiz, A. Antagonistic pleiotropic effect of second-chromosome inversions on body size and early life-history traits in *Drosophila buzzatii*. *Evolution* **1998**, *52*, 144–154.

133. Slotkin, R.K.; Vaughn, M.; Borges, F.; Tanurdžić, M.; Becker, J.D.; Feijó, J.A.; Martienssen, R.A. Epigenetic reprogramming and small, RNA silencing of transposable elements in pollen. *Cell* **2009**, *136*, 461–472.

134. Lippman, Z.; Gendrel, A.V.; Black, M.; Vaughn, M.W.; Dedhia, N.; McCombie, W.R.; Lavine, K.; Mittal, V.; May, B.; Kasschau, K.D.; et al. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **2004**, *430*, 471–476.

135. Gendrel, A.V.; Lippman, Z.; Yordan, C.; Colot, V.; Martienssen, R.A. Dependence of heterochromatic histone H3 methylation patterns on the *Arabidopsis* gene, DDM1. *Science* **2002**, *297*, 1871–1873.

136. Lippman, Z.; May, B.; Yordan, C.; Singer, T.; Martienssen, R. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol.* **2003**, *1*, e0000067.

137. McCormick, S. Control of male gametophyte development. *Plant Cell* **2004**, *16* (Suppl. 1), S142–S153.

138. Kawanabe, T.; Ishikura, S.; Miyaji, N.; Sasaki, T.; Wu, L.M.; Itabashi, E.; Takada, S.; Shimizu, M.; Takasaki-Yasuda, T.; Osabe, K.; et al. Role of DNA methylation in hybrid vigor in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E6704–E6711.

139. Groszmann, M.; Gonzalez-Bayon, R.; Greaves, I.K.; Wang, L.; Huen, A.K.; Peacock, W.J.; Dennis, E.S. Intraspecific *Arabidopsis* hybrids show different patterns of heterosis despite the close relatedness of the parental genomes. *Plant Physiol.* **2014**, *166*, 265–280.

140. Zhang, Q.; Li, Y.; Xu, T.; Srivastava, A.K.; Wang, D.; Zeng, L.; Yang, L.; He, L.; Zhang, H.; Zheng, Z.; et al. The chromatin remodeler, DDM1 promotes hybrid vigor by regulating salicylic acid metabolism. *Cell Discov.* **2016**, *2*, 16027, doi:10.1038/celldisc.2016.27.

141. Rogers, S.M.; Bernatchez, L. The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae) species pairs. *Mol. Biol. Evol.* **2007**, *24*, 1423–1438.

142. Renaut, S.; Bernatchez, L. Transcriptome-wide signature of hybrid breakdown associated with intrinsic reproductive isolation in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Heredity (Edinb.)* **2011**, *106*, 1003–1011.

143. Dion-Côté, A.M.; Renaut, S.; Normandeau, E.; Bernatchez, L. RNA-seq reveals transcriptomic shock involving transposable elements reactivation in hybrids of young lake whitefish species. *Mol. Biol. Evol.* **2014**, *31*, 1188–1199.

144. Metcalfe, C.J.; Bulazel, K.V.; Ferreri, G.C.; Schroeder-Reiter, E.; Wanner, G.; Rens, W.; Obergfell, C.; Eldridge, M.D.; O'Neill, R.J. Genomic instability within centromeres of interspecific marsupial hybrids. *Genetics* **2007**, *177*, 2507–2517.

145. O'Neill, R.J.; O'Neill, M.J.; Graves, J.A. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* **1998**, *393*, 68–72.

146. Brideau, N.J.; Flores, H.A.; Wang, J.; Maheshwari, S.; Wang, X.U.; Barbash, D.A. Two Dobzhansky-Muller Genes interact to cause hybrid lethality in *Drosophila*. *Science* **2006**, *314*, 1292–1295.

147. Phadnis, N.; Baker, E.P.; Cooper, J.C.; Frizzell, K.A.; Hsieh, E.; De La Cruz, A.F.; Shendure, J.; Kitzman, J.O.; Malik, H.S. An essential cell cycle regulation gene causes hybrid inviability in *Drosophila*. *Science* **2015**, *350*, 1552–1555.

148. Cooper, J.C.; Phadnis, N. A genomic approach to identify hybrid incompatibility genes. *Fly (Austin)* **2016**, *10*, 142–148.

149. Satyaki, P.R.; Cuykendall, T.N.; Wei, K.H.; Brideau, N.J.; Kwak, H.; Aruna, S.; Ferree, P.M.; Ji, S.; Barbash, D.A. The *Hmr* and *Lhr* hybrid incompatibility genes suppress a broad range of heterochromatic repeats. *PLoS Genet.* **2014**, *10*, e1004240.

150. Di Giacomo, M.; Comazzetto, S.; Saini, H.; De Fazio, S.; Carrieri, C.; Morgan, M.; Vasiliauskaite, L.; Benes, V.; Enright, A.J.; O'Carroll, D. Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. *Mol. Cell* **2013**, *50*, 601–608.

151. Kelleher, E.S.; Edelman, N.B.; Barbash, D.A. *Drosophila* Interspecific hybrids phenocopy piRNA-pathway mutants. *PLoS Biol.* **2012**, *10*, e1001428.

152. Shpiz, S.; Ryazansky, S.; Olovnikov, I.; Abramov, Y.; Kalmykova, A. Euchromatic transposon insertions trigger production of novel Pi- and Endo-siRNAs at the target sites in the *Drosophila* germline. *PLoS Genet.* **2014**, *10*, e1004138.

153. Senti, K.A.; Jurczak, D.; Sachidanandam, R.; Brennecke, J. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. *Genes Dev.* **2015**, *29*, 1747–1762.

154. Rozhkov, N.V.; Hammell, M.; Hannon, G.J. Multiple roles for Piwi in silencing *Drosophila* transposons. *Genes Dev.* **2013**, *27*, 400–412.

155. Erwin, A.A.; Galdos, M.A.; Wickersheim, M.L.; Harrison, C.C.; Marr, K.D.; Colicchio, J.M.; Blumenstiel, J.P. piRNAs are associated with diverse transgenerational effects on gene and transposon expression in a hybrid dysgenic syndrome of *D. virilis*. *PLoS Genet.* **2015**, *11*, e1005332.

156. Streck, R.D.; Macgaffey, J.E.; Beckendorf, S.K. The structure of hobo transposable elements and their insertion sites. *Embo J.* **1986**, *5*, 3615–3623.

157. Bucheton, A.; Vaury, C.; Chaboissier, M.C.; Abad, P.; Pélisson, A.; Simonelig, M. I elements and the *Drosophila* genome. *Genetica* **1992**, *86*, 175–190.

158. David, J.R.; Capy, P. Genetic variation of *Drosophila melanogaster* natural populations. *Trends Genet.* **1988**, *4*, 106–111.

159. Bregliano, J.C.; Picard, G.; Bucheton, A.; Pelisson, A.; Lavige, J.M.; L'Héritier, P. Hybrid dysgenesis in *Drosophila melanogaster*. *Science* **1980**, *207*, 606–611.

160. Hill, T.; Schlötterer, C.; Betancourt, A.J. Hybrid dysgenesis in *Drosophila simulans* associated with a rapid invasion of the P-Element. *PLoS Genet.* **2016**, *12*, e1005920.

161. Dorogova, N.V.; Bolobolova, E.U.; Zakharenko, L.P. Cellular aspects of gonadal atrophy in *Drosophila* P-M hybrid dysgenesis. *Dev. Biol.* **2017**, *424*, 105–112.

162. Kelleher, E.S. Reexamining the *P*-element invasion of *Drosophila melanogaster* through the lens of piRNA silencing. *Genetics* **2016**, *203*, 1513–1531.

163. Kelleher, E.S.; Barbash, D.A. Analysis of piRNA-mediated silencing of active TEs in Drosophila melanogaster suggests limits on the evolution of host genome defense. *Mol. Biol. Evol.* **2013**, *30*, 1816–1829.

164. Engels, W.R. Hybrid dysgenesis in *Drosophila melanogaster*: Rules of inheritance of female sterility. *Genet. Res.* **1979**, *33*, 219–236.

165. Kofler, R.; Hill, T.; Nolte, V.; Betancourt, A.J.; Schlötterer, C. The recent invasion of natural *Drosophila simulans* populations by the *P*-element. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6659–6663.

166. Anxolabéhère, D.; Kidwell, M.G.; Periquet, G. Molecular characteristics of diverse populations are consistent with the hypothesis of a recent invasion of *Drosophila melanogaster* by mobile P elements. *Mol. Biol. Evol.* **1988**, *5*, 252–269.

167. Clark, J.B.; Altheide, T.K.; Schlosser, M.J.; Kidwell, M.G. Molecular evolution of P transposable elements in the genus *Drosophila*. I. The *saltans* and *willistoni* species groups. *Mol. Biol. Evol.* **1995**, *12*, 902–913.

168. Regner, L.P.; Pereira, M.S.O.; Alonso, C.E.V.; Abdelhay, E.; Valente, V.L. Genomic distribution of P elements in *Drosophila willistoni* and a search for their relationship with chromosomal inversions. *J. Hered.* **1996**, *87*, 191–198.

169. Houck, M.A.; Clark, J.B.; Peterson, K.R.; Kidwell, M.G. Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science* **1991**, *253*, 1125–1128.

170. Vagin, V.V.; Sigova, A.; Li, C.; Seitz, H.; Gvozdev, V.; Zamore, P.D. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **2006**, *313*, 320–324.

171. Brower-Toland, B.; Findley, S.D.; Jiang, L.; Liu, L.; Yin, H.; Dus, M.; Zhou, P.; Elgin, S.C.; Lin, H. *Drosophila*, PIWI associates with chromatin and interacts directly with, HP1a. *Genes Dev.* **2007**, *21*, 2300–2311.

172. Aravin, A.A.; Hannon, G.J.; Brennecke, J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **2007**, *318*, 761–764.

173. Yin, H.; Lin, H. An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*. *Nature* **2007**, *450*, 304–308.

174. Lewis, S.H.; Quarles, K.A.; Yang, Y.; Tanguy, M.; Frézal, L.; Smith, S.A.; Sharma, P.P.; Cordaux, R.; Gilbert, C.; Giraud, I.; et al. Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat. Ecol. Evol*. **2017**, doi:10.1038/s41559-017-0403-4.

175. Srivastav, S.P.; Kelleher, E.S. Paternal induction of hybrid dysgenesis in *Drosophila melanogaster* is weakly correlated with both *p*-element and *hobo* element dosage. *G3 Genes Genomes Genet.* **2017**, *7*, 1487–1497.

176. Evgen'ev, M.B.; Zelentsova, H.; Shostak, N.; Kozitsina, M.; Barskyi, V.; Lankenau, D.H.; Corces, V.G. *Penelope*, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 196–201.

177. Petrov, D.A.; Schutzman, J.L.; Hartl, D.L.; Lozovskaya, E.R. Diverse transposable elements are mobilized in hybrid dysgenesis in *Drosophila virilis*. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8050–8054.

178. Luo, S.; Lu, J. Silencing of transposable elements by piRNAs in *Drosophila*: An evolutionary perspective. *Genom. Proteom. Bioinf.* **2017**, *15*, 164–176.

179. Oliver, K.R.; McComb, J.A.; Greene, W.K. Transposable elements: Powerful contributors to angiosperm evolution and diversity. *Genome Biol. Evol.* **2013**, *5*, 1886–1901.

180. Kinoshita, T. Reproductive barrier and genomic imprinting in the endosperm of flowering plants. *Genes Genet. Syst.* **2007**, *82*, 177–186.

181. Josefsson, C.; Dilkes, B.; Comai, L. Parent-dependent loss of gene silencing during interspecies hybridization. *Curr. Biol.* **2006**, *16*, 1322–1328.

182. Guo, M.; Rupe, M.A.; Yang, X.; Crasta, O.; Zinselmeier, C.; Smith, O.S.; Bowen, B. Genome-wide transcript analysis of maize hybrids: Allelic additive gene expression and yield heterosis. *Theor. Appl. Genet.* **2006**, *113*, 831–845.

183. Pélissier, T.; Tutois, S.; Tourmente, S.; Deragon, J.M.; Picard, G. DNA regions flanking the major *Arabidopsis thaliana* satellite are principally enriched in *Athila* retroelement sequences. *Genetica* **1996**, *97*, 141–151.

184. Pelissier, T.; Tutois, S.; Deragon, J.M.; Tourmente, S.; Genestier, S.; Picard, G. *Athila*, a new retroelement from *Arabidopsis thaliana*. *Plant Mol. Biol.* **1995**, *29*, 441–452.

185. Gehring, M.; Bubb, K.L.; Henikoff, S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* **2009**, *324*, 1447–1451.

186. Waters, A.J.; Makarevitch, I.; Eichten, S.R.; Swanson-Wagner, R.A.; Yeh, C.T.; Xu, W.; Schnable, P.S.; Vaughn, M.W.; Gehring, M.; Springer, N.M. Parent-of-origin effects on gene expression and DNA methylation in the maize endosperm. *Plant Cell* **2011**, *23*, 4221–4233.

187. Kashkush, K.; Feldman, M.; Levy, A.A. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* **2003**, *33*, 102–106.

188. Haun, W.J.; Danilevskaya, O.N.; Meeley, R.B.; Springer, N.M. Disruption of imprinting by *Mutator* transposon insertions in the 5- proximal regions of the *Zea mays Mez1* locus. *Genetics* **2009**, *181*, 1229–1237.

189. Parisod, C.; Salmon, A.; Zerjal, T.; Tenaillon, M.; Grandbastien, M.A.; Ainouche, M. Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in *Spartina*. *New Phytol.* **2009**, *184*, 1003–1015.

190. Castillo, D.M.; Moyle, L.C. Evolutionary implications of mechanistic models of TE-mediated hybrid incompatibility. *Int. J. Evol. Biol.* **2012**, *2012*, 698198, doi:10.1155/2012/698198.

191. Ellstrand, N.C. Is gene flow the most important evolutionary force in plants? *Am. J. Bot.* **2014**, *101*, 737–753.

192. Soucy, S.M.; Huang, J.; Gogarten, J.P. Horizontal gene transfer: Building the web of life. *Nat. Rev. Genet.* **2015**, *16*, 472–482.

193. De la Cruz, F.; Davies, J. Horizontal gene transfer and the origin of species: Lessons from bacteria. *Trends Microbiol.* **2000**, *8*, 128–133.

194. Thomas, C.M.; Nielsen, K.M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* **2005**, *3*, 711–721.

195. Acuña, R.; Padilla, B.E.; Flórez-Ramos, C.P.; Rubio, J.D.; Herrera, J.C.; Benavides, P.; Lee, S.J.; Yeats, T.H.; Egan, A.N.; Doyle, J.J.; et al. Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4197–4202.

196. Pauchet, Y.; Heckel, D.G. The genome of the mustard leaf beetle encodes two active xylanases originally acquired from bacteria through horizontal gene transfer. *Proc. Biol. Sci.* **2013**, *280*, 20131021, doi:10.1098/rspb.2013.1021.

197. Peccoud, J.; Loiseau, V.; Cordaux, R.; Gilbert, C. Massive horizontal transfer of transposable elements in insects. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 4721–4726.

198. El Baidouri, M.; Carpentier, M.C.; Cooke, R.; Gao, D.; Lasserre, E.; Llauro, C.; Mirouze, M.; Picault, N.; Jackson, S.A.; Panaud, O. Widespread and frequent horizontal transfers of transposable elements in plants. *Genome Res.* **2014**, *24*, 831–838.

199. Diao, X.; Freeling, M.; Lisch, D. Horizontal transfer of a plant transposon. *PLoS Biol.* **2006**, *4*, 119–127.

200. Roulin, A.; Piegu, B.; Fortune, P.M.; Sabot, F.; D'hont, A.; Manicacci, D.; Panaud, O. Whole genome surveys of rice, maize and sorghum reveal multiple horizontal transfers of the LTR-retrotransposon *Route66* in Poaceae. *BMC Evol. Biol.* **2009**, *9*, doi:10.1186/1471-2148-9-58.

201. Woodrow, P.; Pontecorvo, G.; Ciarmiello, L.F. Isolation of *Ty1-copia* retrotransposon in myrtle genome and development of S-SAP molecular marker. *Mol. Biol. Rep.* **2012**, *39*, 3409–3418.

202. Wu, Y.; Jiang, T.; Sun, Y.; Wang, Z.; Guo, G.; Sun, S.; Wang, J.; Li, N.; Wang, Z.; Zhang, D.; et al. Mobilization of diverse transposable elements in rice induced by alien pollination without entailing genetic introgression. *Plant Mol. Biol. Rep.* **2015**, *33*, 1181–1191.

203. Tamura, K.; Subramanian, S.; Kumar, S. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol. Biol. Evol.* **2004**, *21*, 36–44.

204. Vershinin, A.V.; Allnutt, T.R.; Knox, M.R.; Ambrose, M.J.; Ellis, T.H. Transposable elements reveal the impact of introgression, rather than transposition, in *Pisum* diversity, evolution, and domestication. *Mol. Biol. Evol.* **2003**, *20*, 2067–2075.

205. Schaack, S.; Gilbert, C.; Feschotte, C. Promiscuous DNA: Horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.* **2010**, *25*, 537–546.

206. McClintock, B. The significance of responses of the genome to challenge. *Science* **1984**, *226*, 792–801.

207. Wang, Z.H.; Zhang, D.; Bai, Y.; Zhang, Y.H.; Liu, Y.; Wu, Y.; Lin, X.Y.; Wen, J.W.; Xu, C.M.; Li, L.F.; et al. Genomewide variation in an introgression line of Rice-*Zizania* revealed by whole-genome re-sequencing. *PLoS ONE* **2013**, *8*, e0074479.

208. Shan, X.; Liu, Z.; Dong, Z.; Wang, Y.; Chen, Y.; Lin, X.; Long, L.; Han, F.; Dong, Y.; Liu, B. Mobilization of the active MITE transposons *mPing* and *Pong* in rice by introgression from wild rice (*Zizania latifolia* Griseb.). *Mol. Biol. Evol.* **2005**, *22*, 976–990.

209. Rosenzweig, B.K.; Pease, J.B.; Besansky, N.J.; Hahn, M.W. Powerful methods for detecting introgressed regions from population genomic data. *Mol. Ecol.* **2016**, *25*, 2387–2397.

210. De Lucca, M., Jr.; Carareto, C.M.A.; Ceron, C.R. Distribution of the *Bari-I* transposable element in stable hybrid strains between *Drosophila melanogaster* and *Drosophila simulans* and in Brazilian populations of these species. *Genet. Mol. Biol.* **2007**, *30*, 676–680.

211. Haring, E.; Hagemann, S.; Pinsker, W. Different evolutionary behaviour of P element subfamilies: M-type and O-type elements in *Drosophila bifasciata* and *D. imaii*. *Gene* **1995**, *163*, 197–202.

212. Lachaise, D.; Cariou, M.L.; David, J.R.; Lemeunier, F.; Tsacas, L.; Ashburner, M. Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evol. Biol.* **1988**, *22*, 159–225.

213. Bock, I.R. The *bipectinata* complex: A study in interspecific hybridization in the genus *Drosophila* (Insecta: Diptera). *Aust. J. Biol. Sci.* **1978**, *31*, 197–208.

214. de Campos Bicudo, H.E.M. Reproductive isolation in the *saltans* group of *Drosophila*. I. The *saltans* subgroup. *Genetica* **1973**, *44*, 313–329.

215. de Campos Bicudo, H.E. Chromosomal polymorphism in the *saltans* group of *Drosophila* I. The *saltans* subgroup. *Genetica* **1973**, *44*, 520–552.

216. Guerreiro, M.P. Interspecific hybridization as a genomic stressor inducing mobilization of transposable elements in *Drosophila*. *Mob. Genet. Elem.* **2014**, *4*, e88992.

217. Brunet, F.; Godin, F.; David, J.R.; Capy, P. The *mariner* transposable element in the drosophilidae family. *Heredity (Edinb.)* **1994**, *73*, 377–385.

218. Silva, J.C.; Kidwell, M.G. Horizontal transfer and selection in the evolution of P elements. *Mol. Biol. Evol.* **2000**, *17*, 1542–1557.

219. De Almeida, L.M.; Carareto, C.M.A. Multiple events of horizontal transfer of the *Minos* transposable element between *Drosophila* species. *Mol. Phylogenet. Evol.* **2005**, *35*, 583–594.

220. Qian, Y.; Cheng, X.; Liu, Y.; Jiang, H.; Zhu, S.; Cheng, B. Reactivation of a silenced minimal Mutator transposable element system following low-energy nitrogen ion implantation in maize. *Plant Cell Rep.* **2010**, *29*, 1365–1376.

221. Rettelbach, A.; Servedio, M.R.; Hermisson, J. Speciation in peripheral populations: Effects of drift load and mating systems. *J. Evol. Biol.* **2016**, *29*, 1073–1090.

222. Le Rouzic, A.; Boutin, T.S.; Capy, P. Long-term evolution of transposable elements. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19375–19380.

223. Petrov, D.A. Mutational equilibrium model of genome size evolution. *Theor. Popul. Biol.* **2002**, *61*, 533–546.

224. Oliver, M.J.; Petrov, D.; Ackerly, D.; Falkowski, P.; Schofield, O.M. The mode and tempo of genome size evolution in eukaryotes. *Genome Res.* **2007**, *17*, 594–601.

225. Finnegan, D.J. Eukaryotic transposable elements and genome evolution. *Trends Genet.* **1989**, *5*, 103–107.

226. Leitch, I.J.; Kahandawala, I.; Suda, J.; Hanson, L.; Ingrouille, M.J.; Chase, M.W.; Fay, M.F. Genome size diversity in orchids: Consequences and evolution. *Ann. Bot.* **2009**, *104*, 469–481.

227. Kraaijeveld, K. Genome size and species diversification. *Evol. Biol.* **2010**, *37*, 227–233.

228. Burleigh, J.G.; Barbazuk, W.B.; Davis, J.M.; Morse, A.M.; Soltis, P.S. Exploring diversification and genome size evolution in extant gymnosperms through phylogenetic synthesis. *J. Bot.* **2012**, *2012*, 1–6, doi:10.1155/2012/292857.

229. Knight, C.A.; Molinari, N.A.; Petrov, D.A. The large genome constraint hypothesis: Evolution, ecology and phenotype. *Ann. Bot.* **2005**, *95*, 177–190.

230. Alfaro, M.E.; Santini, F.; Brock, C.; Alamillo, H.; Dornburg, A.; Rabosky, D.L.; Carnevale, G.; Harmon, L.J. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13410–13414.

231. Santini, F.; Harmon, L.J.; Carnevale, G.; Alfaro, M.E. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evol. Biol.* **2009**, *9*, doi:10.1186/1471-2148-9-194.

232. Amores, A. Zebrafish *hox* clusters and vertebrate genome evolution. *Science* **1998**, *282*, 1711–1714.

233. Tank, D.C.; Eastman, J.M.; Pennell, M.W.; Soltis, P.S.; Soltis, D.E.; Hinchliff, C.E.; Brown, J.W.; Sessa, E.B.; Harmon, L.J. Nested radiations and the pulse of angiosperm diversification: Increased diversification rates often follow whole genome duplications. *New Phytol.* **2015**, *207*, 454–467.

234. Flagel, L.E.; Wendel, J.F. Gene duplication and evolutionary novelty in plants. *New Phytol.* **2009**, *183*, 557–564.

235. Voldoire, E.; Brunet, F.; Naville, M.; Volff, J.N.; Galiana, D. Expansion by whole genome duplication and evolution of the *sox* gene family in teleost fish. *PLoS ONE* **2017**, *12*, e0180936.

236. Pasquier, J.; Braasch, I.; Batzel, P.; Cabau, C.; Montfort, J.; Nguyen, T.; Jouanno, E.; Berthelot, C.; Klopp, C.; Journot, L.; et al. Evolution of gene expression after whole-genome duplication: New insights from the spotted gar genome. *J. Exp. Zool. Part B Mol. Dev. Evol.* **2017**, *328*, 709–721.

237. Hermansen, R.A.; Hvidsten, T.R.; Sandve, S.R.; Liberles, D.A. Extracting functional trends from whole genome duplication events using comparative genomics. *Biol. Proced. Online* **2016**, *18*, doi:10.1186/s12575-016-0041-2.

238. Van De Peer, Y.; Maere, S.; Meyer, A. The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* **2009**, *10*, 725–732.

239. Puttick, M.N.; Clark, J.; Donoghue, P.C.J. Size is not everything: Rates of genome size evolution, not *C*-value, correlate with speciation in angiosperms. *Proc. R. Soc. B Biol. Sci.* **2015**, *282*, 20152289, doi:10.1098/rspb.2015.2289.

240. Soltis, P.S.; Soltis, D.E. Ancient WGD events as drivers of key innovations in angiosperms. *Curr. Opin. Plant Biol.* **2016**, *30*, 159–165.

241. Seehausen, O. African cichlid fish: A model system in adaptive radiation research. *Proc. Biol. Sci.* **2006**, *273*, 1987–1998.

242. Uyeda, J.C.; Zenil-Ferguson, R.; Pennell, M.W. Rethinking phylogenetic comparative methods. *bioRxiv* **2017**, doi:10.1101/222729.

243. Hey, J. Speciation via hybrid dysgenesis: Negative evidence from the *Drosophila affinis* subgroup. *Genetica* **1988**, *78*, 97–103.

244. Hudson, E.J.; Price, T.D. Pervasive reinforcement and the role of sexual selection in biological speciation. *J. Hered.* **2014**, *105*, 821–833.

245. Trizzino, M.; Park, Y.; Holsbach-Beltrame, M.; Aracena, K.; Mika, K.; Caliskan, M.; Perry, G.H.; Lynch, V.J.; Brown, C.D. Transposable elements are the primary source of novelty in primate gene regulation. *Genome Res.* **2017**, *27*, 1623–1633.

246. Copeland, C.S.; Mann, V.H.; Morales, M.E.; Kalinna, B.H.; Brindley, P.J. The *Sinbad* retrotransposon from the genome of the human blood fluke, *Schistosoma mansoni*, and the distribution of related *Pao*-like elements. *BMC Evol. Biol.* **2005**, *5*, doi:10.1186/1471-2148-5-20.

247. De La Chaux, N.; Wagner, A. *BEL/Pao* retrotransposons in metazoan genomes. *BMC Evol. Biol.* **2011**, *11*, doi:10.1186/1471-2148-11-154.

248. Piednoël, M.; Gonçalves, I.R.; Higuet, D.; Bonnivard, E. Eukaryote DIRS1-like retrotransposons: An overview. *BMC Genom.* **2011**, *12*, 621, doi:0.1186/1471-2164-12-621.

249. Cappello, J.; Cohen, S.M.; Lodish, H.F. Dictyostelium transposable element DIRS-1 preferentially inserts into, DIRS-1 sequences. *Mol. Cell. Biol.* **1984**, *4*, 2207–2213.

250. Sandmeyer, S.B. Yeast retrotransposons. *Curr. Opin. Genet. Dev.* **1992**, *2*, 705–711.

251. Devine, S.E.; Boeke, J.D. Integration of the yeast retrotransposon *Ty1* is targeted to regions upstream of genes transcribed by RNA polymerase III. *Genes Dev.* **1996**, *10*, 620–633.

252. Dai, J.; Xie, W.; Brady, T.L.; Gao, J.; Voytas, D.F. Phosphorylation regulates integration of the yeast ty5 retrotransposon into heterochromatin. *Mol. Cell* **2007**, *27*, 289–299.

253. Nekrutenko, A.; Li, W.H. Transposable elements are found in a large number of human protein-coding genes. *Trends Genet.* **2001**, *17*, 619–621.

254. Chaboissier, M.C.; Bucheton, A.; Finnegan, D.J. Copy number control of a transposable element, the I factor, a LINE-like element in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11781–11785.

255. Koyama, T.; Kondo, H.; Aoki, T.; Hirono, I. Identification of two *Penelope*-like elements with different structures and chromosome localization in kuruma shrimp genome. *Mar. Biotechnol.* **2013**, *15*, 115–123.

256. Youngman, S.; Van Luenen, H.G.; Plasterk, R.H. *Rte-1*, a retrotransposon-like element in *Caenorhabditis elegans*. *FEBS Lett.* **1996**, *380*, 1–7.

257. Tay, W.T.; Behere, G.T.; Batterham, P.; Heckel, D.G. Generation of microsatellite repeat families by RTE retrotransposons in lepidopteran genomes. *BMC Evol Biol* 2010, 10, doi:10.1186/1471-2148-10-144.

258. Župunski, V.; Gubenšek, F.; Kordiš, D. Evolutionary dynamics and evolutionary history in the RTE clade of non-LTR retrotransposons. *Mol. Biol. Evol.* **2001**, *18*, 1849–1863.

259. Ponicsan, S.L.; Kugel, J.F.; Goodrich, J.A. Genomic gems: SINE RNAs regulate mRNA production. *Curr. Opin. Genet. Dev.* **2010**, *20*, 149–155.

260. Majewski, J.; Ott, J. Distribution and characterization of regulatory elements in the human genome. *Genome Res.* **2002**, *12*, 1827–1836.

261. Kato, M.; Takashima, K.; Kakutani, T. Epigenetic control of CACTA transposon mobility in *Arabidopsis thaliana*. *Genetics* **2004**, *168*, 961–969.

262. Kojima, K.K.; Jurka, J. Crypton transposons: Identification of new diverse families and ancient domestication events. *Mob. DNA* **2011**, *2*, 12, doi:10.1186/1759-8753-2-12.

263. Lai, J.; Li, Y.; Messing, J.; Dooner, H.K. Gene movement by *Helitron* transposons contributes to the haplotype variability of maize. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9068–9073.

264. Malinsky, S.; Bucheton, A.; Busseau, I. New insights on homology-dependent silencing of I factor activity by transgenes containing ORF1 in *Drosophila melanogaster*. *Genetics* **2000**, *156*, 1147–1155.

265. Jensen, S.; Gassama, M.P.; Dramard, X.; Heidmann, T. Regulation of I-transposon activity in *Drosophila*: Evidence for cosuppression of nonhomologous transgenes and possible role of ancestral I-related pericentromeric elements. *Genetics* **2002**, *162*, 1197–1209.

266. Hartl, D.L. Discovery of the transposable element *mariner*. *Genetics* **2001**, *157*, 471–476.

267. Robertson, H.M. The mariner transposable element is widespread in insects. *Nature* **1993**, *362*, 241–245.

268. Hartl, D.L.; Lohe, A.R.; Lozovskaya, E.R. Regulation of the transposable element *mariner*. *Genetica* **1997**, *100*, 177–184.

269. Nakayashiki, H. The Trickster in the genome: Contribution and control of transposable elements. *Genes Cells* **2011**, *16*, 827–841.

270. Pritham, E.J.; Putliwala, T.; Feschotte, C. *Mavericks*, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene* **2007**, *390*, 3–17.

271. Liu, S.; Yeh, C.T.; Ji, T.; Ying, K.; Wu, H.; Tang, H.M.; Fu, Y.; Nettleton, D.; Schnable, P.S. *Mu* transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome. *PLoS Genet.* **2009**, *5*, e1000733.

272. Nouroz, F.; Noreen, S.; Heslop-Harrison, J.S. Characterization and diversity of novel PIF/Harbinger DNA transposons in *Brassica* genomes. *Pak. J. Bot.* **2016**, *48*, 167–178.

273. Liang, Q.; Kong, J.; Stalker, J.; Bradley, A. Chromosomal mobilization and reintegration of *Sleeping Beauty* and *PiggyBac* transposons. *Genesis* **2009**, *47*, 404–408.

274. Nishizawa-Yokoi, A.; Endo, M.; Osakabe, K.; Saika, H.; Toki, S. Precise marker excision system using an animal-derived *piggyBac* transposon in plants. *Plant J.* **2014**, *77*, 454–463.

275. Tudor, M.; Lobocka, M.; Goodell, M.; Pettitt, J.; O'Hare, K. The *pogo* transposable element family of *Drosophila melanogaster*. *MGG Mol. Gen. Genet.* **1992**, *232*, 126–134.

276. Hencken, C.G.; Li, X.; Craig, N.L. Functional characterization of an active *Rag*-like transposase. *Nat. Struct. Mol. Biol.* **2012**, *19*, 834–836.

277. Muotri, A.R.; Marchetto, M.C.N.; Coufal, N.G.; Gage, F.H. The necessary junk: New functions for transposable elements. *Hum. Mol. Genet.* **2007**, *16*, R159–R167.

278. Kapitonov, V.V.; Jurka, J. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol.* **2005**, *3*, 0998–1011.

279. Volff, J.N. Turning junk into gold: Domestication of transposable elements and the creation of new genes in eukaryotes. *BioEssays* **2006**, *28*, 913–922.

280. Corcoran, P.; Anderson, J.L.; Jacobson, D.J.; Sun, Y.; Ni, P.; Lascoux, M.; Johannesson, H. Introgression maintains the genetic integrity of the mating-type determining chromosome of the fungus *Neurospora tetrasperma*. *Genome Res.* **2016**, *26*, 486–498.

281. Gardner, E.J.; Lam, V.K.; Harris, D.N.; Chuang, N.T.; Scott, E.C.; Pittard, W.S.; Mills, R.E.; Devine, S.E.; 1000 Genomes Project Consortium; Devine, S.E. The Mobile Element Locator Tool (MELT): Population-scale mobile element discovery and biology. *Genome Res.* **2017**, *27*, 1916–1929.

282. Zhuang, J.; Wang, J.; Theurkauf, W.; Weng, Z. TEMP: A computational method for analyzing transposable element polymorphism in populations. *Nucleic Acids Res.* **2014**, *42*, 6826–6838.

283. Hénaff, E.; Zapata, L.; Casacuberta, J.M.; Ossowski, S. Jitterbug: Somatic and germline transposon insertion detection at single-nucleotide resolution. *BMC Genom.* **2015**, *16*, doi:10.1186/s12864-015-1975-5.

284. Kroon, M.; Lameijer, E.W.; Lakenberg, N.; Hehir-Kwa, J.Y.; Thung, D.T.; Slagboom, P.E.; Kok, J.N.; Ye, K. Detecting dispersed duplications in high-throughput sequencing data using a database-free approach. *Bioinformatics* **2015**, *32*, 505–510.

285. Rahman, R.; Chirn, G.W.; Kanodia, A.; Sytnikova, Y.A.; Brembs, B.; Bergman, C.M.; Lau, N.C. Unique transposon landscapes are pervasive across *Drosophila melanogaster* genomes. *Nucleic Acids Res.* **2015**, *43*, 10655–10672.

286. Thung, D.T.; de Ligt, J.; Vissers, L.E.; Steehouwer, M.; Kroon, M.; de Vries, P.; Slagboom, E.P.; Ye, K.; Veltman, J.A.; Hehir-Kwa, J.Y. Mobster: Accurate detection of mobile element insertions in next generation sequencing data. *Genome Biol.* **2014**, *15*, 488, doi:10.1186/s13059-014-0488-x.

287. Wu, J.; Lee, W.P.; Ward, A.; Walker, J.A.; Konkel, M.K.; Batzer, M.A.; Marth, G.T. Tangram: A comprehensive toolbox for mobile element insertion detection. *BMC Genom.* **2014**, *15*, doi:10.1186/1471-2164-15-795.

288. Zook, J.M.; Chapman, B.; Wang, J.; Mittelman, D.; Hofmann, O.; Hide, W.; Salit, M. Integrating human sequence data sets provides a resource of benchmark, SNP and indel genotype calls. *Nat. Biotechnol.* **2014**, *32*, 246–251.

289. Nakagome, M.; Solovieva, E.; Takahashi, A.; Yasue, H.; Hirochika, H.; Miyao, A. Transposon Insertion Finder (TIF): A novel program for detection of *de novo* transpositions of transposable elements. *BMC Bioinf.* **2014**, *15*, doi:10.1186/1471-2105-15-71.

290. Helman, E.; Lawrence, M.S.; Stewart, C.; Sougnez, C.; Getz, G.; Meyerson, M. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Res.* **2014**, *24*, 1053–1063.

291. Tubio, J.M.; Li, Y.; Ju, Y.S.; Martincorena, I.; Cooke, S.L.; Tojo, M.; Gundem, G.; Pipinikas, C.P.; Zamora, J.; Raine, K.; et al. Extensive transduction of nonrepetitive, DNA mediated by L1 retrotransposition in cancer genomes. *Science* **2014**, *345*, doi:10.1126/science.1251343.

292. Chen, K.; Chen, L.; Fan, X.; Wallis, J.; Ding, L.; Weinstock, G. TIGRA: A targeted iterative graph routing assembler for breakpoint assembly. *Genome Res.* **2014**, *24*, 310–317.

293. Gilly, A.; Etcheverry, M.; Madoui, M.A.; Guy, J.; Quadrana, L.; Alberti, A.; Martin, A.; Heitkam, T.; Engelen, S.; Labadie, K.; et al. TE-Tracker: Systematic identification of transposition events through whole-genome resequencing. *BMC Bioinf.* **2014**, *15*, doi:10.1186/s12859-014-0377-z.

294. Ewing, A.D.; Ballinger, T.J.; Earl, D.; Harris, C.C.; Ding, L.; Wilson, R.K.; Haussler, D. Retrotransposition of gene transcripts leads to structural variation in mammalian genomes *Genome Biol.* **2013**, *14*, 1–14.

295. Robb, S.M.; Lu, L.; Valencia, E.; Burnette, J.M.; Okumoto, Y.; Wessler, S.R.; Stajich, J.E. The use of RelocaTE and unassembled short reads to produce high-resolution snapshots of transposable element generated diversity in rice. *G3 Genes Genomes Genet.* **2013**, *3*, 949–957.

296. Lee, E.; Iskow, R.; Yang, L.; Gokcumen, O.; Haseley, P.; Luquette, L.J.; Lohr, J.G.; Harris, C.C.; Ding, L.; Wilson, R.K.; et al. Landscape of somatic retrotransposition in human cancers. *Science* **2012**, *337*, 967–971.

297. Linheiro, R.S.; Bergman, C.M. Whole genome resequencing reveals natural target site preferences of transposable elements in *Drosophila melanogaster*. *PLoS ONE* **2012**, *7*, e0030008.

298. Platzer, A.; Nizhynska, V.; Long, Q. TE-Locate: A tool to locate and group transposable element occurrences using paired-end next-generation sequencing data. *Biology (Basel)* **2012**, *1*, 395–410.

299. Hormozdiari, F.; Hajirasouliha, I.; Dao, P.; Hach, F.; Yorukoglu, D.; Alkan, C.; Eichler, E.E.; Sahinalp, S.C. Next-generation VariationHunter: Combinatorial algorithms for transposon insertion discovery. *Bioinformatics* **2010**, *26*, i350–i357.

300. Quinlan, A.R.; Clark, R.A.; Sokolova, S.; Leibowitz, M.L.; Zhang, Y.; Hurles, M.E.; Mell, J.C.; Hall, I.M. Genome-wide mapping and assembly of structural variant breakpoints in the mouse genome. *Genome Res.* **2010**, *20*, 623–635.

301. Lee, H.; Lee, M.; Mohammed Ismail, W.; Rho, M.; Fox, G.C.; Oh, S.; Tang, H. MGEScan: A Galaxy based system for identifying retrotransposons in genomes. *Bioinformatics* **2016**, *32*, 2502–2504.

302. Novák, P.; Neumann, P.; Macas, J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinf.* **2010**, *11*, 378, doi:10.1186/1471-2105-11-378.

303. Singh, V.; Mishra, R.K. RISCI–Repeat Induced Sequence Changes Identifier: A comprehensive, comparative genomics-based, in silico subtractive hybridization pipeline to identify repeat induced sequence changes in closely related genomes. *BMC Bioinf.* **2010**, *11*, 609, doi:10.1186/1471-2105-11-609.

304. Pestilli, F.; Yeatman, J.D.; Rokem, A.; Kay, K.N.; Wandell, B.A. Evaluation and statistical inference for human connectomes. *Nat. Methods* **2014**, *11*, 1058–1063.

305. Ellinghaus, D.; Kurtz, S.; Willhoeft, U. LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. *BMC Bioinf.* **2008**, *9*, doi:10.1186/1471-2105-9-18.

306. Du, C.; Caronna, J.; He, L.; Dooner, H.K. Computational prediction and molecular confirmation of *Helitron* transposons in the maize genome. *BMC Genom.* **2008**, *9*, 51, doi:10.1186/1471-2164-9-51.

307. Lorenzi, H.; Thiagarajan, M.; Haas, B.; Wortman, J.; Hall, N.; Caler, E. Genome wide survey, discovery and evolution of repetitive elements in three *Entamoeba* species. *BMC Genom.* **2008**, *9*, 595, doi:10.1186/1471-2164-9-595.

Chapter 2: Paternally inherited P-element copy number affects the magnitude of hybrid dysgenesis in *Drosophila simulans* and *D. melanogaster*

Aim 2: If the number of TEs is related to deleterious effects, this provides a window through which TEs can increase their frequency within hosts throughout populations before selection can act on any copy-number dependent deleterious effects. P-elements (PE) in Drosophila provide an ideal system through which to study whether the number of PEs, and TEs generally, are an important factor in determining phenotypic defects that within-species F1s suffer. We address this question by inferring the number of PEs in individual isofemale line genomes and measuring whether there is an effect of the P-element copy number on multiple phenotypes associated with the HD syndrome in two species of *Drosophila*.

INTRODUCTION

Transposable elements (TEs) are common across Eukaryotes and make up a large portion of their genomes across various taxa, yet the extent of their effects as they propagate are poorly understood despite the vital information it provides in understanding how TEs are so prevalent. Insertions of TEs in host genomes have been associated with a variety of effects, mostly deleterious, and include genic disruptions, regulatory changes, and chromosome structural changes [1]. Despite this, TEs can account for a large portion of the genome in both metazoans and plants, making up 69% of the DNA in humans and 85% in maize [2,3], with TE content varying between species and even populations [4,5]. Most TEs are localized within noncoding

and untranslated regions, however the high frequency of elements also results in regularly indirect effects on genes or even direct insertions that selection is more likely to act on [6]–[8]. Transposition and excision of TEs is reduced through evolutionary time, as uncontrolled transposition can have drastic fitness consequences and the host genome evolves TE-repression mechanisms [9–11]. A natural hypothesis is that an increase in copy-number of TEs increases the likelihood of negatively affecting the host through disrupting or altering normal expression, particularly in TEs associated with gametic proliferation as the effects on fitness are direct. Despite this, little is known regarding the relationship between TE copy number and the intensity of deleterious effects, as studies have modeled and explored this prospect to varying degrees but with mixed conclusions [12,13]. Here, we explore how the presence of a single gamete specific TE, that varies in copy number between individuals of two distinct species affects five metrics of fecundity. The aim of which is to test whether there is differential selection acting on individuals at distinct stages of TE invasion, as TEs increase in frequency within a genome, providing an avenue through which TEs can propagate through a population before selection acts against them.

One of the best studied cases of the phenotypic effects of TEs across Eukaryotic systems is P-elements (PEs) in *Drosophila*. PEs have rapidly spread worldwide throughout populations of the genetic model system, *D. melanogaster*. The first report of PE carrying (PE+) *D. melanogaster* individuals was in 1977; PE infection frequencies among populations increased rapidly and no non-PE carrying wild-type lines have been found since 1974 [14,15]. Despite the self-replicating nature of PEs, this spread throughout the entirety of a species is puzzling, due to the drastic negative phenotypic effects they can cause in the host [15]. PEs in *D. melanogaster* lead to F1 sterility when the germ line of the female does not carry the molecular machinery

regulating the expansion of PEs [16,17]. When a female who lacks PEs mates with a male PE carrier, the resulting F1s (both females and males) are sterile, show elevated rates of chromosomal breakage and increased mutation rates, a suite of traits referred to collectively as hybrid dysgenesis (HD) [15]. Conversely, if a female with PEs mates with a PE male carrier, the F1s are fertile and show no signs of HD. In this case, the effects of PEs are silenced through a maternally inherited and germline-specific subclass of small non-coding RNAs, piRNAs (PIWI-interacting RNAs). This RNA facilitated silencing mechanism is not specific to PEs and has been shown to underlie repression throughout multiple classes of TEs and seems to play a role in repression across a variety of plant and mammalian hybrids [16]. However, misexpression of genes in hybrids is common and the extent to which TEs are more or less susceptible to derepression in hybrids remains unknown [18,19]. Ideally, understanding the effects that TEs have on a genome should be carried out in a system with closely related species that can utilize a large foundation of genetic information and can be experimentally manipulated, in order to better decouple causation from correlation among a broad range of genetic differences. We focus on the model system of *Drosophila*. In *Drosophila*, F1 sterility due to PEs is a simple and elegant model of how relatively simple genomic changes (i.e., the invasion of a TE) can induce reproductive isolation between genotypes rapidly and potentially lead to speciation [20].

Recently, PEs were also found in populations of *D. simulans*, another cosmopolitan human commensal of the *melanogaster* species subgroup. PEs spread into *D. simulans*' entire range within 15 years and the phenotypic effects of PEs, in the form of sterility of F1s, is similar in both species [21]. Since the hybrid progeny between *D. melanogaster* and *D. simulans* are sterile [22,23], there is no obvious genetic bridge through which PEs could have entered *D. simulans* from *D. melanogaster* through horizontal transmission. As a result, a natural question is

whether the genomes of related species have also been invaded by PEs and whether there are any conserved patterns in their transmission and intensity of phenotypic effects [20,24,25]. Intriguingly, PEs are not present in *D. sechellia* nor *D. mauritiana*, the other two species that form the *simulans* clade; both species show signatures of introgression from *D. simulans* [26] and PEs could have migrated through hybridization but have not been detected in previous studies despite being present in *D. simulans* [27].

PEs are polymorphic in copy number in terms of their presence or absence across the geographic range in *D. simulans*, but also vary in the number of PE copies per genome both in *D. simulans* and *D. melanogaster*. One of the outstanding questions in our understanding of PEs is whether the number of PE copies affects the magnitude of hybrid dysgenesis. This hypothesis was originally proposed in 1980 by Montgomery et al and stated that the higher the number of PE copies, the more likely that PEs will be misregulated in F1s produced in crosses between PE- mother and PE+ fathers [11]. In spite of the straightforward nature of this hypothesis, the proposition remains largely unresolved. In *D. melanogaster*, there seems to be a positive, but weak, relationship between PE number and the proportion of dysgenic female progeny in crosses between ♀PE- and ♂PE+ males [28]; In *D. simulans* the relationship between these two phenotypes has been reported as strong and biological meaningful. However, other reports argue the correlation does not exist and that the magnitude of variation in gonadal dysgenesis across strains of *D. melanogaster* cannot be explained by the number of PEs in the paternal genome. No consensus has been achieved whether there is a correlation between these two traits. This is an important question because it shows whether PEs are just mildly deleterious when they first invade a genome (i.e., low number of PE copies cause only weak HD) and if, as they increase in copy number, they become deleterious enough that the number of copies might plateau or the

lineage might go extinct [29], preventing the PEs from further propagating throughout populations.

Despite the large body of work that has explored PEs and led to PEs becoming the leading model to understand the evolution of TEs, establishing the extent of PE copy variation in individual genomes is a recent development. Earlier studies have used; Southern blotting, *in situ* hybridization to polytene chromosomes, and qPCR to estimate the number of PE copies in a genome. These approaches, however, are difficult to scale to population levels and provide low resolution. Genome sequencing provides a solution to quantifying the number of PE copies in a genome because, one can infer not only how many copies are present but also verify where they are inserted. Recent work has compared different computational methods and found that estimated copy-number are highly correlated across different bioinformatics approaches. From a technical standpoint, it is now possible to count the number of PE copies per genome to a high degree of confidence.

In this report, we estimate the number of PE copies in the genomes of seven *D. simulans* lines and seven *D. melanogaster* lines. We confirmed the presence of full PEs using PCR and also used Paired-end read data to infer the number of PEs per genome in isofemale lines collected in the island of Bioko. The results indicate the number of PEs per genome in our sample ranges from 5 to 20 in each species. We then measured the magnitude of hybrid dysgenesis in crosses between PE+ and PE- lines in *D. simulans*, by scoring for five traits associated with HD syndromes. We found that HD in *D. simulans* is not restricted to just atrophied gonads in F1 PE-/PE+ individuals but instead can manifest itself as a continuum in the form of: reduced ovariole number in non-sterile females, reduced male fertility, and early onset of reproductive senescence in females. Finally, we tested whether the number of PEs in the

paternal genome was correlated with the strength of each of the 6 metrics that constitute HD. We find that in both species, *D. simulans* and *D. melanogaster,* the number of PEs in the genome is indeed correlated with the magnitude multiple phenotypes associated to the syndrome of HD but the strength of the correlation differs between phenotypes and species.

METHODS

Stocks

D. simulans: DRM collected six *D. simulans* lines in the island of Bioko in the year 2013. These lines were later inferred to be infected with PEs (See immediately below). All flies were collected in locations reported elsewhere [30,31] using yeasted banana traps. We used female flies from this collection to start isofemale lines (i.e., stable stocks derived from the progeny of a single female). All lines have been maintained in corn-meal food bottles since they were started. An additional line was collected in Florida, where the first report of PEs into *D. simulans* was inferred [21].

D. melanogaster: We collected *D. melanogaster* lines from Malawi, Zambia and Namibia in 2015. We obtained five females that were then maintained as isofemale lines. As expected, all these lines were infected with PEs (see below). All collections, isofemale line establishment, and fly rearing and maintenance was done as described above for *D. simulans*. Additionally, we used three lines from the Drosophila Genetic Reference Panel that were infected with PEs.

PE detection

We used PCR and Illumina sequencing to detect PEs;. We assessed whether individuals from these lines had any of the four exons that constitute a full PE. Since PEs require all four exons to be functional, our goal was to type all the individuals for each exon individually using

PCR. We extracted genomic DNA from one female of each isoline (or an individual in ethanol) following the 96-well Gentra Puregene extraction kit protocol. To individually amplify each of the 4 exons that make up the full PE with both positive and negative controls during each run, we used primers described in (Hill, Schlötterer, & Betancourt, 2016). We did all PCRs using NEB reagents in a 10ul reaction (1ul 10x buffer, 1ul 10mM $MgCl^2$, 0.5 ul 10mM dNTPs, 0.3 ul 10mM F+R primers, 1ul DNA, 0.05 Taq Polymerase, 5.85 ul H_20) with a thermocycling cycle of 92° denaturing, 59° annealing, 72° extension for 35 cycles in an Applied Biosystems 2720 Thermal Cycler. To score presence/absence of each exon, we ran 5ul of the PCR product in a 2% (APExBIO) agarose gel for 60 minutes at 120 volts and visualized the results using ethidium bromide staining. Sanger sequencing (Eurofins) was used to verify for PE presence in isolines that amplified for each primer to ensure the presence of the full continuous element.

Second, we used short-read (Illumina) for each isofemale line included in this study from the three focal species (*D. simulans*, and *D. melanogaster*). Paired-end Illumina reads were aligned to the canonical *D. melanogaster* PE (https://flybase.org/reports/FBte0000037.html) using minimap2 in short-read mode ("-cx sr"). Pairs where either read aligned partially or fully to the PE were separately aligned to the *D. simulans* reference sequence r2.02 (ftp://ftp.flybase.net/genomes/Drosophila_simulans/dsim_r2.02_FB2017_04/fasta/dsim-all-chromosome-r2.02.fasta.gz). PE insertions sites were detected using a custom Python script. Briefly, if a single read partially aligned to the P-element and the reference sequence, the exact insertion site can be trivially determined, with most sites supported by several reads. Pairs where one read aligned fully to the P-element and one fully to the reference sequence add support for an insertion site anywhere within ~400bp of the reference-aligned read. Each insertion site is supported by an average of 9.8 read pairs. All lines that showed fewer than 0.5 copies of the PEs

(i.e., a single heterozygote PE copy), were considered PE-negative. This approach also allowed us to infer how many PE copies each sampled isofemale line harbored.

Scoring HD Phenotypes

D. simulans F1 females from crosses between PE- females and PE+ males are likely to show atrophied ovaries [32]. In *D. melanogaster*, HD has been extensively studied and in addition to the number of atrophied ovaries, other defects have been associated to the HD syndrome, namely, dysgenic (sterile) males, reduced number of ovarioles, and rapid reproductive senescence. We studied each of these defects in *D. simulans* using a diallelic design (i.e., all possible crosses in both reciprocal directions) in which we crossed four *D. simulans* lines that carried PEs and four lines that had no evidence of PE infection. All PE+ lines correspond to the P-cytotype as they have intact copies of the PE (see above for the information in the isofemale lines).

Gonad number—Counts: First, we scored whether F1 individuals had zero, one or two developed gonads, with healthy females and males having two ovaries and two testes respectively, at both 23° and 29°C. After 4 to 9 days flies were anesthetized with CO_2 and their gonads removed with metallic forceps [33]. Gonads from each individual were subsequently fixed on a precleaned glass slide with chilled *Drosophila* Ringer's solution (Cold Spring Harbor Protocols). We counted the number of non-atrophied gonads for each individual. Ovaries were considered atrophied if they had no ovarioles. Testes were considered atrophied if they had less than half the length of wild-type testes. In the case of females, we also counted the number of ovarioles (see below) in each mature ovary using a Leica, S6E stereoscopic microscope. We scored >100 females at 23ºC and >40 females at 29ºC, as higher temperatures are associated with an increase

of TE transposition [34]. Table 1 shows the number of females dissected for each genotype. For ovariole counts, we only scored flies for which the dissection contained both left and right gonads.

Ovary number—Statistical analyses: We scored whether each F1 female had zero, one, or two ovaries as described above. To quantify the magnitude of heterogeneity among F1 genotypes, we fitted a multinomial regression using the function *multinom* in the library *nnet* [35] where the number of ovaries was the response of the multinomial assay and the mother and father genotypes were the fixed effects. We also included the interaction between these two effects to account for the interplay between the genome of the two parents. The significance of the effects was inferred using the function *set_sum_contrasts* (library *car* [36]), and a type III ANOVA (library *stats* [37]) in R. Since we did experiments at two different temperatures (23ºC and 29ºC), we fitted two multinomial regressions. To do post-hoc comparisons between crosses, we used a Two-Sample Fisher-Pitman Permutation Test (library *coin*, function '*oneway_test'*; [38]) and adjusted the critical P-values for significance to 0.008 to account for multiple comparisons (6 comparisons).

Ovariole number—statistical analyses: A second potential phenotype of HD is the reduction in the number of ovarioles per ovary in female F1s, in females that did not show atrophied ovaries [39]. In these females, even with two ovaries, their reproductive potential can be limited through a lack of ovarioles [40]. We quantified whether the genotype of the mother, of the father, or the interaction between these two terms affected the number of ovarioles. We analyzed the mean number of ovarioles per ovary (i.e., females with two ovaries will have more total ovarioles than females with one ovary) to account for difference in the number of ovaries. We excluded those females that showed completely atrophied ovarioles from this analysis

because they contain zero ovarioles. We used a Poisson-distributed linear model (library *stats*, function '*glm'* [37]). To assess the significance of interactions, we followed a maximum-likelihood model simplification approach [41]; we first fitted a fully factorial model containing all factors and interactions and then simplified it by a series of stepwise comparisons, starting with the highest-order interaction and progressing to lower-order interaction terms and then to main effects.

Female reproductive senescence—counts. We tested whether the age of the female had an effect on the number of ovarioles in PE^+ and PE^- females. Specifically, we explored whether HD manifested itself as a shorter reproductive period in females that carried PEs [40]. In this scenario PE^+ will show a sharper decline in their ovariole number compared to their PE^- females. To score females of different age, we cleared bottles and collected newly eclosed virgins within 8 hours of clearing as described above (Section '*Crosses'*). To account for heterogeneity across lines, we studied 5 different isolines per population type: 5 PE^+ isofemale lines, and 5 PE^- isofemale lines, for a total of 10 isofemale lines per time point. Female virgins were then dissected every 5 days for 25 days to count the ovariole count as they aged. In total there were 500 observations: 5 time points × 5 isolines × 10 individuals per line × 2 distinct population types.

Female reproductive senescence—Statistical analyses: We used an Analysis of covariance (ANCOVA) to assess whether the presence of PEs affected the reproductive capacity of a female at different ages. We used the function *lm* in the R library *stats* [37]. First, we used the regression coefficients from the ANCOVA to compare the intercept of the linear regressions of females with and without PEs. This test assessed whether genotypes had inherent differences in the number of ovarioles (i.e., whether the effect of genotype—if a female is PE^+ or PE^-—was

significant). Second, we compared the rate of decline of fertility among genotypes. To this end, we quantified differences in the slope of the regressions of number of ovarioles as age progressed (i.e., the interaction between female age and her genotype). To evaluate the significance of the interaction, we used information obtained with the function *lm* as described immediately above and also performed a likelihood ratio test (LRT; function *lrtest*, R library *lmtest* [42]).

Male fertility—sperm motility. We scored whether F1 male progeny produced motile sperm. We dissected the testes of each individual with metallic forceps (Miltex Catalogue number: 17-301, McKesson, Richmond, VA) and mounted them on chilled Ringer's solution. We mounted up to five males per slide and scored whether they had motile sperm within 5 minutes of starting the first dissection. We scored 843 F1 males at 23ºC and 542 F1 males at 29ºC. To quantify the effect of the genotype on sperm motility among F1 genotypes, we fitted a binomial regression (library stats, function '*glm'*). Whether a male had fertile sperm or not was the response of the binomial model, while the mother and father genotypes were the fixed effects. We also included the interaction between these two effects to account for the interplay between the genome of the two parents. We used LRTs (described above) to test whether to retain the interaction and the fixed effects. We found no sterile males at 23ºC, so we fit a single linear model at 29ºC. 95% confidence intervals for point estimates of the proportion of F1 individuals which were sterile were calculated using the conjugate beta prior on the distribution of successes (library binom, function '*binom.cloglog*' [43]).To do posthoc tests, we used a Tukey Honest significance difference test (library multcomp, function '*glht'*).

Male fertility— progeny count. Finally, we scored whether F1 male progeny showed reduced fertility despite showing normal size testes (see above). We collected F1 males from the four F1 genotypes raised at the two studied temperatures (23ºC and 29ºC) and mated them to virgin PE- females. We watched the matings to ensure they were not abnormally short (less than 10 minutes [44]); as soon as the mating was over, we removed the male from the vial. We let the female lay eggs for 10 days. After this period, we removed the females and let the progeny develop at 23ºC. Every two days, we counted the progeny produced by each female until no more flies emerged. We quantified the heterogeneity of the amount of progeny using a generalized linear model similar to the one described above (section 'Ovariole number—statistical analyses') where the number of progeny produced by each individual female was the response, the genotype of the cross and temperature at which the cross was performed were the fixed effects.

Effects of the number of PEs on the magnitude of HD.

We tested whether the number of PEs affected the magnitude of three phenotypes within the HD syndrome. We studied the effect of PEs on F1 female fecundity (i.e., ovary number and mean ovariole number per ovary) and F1 male fertility. For these experiments we used seven lines that showed evidence of PE presence (inferred from the Paired end genome data) and that showed a range in the number of PEs they harbored. In the case of *D. simulans*, we used seven lines that did not contain PEs and in the case of *D. melanogaster* we used four PE- lines. As described above, ovary number can only take values of 0,1, and 2 per individual (see '*Scoring HD Phenotypes'*). We treated this phenotype as a multinomial outcome. We scored 10 females per line combination for a total of 490 females in *D. simulans* and 280 in *D. melanogaster*. To

determine whether the trait was affected by the number of PE copies, we used a multivariate regression as described above.

We used a similar approach to study mean ovariole number and male fertility. For these experiments we used the same combinations of isofemale lines (and sample sizes) used for the number of ovaries per female analyses described immediately above. Ovariole number and male fertility of F1 was scored as described above (see '*Scoring HD Phenotypes'*). All linear regressions were done using the R package *glm* (i.e., day females or males were dissected).

RESULTS

PE insertions show overlap in *D. simulans*

Using PE Illumina reads, we found that all lines collected in Bioko island were PE+ but the number of PEs in each genome differed. We found an estimated number of PEs that range from 6 to 18 (mean number of copies$_{sim}$ = 11.21) This number of copies per genome is slightly lower than the number of insertions found in *D. melanogaster* (mean number of copies$_{mel}$ = 16.44; t = 3.2799, df = 6.2416, p-value = 0.01589). We identified 1,311 exact and 118 approximate insertion positions, of which 45 are shared by multiple lines. In total; 1,215 sites were unique to each genome, 40 sites were shared by 2 genomes, 4 sites shared by 4, and 1 site was shared in 4 of the 6 unique genomes.

Intraspecific effects of PEs within *D. simulans*

Previous reports have found evidence of HD within *D. simulans* in the form of ovary number reduction [45]. We explored whether F1 *D. simulans* from crosses between PE- and PE+ lines showed evidence of an increased of HD associated phenotypic effects (F1 male sterility,

early reproductive senescence, and reduced number of ovarioles) based on PE copy number. We report the results of each phenotype as follows.

Number of ovaries. We assessed whether the mother and father genotype had an effect on the number of functional ovaries in F1 females. We pooled all lines into genotypic categories to perform a multinomial regression. We find that at the two temperatures, the effect of the father is significant (23ºX: $\chi^2_{2} = 40.293$, P= 1.781×10^{-9}; 29ºC: $\chi^2_{2} = 39.972$, P= 2.09×10^{-9}), while the effect of the mother was not significant at either of the two temperatures (23ºX: $\chi^2_{2} = 2.318$, P= 0.31388; 29ºC: $\chi^2_{2} = 2.318$, P= 0.36846). The interaction between mother and father genotypes was significant (23ºX: $\chi^2_{2} = 8.033$, P= 0.01802; 29ºC: $\chi^2_{2} = 7.712$, P= 0.02115). We used permutation-based pairwise comparisons to determine whether F1 *D. simulans* from the PE- × PE+ cross from crosses of lines collected in the island of Bioko had fewer ovaries than females from any of the other three possible genotypes. We find that females from this cross have, on average, fewer ovaries than the rest of the possible crosses at both temperatures (Table 1). The other three crosses did not differ among themselves. This result is in line with a previous report from a *D. simulans* population from Florida also suffers from hybrid dysgenesis [32].

Number of ovarioles: Hybrid dysgenesis can manifest itself not only as the absence of ovaries but also through the development of "rudimentary" ovaries, i.e. ovaries with fewer ovarioles. Since F1s in *D. simulans* crosses might show absent ovaries depending on the genotype of the parents (e.g., ♀PE- × ♂ PE+), we used the mean number of ovaries per individuals and instances of atrophied ovaries were treated as missing data. We found that just as is the case with ovary number, ovarioles are affected by HD. Patterns were similar at both temperatures but were more pronounced at 29ºC. We found that at both temperatures, the mother, and father genotypes affected the number of ovarioles on F1s (Table 2). The father effect

had an effect size more than three times larger than three mother effect. The interaction terms between the two effects were also significant at both temperatures, being larger at 29ºC (Table 2). Notably, we also found differences between temperatures. At 23ºC the line effects were not significant but they were at 29ºC (Table 2). We compared the magnitude of paired crosses using a linear model in which the only effect was the interaction of the genotypes using a linear model that only incorporated the interaction term. As expected by hybrid dysgenesis, PE-/PE+ progeny have fewer ovarioles than the other three types of F1 females (Table 3). We found no significant difference between any of the other six pairwise comparisons at any of the two temperatures (Table 3).

Reproductive senescence: A second potential phenotype in hybrid dysgenesis is that PE-carrying females show a rapid decrease in fertility as they age [46]. Specifically, we tested whether the presence of PEs was predictive of reproductive output throughout the lifespan of females. We tested this possibility by counting the number of ovarioles of females with and without PEs at five different ages for 25 days (Figure 1). The number of ovarioles decreases as females age at similar rates. The intercept was similar for both types of females which indicates the initial reproductive potential is similar in females carrying and not carrying PEs (genotype effect: Table 3). Additionally, the rate of decrease (i.e., the slope of the linear regression) was not different for the two regressions either (genotype by age interaction). These results indicate that PEs in *D. simulans* do not induce early reproductive senescence at 23ºC. Due to high mortality of *D. simulans* at high temperature, we did not test the effect of PEs at 29ºC.

Male sterility. We studied whether PE presence increased male sterility in two ways, scoring sperm motility and counting progeny produced when mated to a PE- *D. simulans* lines. First, we dissected the testes of F1 males from crosses between *D. simulans* PE- and PE+

individuals. We used a diallelic cross scheme in which we crossed all the possible combinations of two PE- and two PE+ lines. At 23ºC, all F1 males, regardless of their genotype, contained functional testes (n = 20 males per line combination). All males had motile sperm at this temperature. At 29ºC, male sterility was most often observed in individuals produced from the crosses that involved a PE^+ parent (PE^-/ PE^+, PE^+/ PE^-, and PE^+/ PE^+) than in males with no PEs (PE^-/ PE^-;). It is worth noting that the number of dissected males at higher temperatures was lower than that at 23ºC (n = 15 males per line combination in average) because *D. simulans* is sensitive to high temperatures [47]. PE-/PE+ males were more likely to be sterile (proportion of sterile males $_{PE-/PE+}$= 0.4722222; 95% CI= [0.3758025-0.5625837]), than any of the other three crosses (proportion of sterile males PE+/PE-= 0.4259259; 95% CI= [0.3318431-0.5166623]; the other two crosses show even lower rates of sterility). There was significant heterogeneity in the F1 male fertility caused by whether the father carried PEs (Mixed model logistic regression followed by a type 'III' ANOVA: χ^2_1 = 6.9187, P = 0.00853), and to a lesser extent of the whether the mother carried PEs (χ^2_1 = 5.1861, P = 0.02277). The effect of the interaction between genotypes was not significant (χ^2_1 = 3.0712, P = 0.07969).

Second, we scored the fertility of the four different genotypes of *D. simulans* F1 males when they were mated to PE- females. When males were raised at 23ºC, we found that number of progeny produced between genotypes was similar and showed no differences (mean of all genotypes $pooled_{25ºC}$= 61.02399, $SD_{25ºC}$ =4.215693; $F_{3,538}$ = 0.6378, P = 0.591). Consistent with previous studies [47,48], we found that crosses at 29ºC produce fewer progeny than crosses at 23ºC in *D. simulans* (mean of all genotypes $pooled_{29ºC}$= 37.26753, $SD_{29ºC}$ = 8.105156; t = 60.674, df = 819.25, $P < 1 \times 10^{-10}$). At 29ºC, the effect of PEs in male fertility is noticeable and both the effects of the mother ($F_{3,538}$ = 262.77, $P < 1 \times 10^{-10}$) and the father ($F_{3,538}$ = 161.74, P < 1

$\times 10^{-10}$) are significant, mirroring similar patterns seen in *D. melanogaster*. The interaction effect was also significant ($F_{3,538} = 262.77$, $P < 1 \times 10^{-10}$). Pairwise comparisons, based on a linear model that includes only the interaction term, indicate that F1 males from the PE-/PE+ cross produce fewer progeny than any of the other three crosses (Table 4). These results are consistent with HD causing reduced male fertility which manifests as slower fecundity but not as complete gonadal atrophy. This defect only manifests itself at higher temperatures.

The number of PEs in the *D. simulans* genome is correlated with the magnitude of HD in *D. simulans* and *D. melanogaster*

The experiments described above revealed that HD in *D. simulans* is not restricted to atrophied ovaries but can also manifest as reduced number of ovarioles in non-dysgenic ovaries, reduced male fertility and no difference in the rate of female reproductive senescence. These are the same phenotypes that have been previously reported for *D. melanogaster*. Using this suite of phenotypes and the estimation of the number of PEs per genome in *D. simulans* and *D. melanogaster*, we tested whether the number of PE copies in the genome was correlated with the strength of hybrid dysgenesis. This hypothesis poses the possibility that in crosses involving a PE- female and a PE+ male, males with more PE copies will cause stronger HD than males with fewer copies of PEs. To this end, we expanded the studies on HD from 4 *D. simulans* lines to 14 *D. simulans* lines (seven with PEs and seven without PEs). We followed a similar approach for *D. melanogaster* (11 lines). Since we identified PE+ *D. simulans* lines with a PE copy number that ranged between 6 and 12 copies per genome, and *D. melanogaster* lines that ranged between 5 and 20 copies per genome, we tested whether there was a positive relationship between the number of PE copies and the strength of HD in both of these species.

Ovary number: As the range of the number of ovaries is narrow (zero, one, or two), we used multinomial regressions to assess whether PEs affects the number of ovaries in F1 females in each of the two species. In the case of *D. melanogaster*, the effect of the number of PEs in the number of ovaries is negative but minimal (23ºC: $\chi^2_2 = 6.1208$, P = 0.04687; 29ºC: $\chi^2_2 = 5.0755$, P = 0.07904). We found a similar, but stronger trend in *D. simulans*; as the number of PEs increases in the paternal line, the number of ovaries per ovary per female decreases at 23ºC ($\chi^2_2 = 19.962$, $P = 4.626 \times 10^{-5}$) and at 29ºC ($\chi^2_2 = 12.772$, P = 0.001685). These results indicate that the number of PE copies in the paternal genome does indeed affect the magnitude of HD, in the form of atrophied ovaries, but that the importance of the effect is contingent on the species.

Ovariole number: Next, we assessed whether PE number on the paternal genome also affected the number of ovarioles in F1 females from crosses between PE- females and PE+ males. Table 5 shows the regression coefficients for each of the four linear models (2 temperatures × 2 species). In the case of *D. melanogaster*, the relationship between PE copy number and the mean number of ovarioles is negative and strong, both at 23ºC ($F_{1,278} = 12.487$, P = 0.0004797) and 29ºC ($F_{1,278} = 29.972$, $P = 9.8 \times 10^{-8}$). We see a similar pattern in *D. simulans*. PE copy number has a negative effect on the number of ovarioles per ovary per female at 23ºC ($F_{1,488} = 35.124$, $P = 5.854 \times 10^{-9}$) and 29ºC ($F_{1,488} = 37.204$, $P = 2.167 \times 10^{-9}$). These results indicate that, just as described with ovary number, that HD is a quantitative syndrome that is affected by the number of PEs in the parental lines. Moreover, it also suggests that ovariole number might be a finer scale phenotype to study HD than ovary number.

Male fertility: Finally, we assessed whether the fertility of males from F1 crosses between PE- females and PE+ males was affected by the number of PE copies in the paternal genome. We scored the progeny production of these males after single matings to PE- females. In the case of *D. melanogaster*, we find a modest effect of the number of PEs in male fertility. This effect is larger at 29ºC ($F_{1,278}$= 5.4951 P=0.01977) than at 23ºC ($F_{1,278}$= 4.1815, P=0.04181). In the case of *D. simulans*, the importance of PEs for male fertility is larger. We find a strong effect of the PE copy number at 23ºC ($F_{1,488}$ = 18.262, P=2.316×10^{-5}) and 29ºC ($F_{1,488}$= 29.544, P = 8.648×10^{-8}). These results indicate that just as is the case with female fertility (i.e., ovary and mean ovariole number), male fertility is affected by the number of PEs in the paternal genome, and that fertility reduction in males, as part of the HD syndrome, is a continuous trait.

DISCUSSION

PEs are arguably the best studied TE in animals that can be experimentally manipulated and their phenotypic effects in *Drosophila* provides an ideal system to understand the organismal fitness effects during their spread within genomes. The most recent invasion of PEs into a novel species has occurred in *D. simulans*, where they lead to atrophied ovaries in PE-/PE+ females [45], mirroring effects seen in *D. melanogaster* [14]. We expanded our understanding of the phenomenon of HD in *D. simulans* by testing whether other HD-associated phenotypes observed in *D. melanogaster* also occur in *D. simulans*. We find that in PE-/PE+ F1s, females with functional gonads have a reduced ovariole number. We also find that PE-/PE+ males show reduced fertility, despite containing functional sperm. This suite of traits is similar to the phenotypic defects associated to HD in *D. melanogaster*, the first species where HD was reported [15]. The discovery of PEs in *D. simulans* [27], and the associated HD caused by the invasion suggest that if PEs succeed at invading a species (i.e., do not get purged or caused

extinction of the lineage) then the outcome of HD might be predictable. Notably, we also find differences between species; HD causes a more noticeable reduction of female and male fitness in *D. simulans* than in *D. melanogaster*. The reasons for this difference remain unknown but might be related to either the recency of the PE invasion or the HD-suppression mechanisms within each species. These possibilities need to be formally addressed, as little is known regarding the evolution of RNAi repression systems in response to TE activity [9,49].

Even though the phenomenon of hybrid dysgenesis has been rigorously characterized in *D. melanogaster* [9], the discovery of PEs in other *Drosophila* species allows us to understand how these elements behave in different genetic backgrounds. Additionally, PEs being incorporated into the genome of a species within a hybridizing species complex provides the first case of PE effects being tested between distinct species. Hybrid dysgenesis is contingent on the genotype of the father and of the mother (i.e., presence or absence of PEs). PEs inherited from the father become overly active in the F1 germ line and lead to sterility when the female mate does not have the PIWI systems required to repress PE transposition. Fine scale assessments have revealed that the nuclear genome of the mother is also involved in the phenomenon of HD. In particular, variation at the *bruno* locus of *D. melanogaster* is associated to variability in the magnitude of female hybrid dysgenesis by modulating germline stem cell (GSC) loss in the presence of P-element activity [50]. There is certainly variability in the phenomenon of HD in *D. simulans*, but whether this locus also has an effect on this species is an open question that can be addressed through the use of QTL (i.e., [50]) or inbred lines panels [51].

One of the aspects that has remained unresolved is the effect of the copy number of PEs, and TEs generally, in the affecting phenotypes within hosts. Our report improves previous experiments for two reasons. First, previous studies had mostly focused on the percentage of

dysgenic females, in the form of atrophied ovaries, per isofemale line. Besides measuring more phenotypes, we also incorporate the magnitude of variation among individuals instead of focusing solely on the variation among isofemale lines. This is an important distinction because isofemale lines are routinely not fully inbred allowing for significant—and to some extent surprising—variation among individuals within a line. Second, we measured ovary number per female, ovariole number per ovary, and individual male fertility. It is worth noting that we focus on the PE number of copies exclusively; however studying whether other different TEs interact to cause HD in F1 individuals (e.g., [52]) will shed light on how different TE infections interact with each other and their host.

The dynamics of PEs, and TEs generally, after they invade a genome is an open field of study that continues to be explored as PEs actively propagate throughout populations and species. While it is clear that PEs have increased their frequency in both *D. simulans* and *D. melanogaster*, little is known about the progression of the copy number per species and its associated effects. PEs seem to increase in copy number as time passes from the original invasion. Nonetheless, some PEs will accumulate differences and will change the cytotype of the infected line: if most PEs are inactive, then they might change from a P-type, able to cause HD, to an intermediate cytotype, which might not cause full HD. One exception to this dearth of knowledge is the follow-up of PEs in experimental cages of *D. simulans* after experimental invasions of PEs at different temperatures. At hot temperatures (mean: ~23ºC), PEs spread rapidly from 1.79 copies per genome but to an average of 31.7 copies per genome after just 20 generations. At this point, the infection plateaued in terms copy number [53]. Since this experiment obtained PE information using poolSeq it is impossible to determine whether this stabilization was truly caused by frequency increase across individuals or an increase in the

number of copies per genome. Some of these PEs were internally deleted, which shows how quickly PEs can degenerate after an invasion and calls into questions whether the PEs invasions of the *paulistorum* and *stutervantii* group, groups that show highly degenerated PEs, are truly as old as previously proposed. The pattern was slightly different at cold temperatures. While there also a monotonic increase in the total number of PE copies per population at cool temperatures (mean: ~15ºC) over 40 generations, there was no plateau even after 40 generations of evolution and copy number continued to increase. This is intriguing because PEs do not cause HD at temperatures lower than 24ºC, so one would expect less fitness costs to the expansion of PEs in populations that would not show HD and thus a faster increase in the number of copies at a population than at higher temperatures who would experience the deleterious effects of HD. This last result means that there are phenotypic aspects of PEs that remain unknown, namely, that in any model that assesses the evolution of TEs, and PEs in particular, needs to incorporate temperature/environment as an important factor to explain evolutionary dynamics.

Notably, also within 20 generations, a piRNA system to counteract the effect of PEs also emerged within these experimental populations, suggesting that some P-elements transposed into piRNA producing loci (i.e. piRNA cluster), facilitating the production of piRNAs complementary to the PEs [53]. These results suggest that the dynamics of PE invasions are very rapid and that even a window of 20 years (as it is the timeline of the *D. simulans* and *D. melanogaster* natural invasions) might to be too long for our understanding of TE invasions in nature. Other studies of the importance of PEs in nature are concordant with this observation. A drastic example of this consists of a TE insertion in the first intron of the gene *cortex* of *Biston betularia* increased rapidly in frequency right after the industrial revolution [54]. The insertion changes the wing color of this moth from white to black, which aids moths camouflage in darker

trees, which became more common after industrialization. In longitudinal studies of genome evolution in maize, some lines showed dramatic decreases in genome size, losing an average of 398 Mb from their genomes over just six generations due to the purging of TEs [55]. Collectively these results indicate that TEs and genome interactions are highly dynamic and that longitudinal studies at a fine scale are sorely needed.

The study of HD has historically acknowledged the importance of moving away from the paradigm of HD being defined by the presence vs. absence of PEs in the genome [9]). Since the discovery of PEs, several cytotypes have been reported to exist in *D. melanogaster*. P and M (PE free) strains, show coupled paternal induction and maternal repression of PEs. While P strains show both, M strains exhibit neither phenotype. Three more cytotypes do exist; Q, P′ and KP. In these three cases, paternal induction from maternal repression is decoupled. Q strains repress but do not induce dysgenesis, making them fully fertile with P and M strains in both directions of crossing (Kidwell 1979). On the other end of the spectrum, P′ strains can induce but do not robustly repress hybrid dysgenesis even in crosses within their own type (Quesneville and Anxolabéhère 1998). KP strains harbor PEs with partially deleted internal sequences and can suppress dysgenesis but do not induce it. KP strains are common as PEs seem to be particularly prone to lose functional segments that are required for the elements to move. Our study addresses a different facet of variation in HD related to the number of PEs, number of PE copies in the paternal genome and finds support for its importance. Nonetheless, there are other factors that are likely to be important in determining how strong HD is in intra and interspecific crosses. We focused on complete PEs (of the P cytotype) but the number of copies of different cytotypes might be of importance in *D. simulans* as well.

A natural follow-up question involves testing whether the strength of HD increases with genetic distance. Previous experiments in our lab have shown that PEs increase some barriers between species stronger in the *simulans* species group. The experiments we present in this report involve only intraspecific crosses information we can gather across interspecific crosses would be limited. *D. simulans* can hybridize with six of the eight other species in the *melanogaster* species subgroup [56]. Nonetheless, only two of these crosses yield fertile progeny: the crosses with *D. sechellia* and *D. mauritiana*. These two species, however, form a hard polytomy with *D. simulans* and the genetic distances between the three species are almost identical [57]. This curtails the possibility of using these two hybridizations to assess the effect of PEs along a continuous range of genetic divergences. All the other hybridizations, including the ones with *D. melanogaster*, yield sterile progeny (i.e., with already atrophied gonads) and for that reason they are not conducive to the study of HD.

PEs are arguably the best characterized transposable element, not only in their molecular function but also in the phenotypes they can induce. Here we study the importance of paternal copy number for HD in *D. melanogaster* and *D. simulans*, expanding previous studies that have found a relationship between these two traits [13]. The genomic features that might affect the strength of HD, besides the presence and absence of PEs remain largely unknown (but see [50]). A precise quantification of the frequency of PEs, and TEs in general, as well as an overlap in their effects across different species is still in its infancy [20,24]. The assessment of allele frequencies of different PE insertions in natural populations will serve to inform the forces that foster or hamper the spread of PEs across populations and across species. The number of PEs in the parental genome is correlated with the strength of HD but it is not the only factor that determines the degree of the phenotypes experienced. Polymorphism at the gene Bruno and other

loci associated with TE regulation within the Piwi-RNA pathway can have drastic effects as well, but the interaction between copy-number and distinct natural polymorphisms remains unexplored [50,58].

TABLE 1. The presence of PEs affects the number of ovaries in F1 *D. simulans* female (from intraspecific matings) at 23ºC and 29ºC. *N* is the number of dissected females that produced the means (percentage of females mated) and standard deviations (SD). The last four columns show pairwise comparisons as 4 × 4 matrices for each cross. The upper triangular matrix shows the Z value from an approximate Two-Sample Fisher-Pitman Permutation Test (9,999 permutations). The lower triangular matrix shows the P-value associated to the comparison. Only pairwise comparisons with $P < 0.008$ were considered significant.

23ºC							
Cross	***N***	**Mean # of Ovaries**	**SD**	**Pairwise comparisons**			
				♀ **PE$^+$** × ♂ **PE$^+$**	♀ **PE$^+$** × ♂ **PE$^-$**	♀ **PE$^-$** × ♂ **PE$^+$**	♀ **PE$^-$** × ♂ **PE$^-$**
♀ **PE$^+$** × ♂ **PE$^+$**	40	1.800	0.564	*	1.897	-3.669	2.188
♀ **PE$^+$** × ♂ **PE$^-$**	40	1.962	0.194	0.071	*	-5.536	1.247
♀ **PE$^-$** × ♂ **PE$^+$**	40	1.147	0.821	1×10^{-4}	$< 1 \times 10^{-10}$	*	5.234
♀ **PE$^-$** × ♂ **PE$^-$**	40	2.000	0.000	0.054	0.493	$< 1 \times 10^{-10}$	*
29ºC							
Cross	***N***	**Mean**	**SD**	**Pairwise comparisons**			

				♀ **PE$^+$** × ♂ **PE$^+$**	♀ **PE$^+$** × ♂ **PE$^-$**	♀ **PE$^-$** × ♂ **PE$^+$**	♀ **PE$^-$** × ♂ **PE$^-$**
♀ **PE$^+$** × ♂ **PE$^+$**	40	1.675	0.616	*	0.278	4.466	3.144
♀ **PE$^+$** × ♂ **PE$^-$**	40	1.712	0.637	0.869	*	-4.88	2.755
♀ **PE$^-$** × ♂ **PE$^+$**	40	0.853	0.744	$< 1 \times 10^{-10}$	$< 1 \times 10^{-10}$	*	6.449
♀ **PE$^-$** × ♂ **PE$^-$**	40	2.000	0.000	6.001×10^{-4}	3.3×10^{-3}	$< 1 \times 10^{-10}$	*

TABLE 2. Ovariole number in F1 *D. simulans* females depends on PE status of the parents. Fully-factorial linear models for each temperature shows that female fecundity is affected by the interaction between the mother and father genotype (i.e., whether the parent harbors PEs). Df: degrees of freedom.

23ºC					
	Df	**Sum of the squared differences**	**Mean squared error**	**F-value**	**P-value**
mother	1	228.3	228.28	6.9789	0.009128
father	1	767.6	767.57	23.4660	3.155×10^{-6}
mother:line	5	332.7	66.55	2.0345	0.076998
father:line	3	210.3	70.09	2.1428	0.097274
mother:father	1	261.4	261.38	7.9907	0.005348
mother:father:line	2	297.4	148.70	4.5460	0.012123
mother:line1:father	1	34.4	34.35	1.0502	0.307114
mother:line1:father:line	2	153.0	76.51	2.3390	0.099957
Residuals	149	4873.8	32.71		
29ºC					
	Df	**Sum of the squared differences**	**Mean squared error**	**F-value**	**P-value**

mother	1	68.1	68.15	2.9049	0.0903966
father	1	271.9	271.88	11.5888	0.0008526
mother:line	5	566.8	113.35	4.8315	0.0004007
father:line	3	306.5	102.15	4.3542	0.0056833
mother:father	1	638.2	638.25	27.2051	6.022×10^{-7}
mother:father:line	2	347.7	173.86	7.4108	0.0008545
mother:line1:father	1	193.6	193.59	8.2518	0.0046659
mother:line1:father:line	2	288.0	144.02	6.1386	0.0027427
Residuals	149	3495.6	23.46		

TABLE 3. Pairwise comparisons for the average number of ovarioles at two different temperatures between the four possible genotypes. *N* is the number of dissected females that produced the means (average number of ovarioles per females) and standard deviations (SD). The last four columns show pairwise comparisons as 4 × 4 matrices for each cross. The upper triangular matrix shows the t value from a HSD Tukey test following a linear model that includes only the interaction between fixed effects. The lower triangular matrix shows the P-value associated to the comparison. Only pairwise comparisons with $P < 0.008$ were considered significant.

29ºC							
Cross	***N***	**Mean**	**SD**	**Pairwise comparisons**			
				♀ **PE$^+$** × ♂ **PE$^+$**	♀ **PE$^+$** × ♂ **PE$^-$**	♀ **PE$^-$** × ♂ **PE$^+$**	♀ **PE$^-$** × ♂ **PE$^-$**
♀ **PE$^+$** × ♂ **PE$^+$**	40	18.825	4.733	*	-0.099	4.614	-1.038
♀ **PE$^+$** × ♂ **PE$^-$**	41	18.952	4.400	1.000	*	4.769	-0.952
♀ **PE$^-$** × ♂ **PE$^+$**	40	12.825	6.594	3.94×10^{-5}	2.30×10^{-5}	*	-5.652
♀ **PE$^-$** × ♂ **PE$^-$**	40	20.175	7.125	0.727	0.777	$< 1 \times 10^{-5}$	*
23ºC							

Cross	*N*	Mean	SD	Pairwise comparisons			
				♀ **PE$^+$** × ♂ **PE$^+$**	♀ **PE$^+$** × ♂ **PE$^-$**	♀ **PE$^-$** × ♂ **PE$^+$**	♀ **PE$^-$** × ♂ **PE$^-$**
♀ **PE$^+$** × ♂ **PE$^+$**	40	23.450	6.477	*	-1.218	4.075	-1.417
♀ **PE$^+$** × ♂ **PE$^-$**	40	25.000	4.049	0.617	*	5.342	-0.216
♀ **PE$^-$** × ♂ **PE$^+$**	40	18.200	6.182	3.8×10^{-4}	$< 1 \times 10^{-4}$	*	-5.492
♀ **PE$^-$** × ♂ **PE$^-$**	40	25.275	6.089	0.491	0.996	$< 1 \times 10^{-4}$	*

Figure 1: Number of ovarioles observed in PE+ and PE- females as they age. Each isoline is represented by 10 individuals at each age (days). The red line shows the linear regression for 5 PE- lines. The blue line shows the linear regression for 5 PE+ lines. We found no difference in the intercept or the slope of the two regressions, which indicates that PE elements have no discernable effect on the rate of reproductive senescence.

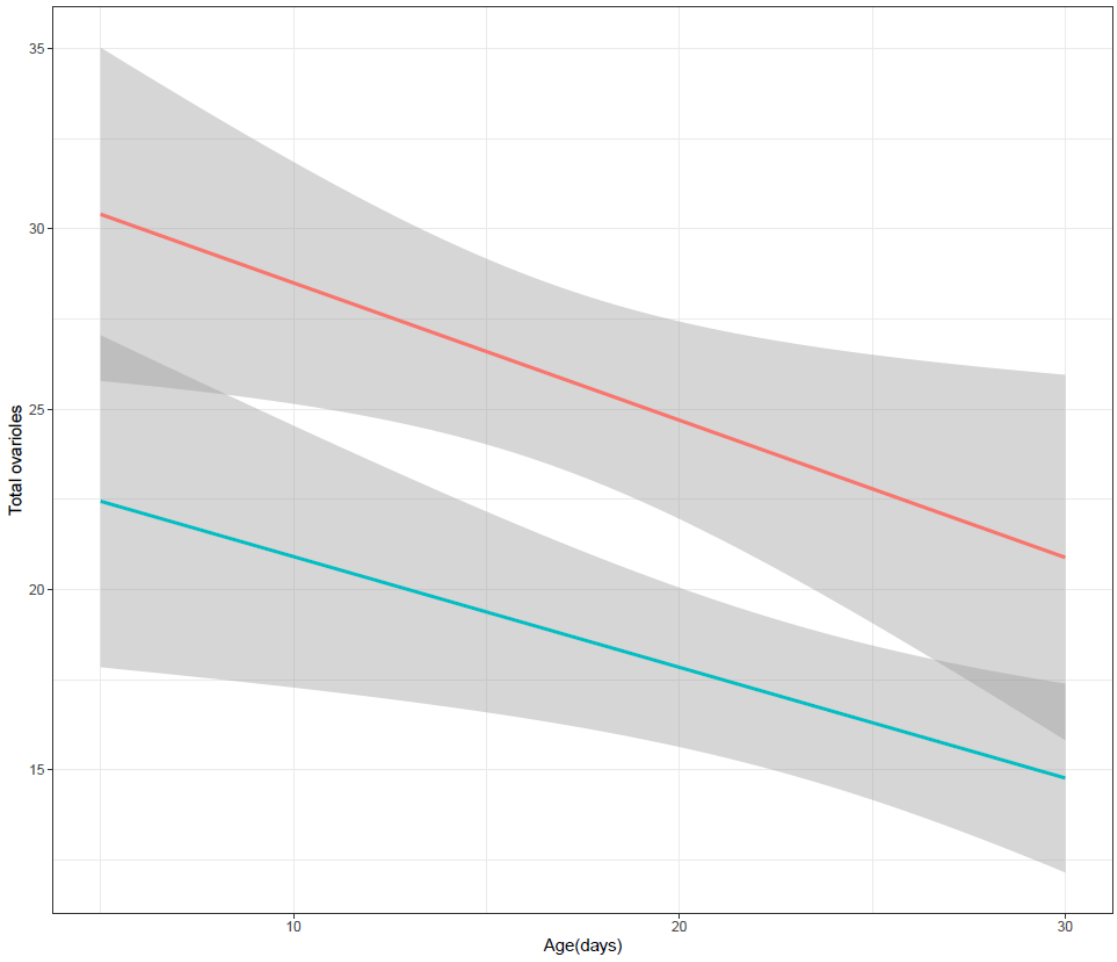


TABLE 4. ♀ PE$^-$ × ♂ PE$^+$ male fertility in *D. simulans* is lower than the fertility of the other three possible crosses genotype at 29ºC. Fertility was scored as the number of progeny produced when a male of each of the four shown genotypes was mated to a PE- *D. simulans* female. *N* is the number of dissected females that produced the means (average number of ovarioles per females) and standard deviations (SD). The last four columns show pairwise comparisons as 4 × 4 matrices for each cross. The upper triangular matrix shows the t value from a HSD Tukey test following a linear model that includes only the interaction between fixed effects. The lower triangular matrix shows the P-value associated to the comparison. Only pairwise comparisons with P < 0.008 were considered significant.

29ºC							
Cross	***N***	**Mean**	**SD**	**Pairwise comparisons**			
				♀ PE$^+$ × ♂ PE$^+$	**♀ PE$^+$ × ♂ PE$^-$**	**♀ PE$^-$ × ♂ PE$^+$**	**♀ PE$^-$ × ♂ PE$^-$**
♀ PE$^+$ × ♂ PE$^+$	110	40.63889	3.993319	*	-1.154	17.857	-0.344
♀ PE$^+$ × ♂ PE$^-$	120	41.45098	3.502892	0.656	*	20.942	0.820
♀ PE$^-$ × ♂ PE$^+$	120	28.16981	8.469758	$<1 \times 10^{-4}$	$<1 \times 10^{-4}$	*	-18.878
♀ PE$^-$ × ♂ PE$^-$	119	40.89344	3.856533	0.986	0.844	$<1 \times 10^{-4}$	*

TABLE 5. The number of PEs in the paternal genome reduces the number of mean ovarioles per ovary in F1s from PE- × PE+ in two Drosophila species and at two different temperatures.

Species	Temperature	Estimate	Std. Error	t value	P
D. melanogaster	23ºC	-0.34857	0.09864	-3.534	0.00048
D. melanogaster	29ºC	-0.36756	0.06714	-5.475	9.8×10^{-8}
D. simulans	23ºC	-0.58503	0.09871	-5.927	5.85×10^{-9}
D. simulans	29ºC	-0.4093	0.0671	-6.10	2.17×10^{-9}

TABLE 6. Male fertility is negatively affected by the number of PEs in the paternal genome. We fitted four linear models (2 species × 2 temperatures) and found that increasing the number of PEs leads to a reduction in male fertility in PE-/PE+ F1 males of *D. simulans* and *D. melanogaster*.

Species	Temperature	Estimate	Std. Error	t value	P
D. melanogaster	23ºC	-0.2955	0.1445	-2.045	0.0418
D. melanogaster	29ºC	-0.3236	0.1381	-2.344	0.0198
D. simulans	23ºC	-0.5928	0.1387	-4.273	2.32×10^{-5}
D. simulans	29ºC	-0.6311	0.1161	-5.435	8.65×10^{-8}

REFERENCES

1. Kent T V, Uzunović J, Wright SI. Coevolution between transposable elements and recombination. *Philos Trans R Soc B Biol Sci*. 2017;372(1736):20160458. doi:10.1098/rstb.2016.0458

2. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. Copenhaver GP, ed. *PLoS Genet*. 2011;7(12):e1002384. doi:10.1371/journal.pgen.1002384

3. Schnable PS, Ware D, Fulton RS, et al. The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science (80-)*. 2009;326(5956):1112-1115. doi:10.1126/science.1178534

4. Stapley J, Santure AW, Dennis SR. Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Mol Ecol*. 2015;24(9):2241-2252. doi:10.1111/mec.13089

5. Vicient CM, Casacuberta JM. Impact of transposable elements on polyploid plant genomes. *Ann Bot*. 2017;120(2):195-207. doi:10.1093/aob/mcx078

6. Staton SE, Bakken BH, Blackman BK, et al. The sunflower (Helianthus annuus L.) genome reflects a recent history of biased accumulation of transposable elements. *Plant J*. 2012;72(1):142-153. doi:10.1111/j.1365-313X.2012.05072.x

7. Song MJ, Schaack S. Evolutionary Conflict between Mobile DNA and Host Genomes. *Am Nat*. 2018;192(2):263-273. doi:10.1086/698482

8. Werren JH. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc Natl Acad Sci*. 2011;108(Supplement_2):10863-10870. doi:10.1073/pnas.1102343108

9. Kelleher ES. Reexamining the P-element invasion of Drosophila melanogaster through the lens of piRNA silencing. *Genetics*. 2016;203(4):1513-1531. doi:10.1534/genetics.115.184119

10. Quesneville H, Anxolabéhère D. Dynamics of Transposable Elements in Metapopulations: A Model of P Element Invasion in Drosophila. *Theor Popul Biol*. 1998;54(2):175-193. doi:10.1006/TPBI.1997.1353

11. Montgomery E, Charlesworth B, Langley CH. A test for the role of natural selection in the stabilization of transposable element copy number in a population of Drosophila melanogaster. *Genet Res (Camb)*. 1987;89(5-6):435-445. doi:10.1017/S0016672308009634

12. Menne D. Effects of P Element Insertions on Quantitative Traits in Drosophila melanogaster Trudy. 1977;312(1):301-312. doi:10.1016/S0168-9525(98)80015-2

13. Srivastav SP, Kelleher ES. Paternal Induction of Hybrid Dysgenesis in *Drosophila melanogaster* Is Weakly Correlated with Both *P* -Element and *hobo* Element Dosage. *G3: Genes|Genomes|Genetics*. 2017;7(May):g3.117.040634. doi:10.1534/g3.117.040634

14. Engels WR, Preston CR. Components of hybrid dysgenesis in a wild population of

Drosophila melanogaster. *Genetics*. 1980;95(1):111-128. http://www.genetics.org/content/genetics/95/1/111.full.pdf. Accessed April 30, 2018.

15. Kidwell MG, Kidwell JF, Sved JA. Hybrid Dysgenesis in Drosophila melanogaster: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination. *Genetics*. 1977;86(4):813-833. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1213713&tool=pmcentrez&rendertype=abstract.

16. Michalak P. Epigenetic, transposon and small RNA determinants of hybrid dysfunctions. *Heredity (Edinb)*. 2009;102(1):45-50. doi:10.1038/hdy.2008.48

17. Tasnim S, Kelleher ES. p53 is required for female germline stem cell maintenance in P - element hybrid dysgenesis. *doi.org*. 2017:204800. doi:10.1101/204800

18. Carnelossi EAG, Lerat E, Henri H, Martinez S, Carareto CMA, Vieira C. Specific activation of an I-like element in drosophila interspecific hybrids. *Genome Biol Evol*. 2014;6(7):1806-1817. doi:10.1093/gbe/evu141

19. McManus CJ, Coolon JD, Duff MO, Eipper-Mains J, Graveley BR, Wittkopp PJ. Regulatory divergence in Drosophila revealed by mRNA-seq. *Genome Res*. 2010;20(6):816-825. doi:10.1101/gr.102491.109

20. Serrato-Capuchina A, Matute DR. The role of transposable elements in speciation. *Genes (Basel)*. 2018;9(5):254. doi:10.3390/genes9050254

21. Kofler R, Hill T, Nolte V, Betancourt AJ, Schlötterer C. The recent invasion of natural Drosophila simulans populations by the P-element. *Proc Natl Acad Sci*. 2015;2015(28):201500758. doi:10.1073/pnas.1500758112

22. Sturtevant AH. Genetic Studies on Drosophila Simulans. I. Introduction. Hybrids With Drosophila Imelanogaster. *Genetics*. 1920;5:488-500. doi:10.1002/jez.b.21128

23. Ranz JM, Namgyal K, Gibson G, Hartl DL. Anomalies in the expression profile of interspecific hybrids of Drosophila melanogaster and Drosophila simulans. *Genome Res*. 2004;14(3):373-379. doi:10.1101/gr.2019804

24. Watson ET, Demuth JP. The role of transposable elements in speciation. December 2013. https://uta.influuent.utsystem.edu/en/publications/the-role-of-transposable-elements-in-speciation. Accessed June 8, 2018.

25. Marco R, Valentina P, Etienne G, Cristian T, Alessio B. Transposable Elements Activity is Positively Related to Rate of Speciation in Mammals. *J Mol Evol*. May 2018:1-8. doi:10.1007/s00239-018-9847-7

26. Matute DR, Ayroles JF. Hybridization occurs between Drosophila simulans and D. sechellia in the Seychelles archipelago. *J Evol Biol*. 2014;27(6):1057-1068. doi:10.1111/jeb.12391

27. Kofler R, Hill T, Nolte V, Betancourt AJ, Schlötterer C. The recent invasion of natural *Drosophila simulans* populations by the P-element. *Proc Natl Acad Sci*. 2015;112(21):6659-6663. doi:10.1073/pnas.1500758112

28. Eanes WF, Wesley C, Hey J, Houle D, Ajioka JW. The fitness consequences of P element insertion in Drosophila melanogaster. *Genet Res*. 1988;52(01):17. doi:10.1017/S0016672300027269

29. Arkhipova I, Meselson M. Deleterious transposable elements and the extinction of asexuals. *BioEssays*. 2005;27(1):76-85. doi:10.1002/bies.20159

30. Comeault AA, Serrato Capuchina A, Turissini DA, McLaughlin PJ, David JR, Matute DR. A nonrandom subset of olfactory genes is associated with host preference in the fruit fly Drosophila orena. *Evol Lett*. 2017;63:16.

31. Cooper BS, Sedghifar A, Nash WT, Comeault AA, Matute DR. A Maladaptive Combination of Traits Contributes to the Maintenance of a Drosophila Hybrid Zone. *Curr Biol*. 2018. doi:10.1016/j.cub.2018.07.005

32. Hill T, Schlötterer C, Betancourt AJ. Hybrid Dysgenesis in Drosophila simulans Associated with a Rapid Invasion of the P-Element. *PLoS Genet*. 2016;12(3):1-17. doi:10.1371/journal.pgen.1005920

33. Wong LC, Schedl P. Dissection of Drosophila ovaries. *J Vis Exp*. 2006;8(1):52. doi:10.3791/52

34. Schoville SD, Barreto FS, Moy GW, Wolff A, Burton RS. Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod Tigriopus californicus. *BMC Evol Biol*. 2012;12(1):170. doi:10.1186/1471-2148-12-170

35. Venables WN (William N., Ripley BD, Venables WN (William N). Modern Applied Statistics With S. *Technometrics*. 2003;45(1):111-111. doi:10.1198/tech.2003.s33

36. Fox J, Sanford W. *Car: An R Companion to Applied Regression (R Package).*; 2011. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion. Accessed June 12, 2018.

37. R-Core-Team. R: A language and environment for statistical computing. *JComputGraphStat*. 2013;5(3). http://www.r-project.org/.

38. Hothorn T, Hornik K, Van De Wiel MA, Zeileis A. A lego system for conditional inference. *Am Stat*. 2006;60(3):257-263. doi:10.1198/000313006X118430

39. Khurana JS, Wang J, Xu J, et al. Adaptation to P element transposon invasion in drosophila melanogaster. *Cell*. 2011;147(7):1551-1563. doi:10.1016/j.cell.2011.11.042

40. Lobell AS, Kaspari RR, Serrano Negron YL, Harbison ST. The Genetic Architecture of Ovariole Number in Drosophila melanogaster: Genes with Major, Quantitative, and Pleiotropic Effects. *G3 (Bethesda)*. 2017;7(7):2391-2403. doi:10.1534/g3.117.042390

41. Crawley MJ. *GLIM for Ecologists*.; 1993.

42. Kuznetsova A, Brockhoff PB, Christensen RHB. Tests in Linear Mixed Effects Models. 2015. http://cran.uib.no/web/packages/lmerTest/lmerTest.pdf. Accessed May 6, 2018.

43. Sundar Dorai-Raj A, Sundar Dorai-Raj M. The binom Package Title Binomial Confidence Intervals For Several Parameterizations Description Constructs confidence intervals on the

probability of success in a binomial experiment via several parameterizations. 2006. http://ftp.auckland.ac.nz/software/CRAN/doc/packages/binom.pdf. Accessed June 8, 2018.

44. Matute DR, Coyne JA. Intrinsic reproductive isolation between two sister species of drosophila. *Evolution (N Y)*. 2010;64(4):903-920. doi:10.1111/j.1558-5646.2009.00879.x

45. Hill T, Schlotterer C, Betancourt AJ. Hybrid dysgenesis in Drosophila simulans due to a rapid global invasion of the P-element. *PLoS Genet*. 2016:1-17. doi:10.1101/031781

46. Schnebel EM, Grossfield J. ANTAGONISTIC PLEIOTROPY: AN INTERSPECIFIC DROSOPHILA COMPARISON. *Evolution (N Y)*. 1988;42(2):306-311. doi:10.1111/j.1558-5646.1988.tb04134.x

47. Matute DR, Novak CJ, Coyne JA. TEMPERATURE-BASED EXTRINSIC REPRODUCTIVE ISOLATION IN TWO SPECIES OF DROSOPHILA. *Evolution*. 2009;63(3):595-612.

48. Stanley SM, Parsons PA, Spence GE, Weber L. Resistance of species of the Drosophila melanogaster subgroup to environmental extremes. *Aust J Zool*. 1980. doi:10.1071/ZO9800413

49. Kelleher ES, Barbash DA. Analysis of piRNA-mediated silencing of active TEs in Drosophila melanogaster suggests limits on the evolution of host genome defense. *Mol Biol Evol*. 2013;30(8):1816-1829. doi:10.1093/molbev/mst081

50. Kelleher ES, Jaweria J, Akoma U, Ortega L, Tang W. QTL mapping of natural variation reveals that the developmental regulator bruno reduces tolerance to P-element transposition in the Drosophila female germline. *PLoS Biol*. 2018. doi:10.1371/journal.pbio.2006040

51. Signor SA, New FN, Nuzhdin S. A Large Panel of Drosophila simulans Reveals an Abundance of Common Variants. *Genome Biol Evol*. 2018;10(1):189-206. doi:10.1093/gbe/evx262

52. Srivastav SP, Kelleher ES. Paternal Induction of Hybrid Dysgenesis in Drosophila melanogaster Is Weakly Correlated with Both P -Element and hobo Element Dosage . *G3: Genes|Genomes|Genetics*. 2017. doi:10.1534/g3.117.040634

53. Kofler R, Senti K-A, Nolte V, Tobler R, Schlötterer C. Molecular dissection of a natural transposable element invasion. *Genome Res*. 2018;28(6):gr.228627.117. doi:10.1101/gr.228627.117

54. Van't Hof AE, Campagne P, Rigden DJ, et al. The industrial melanism mutation in British peppered moths is a transposable element. *Nature*. 2016;534(7605):102-105. doi:10.1038/nature17951

55. Roessler K, Muyle A, Diez CM, et al. The Genomics of Selfing in Maize (Zea mays ssp. mays): Catching Purging in the Act. *bioRxiv*. April 2019:594812. doi:10.1101/594812

56. Turissini DA, McGirr JA, Patel SS, David JR, Matute DR. The Rate Of Evolution Of Postmating-Prezygotic Reproductive Isolation In *Drosophila*. *bioRxiv*. May 2017:142059.

57. Garrigan D, Kingan SB, Geneva AJ, et al. Genome sequencing reveals complex speciation in the Drosophila simulans clade. *Genome Res*. 2012;22(8):1499-1511. doi:10.1101/gr.130922.111

58. Khurana JS, Wang J, Xu J, et al. Adaptation to P Element Transposon Invasion in Drosophila melanogaster. *Cell*. 2011;147(7):1551-1563. doi:10.1016/J.CELL.2011.11.042

Chapter 3: P-elements strengthen reproductive isolation within the *Drosophila simulans* species complex

Aim 3: PEs can cause reproductive isolation between PE+ and PE- lineages of the same species. However, it is unclear whether they can also contribute to the magnitude of RI in crosses between species. We use the simulans species complex to assess whether differences in PE load between *D. simulans* and its sister species, *D. sechellia*/*D. mauritiana,* contribute to F1 fitness.

INTRODUCTION

Gene flow is a homogenizing force that opposes genome divergence and species diversification. The development of reproductive isolation (RI), which reduces gene flow, signals that speciation is in process or has been completed [1,2] and allows for genetic variation to be partitioned in nature. For this reason, the genetic factors that determine the strength of reproductive isolation between species in nature are a central focus of studies throughout speciation research. Barriers to gene flow can be categorized depending on where in the reproductive cycle they occur [3,4]. Prezygotic isolation includes all RI traits that occur before a zygote is formed and include ecological, behavioral and gametic incompatibilities [4]. Postzygotic barriers occur after a hybrid zygote is formed and include phenotypes as extreme as hybrid sterility and inviability [5,6], but also more nuanced traits, such as hybrid behavioral defects [7,8] and delays in development. Prezygotic isolation traits seem to evolve faster between species and have commonly been implicated in setting the speciation process in motion [9–12]. Nonetheless, prezygotic isolation barriers are not impervious to the effect of gene flow, as

species that only show behavioral isolation are more likely to collapse into a single gene pool than species that also show some postzygotic isolation [13]. Postzygotic isolation traits tend to evolve slower than prezygotic traits and are thought to be crucial in the persistence of species in the face of gene flow once they come into secondary contact. Both prezygotic and postzygotic isolation are important for the origin and persistence of species, and their relative importance likely differs between different taxa [11].

The last three decades have seen significant progress on identifying alleles involved in isolation (reviewed in [14,15]). Several hypotheses have emerged regarding the types of alleles involved in species divergence. First, intrinsic RI is commonly caused by epistatic interactions that go awry in hybrids (i.e., the Dobzhansky Muller model). In some (but not all) instances, selection causes the divergence of epistatic partners. In this case, RI is a by-product of local adaptation [1,14,15]. Genetic conflict can also lead to hybrid sterility [16]; a handful of cases in which a single allele can cause hybrid sterility and segregation distortion support this idea [17,18]. Some alleles involved in RI show the signature of positive selection at the molecular level, lending some support to these hypotheses [19–21]. Gene movement within a genome can also lead to hybrid incompatibilities, as crosses between individuals with chromosomal translocations can produce some aneuploid progeny (Theory: [19–21], Empirical: [22–25]).

A different hypothesis is that particular types of molecular changes result in deleterious interactions in hybrids not at the individual loci level but at the genome level (i.e., ‘genome clashes’), which might result in RI [29–31]. Transposable elements (TEs) are a potential candidate to cause such clashes. TEs are repetitive genetic units found ubiquitously across life and have been regularly linked to molecular and phenotypic novelty [32,33]. TEs can facilitate rapid genotypic change by causing genic interruptions, duplications, or gene expression changes

through their transposition [34]. Derepression of TEs through a mismatch of regulatory systems might activate dormant transposons which in turn can have deleterious fitness implications in hybrids (see below). This prediction has been in place since 1984 [35] and the support for it has been mixed. In *Arabidopsis* hybrids, seed viability is inversely correlated with the expression of pericentromeric retroelements (*ATHILA*) that are exclusively derepressed in hybrids. TE mobilization and proliferation are observed in rice [36] and sunflowers [37]. In animals, the evidence is less clear (reviewed in [38]) but TEs do cause hybrid dysfunction in some cases. Interspecific hybrids between Australian wallaby species (*Macropus eugenii* and *Wallabia bicolor*) show dramatically extended centromeres due to proliferation of additional centromeric material consisting of unmethylated retroelements [39,40]. These studies show that hybrids exhibit derepression of TEs but do not provide evidence that hybrids suffer fitness defects because of that derepression. A natural hypothesis, given their prevalence, is that TEs might strengthen reproductive isolating barriers between species [41]. The potential connection between TEs and RI is tantalizing, because TEs can readily move within and across genomes, generate phenotypic novelty [42–44], and show a higher activity in hybrids than in pure species [41].

One example of the connection between TEs and RI is P-elements (PEs). PEs in *Drosophila melanogaster* might arguably be the best studied TE in terms of molecular action and phenotypic consequences. The invasion of PEs in all populations of *D. melanogaster* occurred over the course of less than 40 years [45]. PEs in *D. melanogaster* cause a variety of phenotypes described as hybrid dysgenesis (HD), a syndrome of deleterious effects in F1s that occurs when the mother lacks the PE and the father carries a PE. F1s from these crosses show a suite of defects that include sterility, chromosomal breaks, and increased mutation rates [45–49].

Conversely, in crosses where the mother carries PEs, F1s do not show hybrid dysgenesis regardless of the genotype of the father. This asymmetry is explained by maternally mediated PE repression in the germoplasm, a germline-specific subclass of small non-coding RNAs, piRNAs (PIWI-interacting RNAs). This silencing mechanism is not specific to PEs and underlies the repression of multiple classes of TEs across multiple taxa (reviewed in Michalak, 2009).

While it is clear that PEs can cause hybrid dysgenesis within species, the role of PEs in strengthening RI in between well-formed species remains largely untested. This is an important question because nascent species come into contact often and if RI is not complete, lineages might merge into a single population [50]. Testing this hypothesis requires the study of PE effects on a group of species in which: *i*) some species harbor PEs while others do not, *ii*) species hybridize and produce viable progeny, and *iii*) since PEs mostly affect the reproductive fitness of F1s, hybrids must be at least partially fertile in order to propagate through the population. The closest known extant species to *D. melanogaster* are the three species of the *simulans* species complex, *D. simulans, D. sechellia, and D. mauritiana*. *D. melanogaster* can intercross with these three species but all crosses produce sterile hybrids, with only certain very rare strains being able to produce fertile F1 progeny [51–53]. PEs are also present in species of the *willinstoni* and *saltans* species group, but all species (and semi-species) within the group have been infected by PEs [54–56], preventing comparisons of PE status influencing the magnitude of RI in well-formed species.

The *simulans* species group is an ideal system to assess the potential effects of TEs in RI for at least three reasons. First, PEs recently invaded *D. simulans* [57] and are only present in some lines. While many *D. simulans* individuals collected after 2015 harbor functional PEs, neither *D. sechellia* nor *D. mauritiana* show any evidence for their presence, past or present [54].

Second, the three sister species, *D. simulans*, *D. sechellia* and *D. mauritiana*, produce fertile female and sterile male F1 hybrids when crossed in all their pairwise crosses [58,59]. Third, there is evidence of hybridization and admixture in nature between these three species. *D. simulans* and *D. sechellia* hybridize in the central islands of the Seychelles archipelago [60] and show signatures of gene exchange [18,61,62]. In spite of the absence of a contemporary hybrid zone between *D. mauritiana* and *D. simulans*, these two species also show evidence of introgression [61,63]. The *simulans* species groups is thus a powerful system to assess RI conferred by TEs in nature, because the natural variation in D. *simulans* lines allows for the test of whether populations with PEs show stronger RI towards the sister species than populations without PEs.

Here, we leverage the recent PE invasion of *D. simulans* to test whether the magnitude of pre- and post-zygotic reproductive isolation barriers was affected by the presence of PEs. We found that PEs have no effects on the magnitude of prezygotic isolation. PEs also have no effect on hybrid viability, a measure of somatic effects. On the other hand, PEs do have an effect on F1 fecundity, mirroring observations of hybrid dysgenesis described between genotypes within-species (where F1s from the cross where the father harbors PEs but the mother does not) show more pronounced fitness reduction than the reciprocal direction. This effect is only observable in F1 females, as F1 males from interspecific crosses between species of the *simulans* species complex are invariably sterile. Our results serve as a formal test on the putative role of PEs on RI between sister species. We discuss the next steps required to address the potential role of TEs on setting the speciation process in motion and on the persistence of species in the face of gene flow.

MATERIALS AND METHODS

Isofemale lines

All experiments reported in this manuscript used isofemale lines, and all have been previously reported. We used the three different species of the *simulans* species group. We used seven lines for each species. The details for each species lines are reported as follows.

D. simulans: *D. simulans* started showing evidence of PEs in the last 15 years and this TE has rapidly increased in frequency [54]. We used six PE+ lines collected in the island of Bioko in 2013, which showed evidence of P-elements: Riaba_1, Riaba_9, H9, Cascade1, H1, and LB1. DRM collected all these lines in the northern rim of the Lago Biao Caldera at 1,400 meters above sea level. These lines vary in their PE load (See: Dissertation Chapter 2). We also used seven lines with no evidence for the presence of PEs: MD19, MD98, NC105, MD99, Anro6, Anro10, and Anro43. All of these lines were collected using banana-yeast traps.

D. sechellia: Unlike *D. simulans*, *D. sechellia* lacks functional P-elements ([64], see below). We used 4 lines collected on the island of Mahé (Anse Royale Beach) by DRM and J.F. Ayroles in 2012: LD10, AnRo104, Denis100, CEnisNF10. Instead of banana-yeast traps, females were collected using bottles seeded with ripe *Morinda* fruits, due to resource preference. Additional details for the lines have been published elsewhere [60].

D. mauritiana: Similar to *D. sechellia*, *D. mauritiana* show no evidence of PE presence. We used four lines collected in the island of Mauritius by Womack in 2009: M13, R13, M12, and M32. Lines were collected using banana traps.

F1 Production

Hybrids were produced by placing 4 to 9 day-old virgins of each species in 30mL bottles with yeast. Due to females being choosier sex, we used a 1:2 ratio of females to males in each cross, with no more than 50 total flies per bottle. We allowed each cross the opportunity to mate for 6 days before removing all flies and inserting KimWipes treated with 0.5% propionic acid to provide a laying substrate. F1s produced were collected for up to 15 days after parentals were removed.

Male hybrid sterility. F_1 hybrids between all interspecific crosses in the *melanogaster* species subgroup are sterile at 23ºC [11]. We evaluated whether PE infection status made a difference on this barrier, though unlikely. Hybrid males were produced as described above at two different temperatures. Four to eight-day-old males from conspecific and heterospecific crosses were then lightly anesthetized with CO_2. Their testes were then extracted with forceps and mounted in chilled Ringer's solution. An average of 5 to 10 pairs of testes were mounted per slide. Using a microscope, we scored whether each pair of testes had motile sperm within 5 minutes of having mounted the samples. For each interspecific genotype combination, we scored sperm motility for 100 males per cross (398 combinations: 39,800 total males).

Female fecundity, generalities. One of the most pronounced phenotypes of the HD syndrome is reduced female fecundity as evidence by lower number of gonads (ovaries) and ovarioles per ovary. We scored these two traits in conspecific and heterospecific F1 females. We scored at least 20 females for each of the conspecific and heterospecific combinations which yielded a total of ~11,760 scored females (588 combinations × ~20 females per combination). The procedure for conspecific and heterospecific crosses was identical. We produced F1 females as described above (See 'F1 Production'). To score female fecundity, we counted the number of gonads in each F1 female (i.e., 0, 1 or 2 ovaries). We collected females aged 4 to 9 days post eclosion, anesthetized them with CO_2, and extracted their female reproductive tracts. While females were under anesthesia, we removed their reproductive tract using forceps [65]. The ovaries from each individual were then fixed on a precleaned glass slide with chilled *Drosophila* Ringer's solution (Cold Spring Harbor Protocols, Serrato-Capuchina et al. 2018). We counted the number of functional and atrophied gonads for each individual. An ovary was considered functional if it contained at least one ovariole. We also recorded the number of ovarioles (egg chambers) in each ovary. We used a Leica, S6E stereoscopic microscope to score all dissections. Table S2 shows the number of females dissected for each cross and each temperature. We used this dataset to study the effect of PEs on two phenotypes, ovary number and ovariole number. We describe each of these two traits as follows.

Female fecundity, ovary number. One of the classical traits of HD is a reduced number of ovaries in PE-/PE+ females. Our goal was to assess whether presence of PEs led to a more pronounced HD phenotype in heterospecifics than in conspecifics. A previous study reported that PE copy number has a negative effect on the number of ovaries [57].

First, we compared whether different crosses showed differences in the proportion of dysgenic females they produced. For crosses that showed more than 1% of dysgenic females, we used a proportion test (function '*prop.test*', R library *stats*). For instances where female dysgenesis was rare, we used a Two-Sample Fisher-Pitman Permutation Test (library *coin,* function '*oneway_test'*; [66]), adjusting the critical P-values for significance to account for multiple comparisons (8 comparisons).

For each of two types of interspecific crosses, we fitted a linear mixed model where the response of the model was whether a female showed evidence of dysgenesis or not, the type of cross (a cross was either conspecific or heterospecific) was a fixed effect, and the number of PE copies in the *D. simulans* genome was a continuous trait. We fitted a binomial regression using the function glmer in the library *lme4*. We included the interaction between these two effects to account for differences in HD caused by the number of PE copies and assessed its significance using a likelihood ratio test (function '*lrtest',* R library *lmtest*) by comparing a model with the interaction and a model without it. This tested whether the slope of the regression differed between same species F1s and heterospecific F1s by comparing the fully factorial model with one that did not have the interaction. The fully factorial model followed the form:

$$\text{ovary number} \sim \text{type of cross}_i + \text{PE copy number}_j + (\text{type of cross X PE copy number})_{ij} + \text{Error}_{ij}$$

To quantify the significance of each effect, we used a type III ANOVA (library *stats* [67]) in R. In total we fitted 8 linear models, one for each combination of cross and temperature (4 types of crosses × 2 temperatures). Additionally, we fitted linear models identical to the ones described above with the only difference being that instead of the number of PE copies, we

studied whether the presence or absence of PE elements affected whether a female was dysgenic or not. The full factorial model for this analysis followed the form:

ovary number ~ type of $cross_i$ + PE copy $number_j$ + $Error_{ij}$

Female fecundity, ovariole number. A second aspect of hybrid dysgenesis is a reduced number of ovarioles in PE-/PE+ females. We quantified whether there were differences in the mean number of ovarioles in within-species and interspecific hybrid F1 females. For these analyses, we excluded females that showed completely atrophied ovarioles. The procedure is similar to the one described above for ovary numbers (See '**Female fecundity, ovary number'**). We used the mean number of ovarioles per female, treating dysgenic ovaries as missing data to control for the effect of only having one ovary. We used a Poisson-distributed linear model (function '*glmer'*, library '*lme4'*) to fit a linear regression. We fitted two linear models, one for each type of interspecific cross. The mean number of ovarioles per female was the response, while the type of cross and the PE infection status were fixed effects. To fit the model, we approximated to the next nearest integer number of mean ovariole number. We included the interaction between these two effects to account for differences in HD caused by the number of PE copies and assessed its significance using a likelihood ratio test (function '*lrtest',* R library *lmtest*) by comparing a model with the interaction and a model without it. The full factorial model for this analysis followed the form:

ovariole number~type of $cross_i$ + PE copy $number_j$ + (type of cross X PE copy $number)_{ij}$ + $Error_{ij}$

Additionally, we fitted linear models identical to the ones described above with the only difference being that instead of the number of PE copies, we studied whether the presence or absence of PE elements affected the mean number of ovarioles per ovary per female. The full factorial model for this analysis followed the form:

ovariole number ~ type of $cross_i$ + PE copy $number_j$ + $Error_{ij}$

RESULTS

Male fertility

Consistent with previous work, all F1 males from all interspecific crosses (N >100 per cross), were completely sterile and produced no sperm.

Female fecundity

The classical phenotype in hybrid dysgenesis (HD) is a reduction in F1 female fecundity in crosses between PE- mothers and the PE+ fathers. We assessed whether the presence of PEs had an effect on hybrid female fecundity in two ways, counting the number of ovaries in females from crosses between PE+ and PE- individuals, and scoring the number of ovarioles in non-atrophied ovaries. We describe these results as follows:

Ovary number. The regressions of the number of ovaries against the number of PE copies in heterospecific crosses and conspecific crosses are shown in Figure 1. We studied whether the presence/absence of PEs in the paternal genome had an effect on the number of ovaries in conspecific crosses involving a *D. simulans* PE+ father and a *D. simulans* PE- mother, and in hybrid crosses involving a *D. simulans* PE+ father and a mother from a different species (i.e., *D. sechellia* or *D. mauritiana*). We did this in two ways. First, we studied whether PE- × PE+ F1s were more likely to be dysgenic than females from the PE- × PE- cross. Second, we measured the effect of the PE copy number in the paternal genome on the number of ovaries produced in each cross.

At 23ºC there is a small signal of HD in conspecific crosses as F1 females from the PE- × PE+ crosses were more likely to show atrophied ovaries than females from PE+ × PE- crosses (Table 1). At 29ºC, the signature of HD was even more pronounced as a higher proportion of F1 females from the PE- × PE+ crosses showed atrophied ovaries than females from PE+ × PE- crosses (Table 1). These results are concordant with the expectations from previous reports that show that HD is stronger at higher temperatures in conspecific crosses.

Second, we studied whether the presence of PEs in the maternal genome had an effect on the number of dysgenic females in hybrid crosses involving *D. simulans* PE+ and *D. simulans* PE- females and *D. sechellia* and *D. mauritiana* males. In general, all crosses where the *D. simulans* was the mother showed low proportion of dysgenic females; no cross showed more than 1% of dysgenic females (Table 2). Regardless of the cross and the temperature, the vast majority of conspecific and hybrid females had two ovaries (i.e., were non-dysgenic). We avoided doing pairwise comparisons where one of the crosses had no dysgenesis (NA in Table 2). Crosses between *D. simulans* PE+ females and *D. sechellia* males show a slightly higher

proportion of dysgenic hybrid females than crosses between *D. simulans* PE- and *D. sechellia* males. This result is puzzling and not consistent with HD but may be governed by the exceedingly low proportions of dysgenic females observed in this direction of the cross.

We furthered analyzed whether in addition to the number of dysgenic females in each cross, PEs in the paternal genome also affected the mean number of ovaries per female. This correlation has been previously observed in within-species HD syndromes in *D. melanogaster* [45] and *D. simulans* [68]. We compared the number of ovaries in females from *D. mauritiana* × *D. simulans* crosses to that of females from conspecific (*D. simulans* × *D. simulans*) crosses. The results at the two temperatures, 23ºC and 29ºC, were qualitatively similar (Figure 1A, B; Table 3). In both cases, the cross effect and the number of PEs were significant indicating that both the type of cross (i.e., conspecific vs heterospecific) and the number of PEs affected the number of ovarioles in conspecific and heterospecific matings. *D. simulans* × *D. simulans* F1 females are at least 2.5 times more likely than *D. mauritiana* × *D. simulans* females to be non-dysgenic at both 23ºC and 29ºC (Table 2). As the PE copy number in the paternal genome increases, the number of ovaries in the progeny decreases (23ºC: Coefficient estimate = -0.21330, SE = 0.02052, Z= -10.393, $P < 1 \times 10^{-10}$; 29ºC: Coefficient estimate = -0.29418, SE = 0.01516, Z= -19.401, $P < 1 \times 10^{-10}$). The rate of decrease in ovary number as the number of PEs increased in the paternal genome did not differ between conspecific and heterospecific crosses (i.e., the two regressions had a similar slope, Table 3) suggesting that the effect of PE copy number is similar for both types of crosses.

The comparisons of the number of ovaries in *D. sechellia* × *D. simulans* and *D. simulans* × *D. simulans* crosses yielded similar results (Figure 1A, B). The cross effect was significant indicating that both the type of cross and the number of PEs affected the number of ovarioles in

conspecific and heterospecific matings. *D. simulans* × *D. simulans* F1 females are at least twice more likely than *D. sechellia* × *D. simulans* females to be non-dysgenic at both 23ºC and 29ºC (Table 3). The PE copy number was also significant indicating the existence of a negative relationship between paternally inherited PEs and ovary number in the female PE-/PE+ progeny (23ºC: Coefficient estimate = -0.24387, SE = 0.01938, Z= -12.58, $P < 1 \times 10^{-10}$; 29ºC: Coefficient estimate = - 0.23846, SE = 0.01741, Z= - 13.694, $P < 1 \times 10^{-10}$). Unlike all other regressions, the interaction in the comparison at 29ºC was tangentially significant (Likelihood Ratio test; $\chi 2 = 4.072$, df=1, P=0.044) which suggests that the slope between conspecific and heterospecific crosses differs (Slope $_{sim \times sim}$ - slope $_{sech \times sim}$ = -0.06157, SD=0.03051, Z= -2.018). This in turn indicates that the effect of PEs in decreasing ovary number is more severe in some heterospecific crosses (*D. sechellia* × *D. simulans*) than in conspecifics (*D. simulans* × *D. simulans*) or other heterospecific crosses (*D. mauritiana* × *D. simulans*). These results suggest that hybrid females from crosses between PE- females and PE+ males suffer the effect of hybrid dysgenesis.

We also assessed whether the maternal number of PEs had an effect on ovary number on crosses involving *D. simulans* females. At 23ºC, all *D. simulans* × *D. sechellia* females (regardless of the PE-status of *D. simulans*) had two ovaries (i.e., we dissected almost 4,000 females and did not see a single dysgenic fly), strongly suggesting that there was no effect on the number of ovaries by the maternal number of PE copies. (Attempts to fit linear models to this dataset failed as the model never converged due to the low variability in the data.) We observed a similar—but not as dramatic— low level of dysgenic progeny in *D. simulans* × *D. mauritiana* crosses at both temperatures. The factors were not significant in any of these comparisons (Table 2, Figure 1C, D). This suggests that the conspecific and heterospecific crosses have similar

fecundities (at least in their number of ovaries), and that the PE copy number has no effect on the number of ovaries in this direction of the cross. Notably, the two effects, PE-copy number and type of cross, and the interaction between them were significant in *D. simulans* × *D. sechellia* crosses at 29ºC (Table 3). These results are mostly consistent with the expectation of HD in which females from crosses between PE+ females and PE- males have no deleterious effects.

The overall results from these analyses, which involve studying HD either as the proportion of females with atrophied gonads or the effect of the PE copy, are consistent with the expectations of the patterns of HD observed within species in *D. simulans* and *D. melanogaster,* in which females from crosses between PE- females and PE+ males show a reduction in the number of functional gonads, while females from crosses between PE+ females and PE- males tend to not produce dysgenic progeny. These results suggest that HD induced by PEs is also present in heterospecific matings.

Ovariole number. A second aspect of hybrid dysgenesis is that even when females show non-atrophied gonads, their ovaries might have fewer ovarioles than females from non-dysgenic crosses. Our hypothesis for this trait was similar to that described for ovary number: females from crosses between PE- females (either conspecific or heterospecific) and *D. simulans* PE+ males should show fewer mean number of ovarioles per ovary than females from the reciprocal crosses. In fact, the reciprocal cross should show no differences regardless of the PE status of the mother. First, we studied whether the presence of PEs in the paternal genome had a different effect on the mean number of ovarioles in hybrid crosses involving a *D. simulans* PE+ father (i.e., with *D. sechellia* and *D. mauritiana* females) than in conspecific crosses involving a *D. simulans* PE+ father and a *D. simulans* PE- mother. The results at 23ºC and 29ºC are

qualitatively similar. When we treated PEs as a two-level fixed effect (PE+ vs PE-), we found that the number of ovarioles per ovary per female is lower in *D. mauritiana* × *D. simulans* PE+ than in *D. mauritiana* × *D. simulans* PE- (Table 4). We observed the same pattern in *D. sechellia* × *D. simulans* heterospecific crosses. These results suggest that, similar to the observations from conspecific crosses, the presence of PEs in the paternal genome reduced the number of ovarioles in F1 females that result from heterospecific crosses.

Next, we studied whether progeny from *D. simulans* females and males from other species, showed reduced mean number of ovarioles. This is a test on whether maternal number of PE copies have an effect of the number of ovarioles in F1 females. We found that in three out of the four crosses, females from PE+/PE- crosses do not show reduced number of ovarioles. In one case, *D. simulans* × *D. sechellia*, females did show a reduced number of ovarioles per ovary at 29ºC. With the exception of this last case, these results are in line with the expectation of no reduced fecundity in this direction of the cross in conspecific crosses (Table 5).

We also studied whether the mean number of ovarioles per ovary per female was affected by the PE copy number in the paternal genome. The regressions of the number of the mean number of ovarioles per female against the number of PE copies in heterospecific crosses and conspecific crosses are shown in Figures 2 and 3. First, we compared the number of ovarioles in females from *D. mauritiana* × *D. simulans* crosses to that of females from conspecific, *D. simulans* × *D. simulans.* The results at the two temperatures, 23ºC and 29ºC were qualitatively similar (Table 5). Similar to the results on ovary number (see above), we found strong differences between conspecific and heterospecific crosses. *D. mauritiana* × *D. simulans* crosses had lower number of ovaries (23ºC: Coefficient estimate $_{sim \times sim}$ – Coefficient estimate $_{mau \times sim}$ = 0.202787, SE = 0.010140, Z = 20.00, P <1×10^{-10}; 29ºC: Coefficient estimate $_{sim \times sim}$ –

Coefficient estimate $_{sech \times sim}$ = 0.288953, SE = 0.011463, Z = 25.21, P <1 × 10^{-10}), and increases in PE copies led to decreases in ovarioles (23ºC: Coefficient estimate = -0.086018, SE = 0.001915, Z = -44.93, P <1 ×10^{-10}; 29ºC: Coefficient estimate = -0.113181, SE = 0.002368, Z = -47.80, P <1 ×10^{-10}). Unlike, ovary number, the interaction between PE-copies and cross was significant suggesting differences in the slope of the regression (Table 5) which in turns means that the PE-copy effect on ovariole number was more pronounced in the hybrid females than in the pure species temperatures (23ºC: Slope estimate$_{sim \times sim}$ - slope estimate$_{mau \times sim}$ = 0.072919, SE= 0.002589, Z= 28.17, P <1 ×10^{-10}; 29ºC: Slope estimate$_{sim \times sim}$ - slope estimate$_{mau \times sim}$ =, SE =, Z =, P <1 ×10^{-10}).

Comparisons of the number of the mean number of ovarioles between *D. sechellia* × *D. simulans* and *D. simulans* × *D. simulans* crosses yielded similar results at both 23ºC and 29ºC. The cross effect and the number of PEs were both significant indicating that both the type of cross and the number of PEs both affected the number of ovarioles in conspecific and heterospecific matings. Namely, *D. sechellia* × *D. simulans* crosses had lower number of ovaries (23ºC: Coefficient estimate $_{sim \times sim}$ – Coefficient estimate $_{sech \times sim}$ = 0.313862, SE = 0.010432, Z = 30.09, P <1 × 10^{-10}; 29ºC: Coefficient estimate $_{sim \times sim}$ – Coefficient estimate $_{sech \times sim}$ = 0.384615, SE = 0.011697, Z = 32.88, P <1 × 10^{-10}), and increases in PE copies led to decreases in ovarioles (23ºC: Coefficient estimate = -0.042500, SE = 0.001762, Z = -24.12, P <1 ×10^{-10}; 29ºC: Coefficient estimate = -0.058029, SE = 0.002079, Z = -27.92, P <1 ×10^{-10}). The interaction between these two effects in this case was significant (Table 2) which suggests that the slope between conspecific and heterospecific crosses differs at both temperatures (23ºC: Slope estimate$_{sim \times sim}$ - slope estimate$_{sech \times sim}$ = 0.028546, SE=0.002500, Z= 11.42, P <1 ×10^{-10}; 29ºC:

Slope estimate$_{sim \times sim}$ - slope estimate$_{sech \times sim}$ = 0.036049, SE = 0.003117, Z = 11.56, P <1 $\times 10^{-10}$). This result indicates that, just as it occurs in *D. mauritiana x D. simulans* crosses, the effect of PEs in decreasing ovariole number per ovary is more severe in *D. sechellia* $\times$ *D. simulans* crosses than in conspecific crosses. We found no difference in the slope of the regression between the number of ovarioles and PE copies between *D. mauritiana* $\times$ *D. simulans* and its reciprocal cross. These results indicate that HD is more pronounced in heterospecific crosses and the effect of PE copy number is stronger in heterospecific crosses than in conspecific crosses.

Second, we studied whether the presence of PEs in the maternal genome had an effect on the mean number of ovarioles per female involving *D. sechellia* and *D. mauritiana* females with *D. simulans* PE+ and *D. simulans* PE- males compared to conspecific crosses within *D. simulans*. Females from this direction of the cross—from either conspecific or heterospecific crosses—generally do not show reductions in the mean number of ovarioles regardless on the number of PEs in the genome of the *D. simulans* mother. Table 5 lists the results for the linear models that assessed the effect of the number of PE copies in the maternal genome fitted for each combination of cross and temperature and are consistent with the linear models that studied the effect of PE presence in the maternal genome (Table 4). With the exception of *D. simulans* $\times$ *D. sechellia* at 29ºC, increases in PE copy number in the maternal genome did not lead to reductions in the mean number of ovariole number per ovary. As mentioned above, the reasons behind the decrease in ovariole number in *D. simulans* $\times$ *D. sechellia* at 29ºC remain unknown.

Collectively, these results suggest that the magnitude of HD in the form of reduced number of ovarioles per ovary is higher in heterospecific crosses than in conspecific crosses when PEs are present in the genome of the father but not when they are present in the genome of the mother, consistent with the expectations of the phenomenon of PE-induced HD.

DISCUSSION

Hybrid dysgenesis in *Drosophila* is caused by PEs and remains one of the best understood cases of how TEs interact with their host genome. The vast majority of studies of PEs have focused on their effects on pure species and few have addressed the effects of PEs on interspecific crosses. We studied the role of PEs in RI between three closely related species of the *simulans* species complex by crossing *D. simulans* lines that vary in whether they carry PEs to its two sister species which do not carry PEs. We find that PE status of the father (and the number of copies) affected F1 female fertility in interspecific crosses. This effect is stronger than the effect of PEs in within-species crosses, suggesting that hybrids are more susceptible to suffer the effects of HD than pure species individuals.

The importance of PE invasions for ensuring or furthering the speciation process remains largely unexplored. One of the possible effects of TEs is the rapid onset of reproductive isolation between populations. HD dysgenesis is an example in which PE+ and PE- 'populations' can be moderately reproductively isolated. In the case of *D. melanogaster*, PEs have taken over most of the species ranges and no PE-free wildtype lines have been found since 1974 [49]. Since PEs have spread over the whole range of this species, and no PE-free population seems to exist, PEs have played no role in generating stable RI. In the case of *D. simulans*, it remains to be seen if PEs also overtake the whole geographic range but there is no evidence that PE+ and PE- populations might be isolated from each other. Understanding whether PEs can ensure differentiation is a question of whether the isolation provided by HD in PE-/PE+ F1s is strong enough to overcome the gene flow that might occur through the reciprocal direction (PE+/PE-).

PEs can also make existing reproductive isolating mechanisms stronger. One way in which TEs could foster and complete the speciation process would be by reducing the fitness of

hybrids. This would thus provide an indirect source of selection for prezygotic-isolation to be strengthened. The process in which speciation can be completed as a byproduct of selection against hybrids, coined reinforcement, seems to be a common step towards the completion of speciation in nature [69]. If PEs cause F1 interspecific hybrids to suffer fitness consequences due to HD, then natural selection could lead to an eventual decrease of gene flow between *D. simulans* and *D. sechellia* that coexist in the central islands of the Seychelles archipelago. Yet, *D. simulans* has not been reported in the Mauritius islands, the endemic range of *D. mauritiana*. Thus far, there is no evidence of reinforcement in the *simulans* species group. However, there is extensive variation in the magnitude of isolation between species in this group across all barriers that have been systematically studied, which in turn suggests there is potential for reinforcement to occur. Notably, PEs primarily affect the female progeny from the PE- mother and PE+ fathers, which would allow for extensive gene flow in the reciprocal direction of the cross. Reinforcement can be completed in the face of gene flow, but too much hybridization and introgression can hamper the effect of indirect natural selection favoring the completion of speciation [70] and will eventually lead to the collapse of the species boundary. In the *simulans* species group, the PE- × PE+ cross that would produce dysgenic individuals are crosses between island endemics (*D. sechellia* and *D. mauritiana*) and *D. simulans*, which occur rarely. On the other hand, matings between *D. simulans* females and the island endemic species are much more common [71]. Understanding whether PEs can induce reinforcing selection is a similar question to whether PEs can generate stable RI within species (see above) but with the key difference that while in intraspecific matings, mating choice tends to be weak, in interspecific crosses mating choice—and the asymmetry in the direction of the mating—is usually strong. The dynamics and potential involvement of PEs in reinforcement merits a formal theoretical treatment which

models the intraspecific fitness costs of a PE invasion (i.e., intraspecific HD), the likelihood of hybridization (including potential asymmetry), and the costs associated with hybridization (i.e., interspecific HD). This treatment also establishes under which circumstances the potential advantages of interspecific HD (through reducing gene flow) can outweigh the intraspecific HD costs, ultimately serving as a possible mechanism to facilitate speciation by reinforcement. Hybrids often result in the misregulation of transposable elements, a phenomenon not observed in the parental species. In whitefish for example, transposable elements reactivation is much more common in malformed F1s than in properly developed F1 embryos, strongly suggesting that TEs are involved in the massive dysregulation specific to the F1s [72].

Besides the effect of PEs on RI, an additional aspect that deserves consideration is whether hybridization facilitates the transmission of PEs across species boundaries. The rapid increase in PE frequency in both *D. simulans* and *D. melanogaster*, despite fitness costs, remains a puzzle. A similarly intriguing question is why PEs have not been found in *D. mauritiana* and *D. sechellia*. In the case of *D. mauritiana*, there is evidence for interspecific introgression with *D. simulans* [18,61,72]. This begs the question of why PEs have not been transferred from *D. simulans* to *D. mauritiana* while other alleles have crossed those species boundaries. One likely possibility is that hybridization and introgression is ancient and precedes the invasion of PEs in *D. simulans*. The case of *D. sechellia* is even more puzzling. Not only do *D. sechellia* and *D. simulans* show signatures of introgression [74], these two species form a contemporary hybrid zone in the central islands of the Seychelles archipelago [60]. Although as of 2014 *D. simulans* in the Seychelles Islands did not contain the PEs, there is precedence for gene exchange between the species [64]. This indicates that, at present, there is ample opportunity for gene exchange—and for interspecific transfer of PEs.

Introgressive hybridization cannot explain the invasion of PEs in *D. melanogaster*, as this species does not produce fertile hybrid progeny with any of the species where PEs have been found (i.e., species from the *willinstoni* and *saltans* group). In the case of *D. simulans*, invasions of PEs following hybridization with *D. melanogaster* is a very unlikely (but not impossible) event as some *D. simulans* mutants can produce weakly fertile progeny when they are mated to *Hmr- D. melanogaster* mutants [51,52]. These mutants seem to segregate at very low frequency in nature and thus the most parsimonious explanation is that PEs invaded *D. simulans* through a horizontal gene transfer event that did not involve hybridization. The precise mechanisms of transmission of TEs across species lines remain unknown, but exploring this within a well-documented element such as PEs can provide insight in this and other longstanding evolutionary questions regarding TEs. A prevailing hypothesis regarding horizontal transfer of PEs has involved mites, a natural predator for *Drosophila* eggs, through their syringe-like feeding transferring DNA between species, but the evidence for this animal-mediated horizontal gene transfer remains circumstantial (Houck et al., 1991). It is likely that multiple mechanisms play a role in transmission of PEs and only a robust systematic assessment of the relative frequency of PEs on different species of *Drosophila* and the putative vectors will reveal what is each mechanism's importance in the spread of PEs within and between species. Our results show that the dynamics of PEs, and possibly of TEs in general, should not only be addressed with a lens on the fitness effects that PEs have on within-species crosses, but also on the effects that they might have on interspecific hybrids.

TABLE 1. PE-/PE+ are more likely to produce dysgenic females than PE-/PE- crosses in conspecific and heterospecific crosses. Females with atrophied ovaries are more common in heterospecific than in conspecific crosses in this direction of the cross. Each row shows a type of cross. '*mother'* represents the conspecific or heterospecific female used in the cross, as indicated under "Cross". We used a 2-sample test for equality of proportions for this set of pairwise comparisons.

Cross	**23ºC**			**29ºC**		
	Proportion of dysgenic females		**χ^2, df=1**	**Proportion of dysgenic females**		**χ^2, df=1**
	♀ *mother* × ♂*sim-PE+*	♀ *mother* × ♂*sim-PE-*		♀ *mother* × ♂*sim-PE+*	♀ *mother* × ♂*sim-PE-*	
D. simulans × *D. simulans*	0.002	0.0170	10.299, $P = 1.331 \times 10^{-3}$	0.126	0.057	28.257, $P = 1.062 \times 10^{-7}$
D. mauritiana × *D. simulans*	0.149	0.001	152.5, $P < 1 \times 10^{-10}$	0.272	0.000	306.8, $P < 1 \times 10^{-10}$
D. sechellia × *D. simulans*	0.179	0.014	149.9, $P < 1 \times 10^{-10}$	0.260	0.024	223.67, $P < 1 \times 10^{-10}$

TABLE 2. PE+/PE- female progeny show low proportion of dysgenesis, both in conspecific and heterospecific crosses. Each row shows a type of cross. '*father'* means whether the cross is conspecific or heterospecific as indicated in "Cross" column, with the mother *D. simulans* PE status indicated first. Unlike the comparisons in Table 1 which used a Wald test (based on a $\chi 2$ test), here we used a permutations-based test (approximative Two-Sample Fisher-Pitman) because the number of dysgenic females was low in all crosses. NA refers to comparisons we did not perform as one or two values were zero.

Cross	23ºC			29ºC		
	Mean HD ♀ *sim-PE+* × ♂father	Mean HD ♀ *sim-PE-* × ♂father		Mean HD ♀ *sim-PE+* × ♂father	Mean HD ♀ *sim-PE-* × ♂father	Approximative Two-Sample Fisher-Pitman Permutation Test
D. simulans × *D. simulans*	4.082×10^{-3}	0.000	Z = 1.891, P = 0.121	1.020×10^{-3}	0.000	NA
D. simulans × *D. mauritiana*	2.041×10^{-3}	1.020×10^{-3}	NA	9.183×10^{-3}	0.000	NA
D. simulans × *D. sechellia*	0.000	0.000	NA	0.013	0.0173	Z = 3.1659, P = 2.2×10^{-3}

TABLE 3. PE copy number in the paternal genome affects the number of functional ovaries in conspecific and heterospecific produced F1 females. Each row shows the results for each cross at one of two temperatures (ºC). The linear models for each combination of cross and temperatures were binomial models fitted to compare the effect of paternally inherited PEs on conspecific and hybrid F1 females. The effect 'cross' had two levels (sim x sim and a heterospecific cross) while the PE copies effect was continuous and ranged from 0 (i.e., no PEs) to 11.58.

		Odds ratio		**Linear model**		
Cross	**ºC**	**Estimate**	**95%CI**	**PE copies**	**cross**	**PE copies × cross**
D. sechellia × *D. simulans*	23	13.702	8.684-22.922	158.35, $P < 1 \times 10^{-10}$	112.98, $P < 1 \times 10^{-10}$	2.644, P = 0.104
D. sechellia × *D. simulans*	29	2.395	1.68-3.787	187.514, $P < 1 \times 10^{-10}$	24.200, $P = 8.684 \times 10^{-7}$	3.926, P = 0.048
D. mauritiana × *D. simulans*	23	8.931	5.06-15.20	108.004, $P < 1 \times 10^{-10}$	72.817, $P < 1 \times 10^{-10}$	3.493, P = 0.062
D. mauritiana × *D. simulans*	29	2.077	1.523-3.164	376.417, $P < 1 \times 10^{-10}$	15.657 7.593e-05	7×10^{-4}, P = 0.979
D. simulans × *D. sechellia*	23	NA	NA	0.387, P = 0.534	1.6491 0.199081	1.332, P= 0.249

D. simulans × *D. sechellia*	29	1.453	0.826-2.600	0.628, P = 0.428	0.315, P = 0.575	NA
D. simulans × *D. mauritiana*	23	0.749,	0.147-3.410	2.498, P = 0.114	0.142, P = 0.706	1.235, P = 0.266
D. simulans × *D. mauritiana*	29	9.07	1.700-167.410	8.891, P =0.012	7.542, P = 0.023	2.713, P = 0.100

TABLE 4. **The presence of PEs leads to a lower number of ovarioles in the progeny of PE- × PE+ in hybrid crosses but not in most PE- × PE+ crosses.** Linear models are described in the methods section. 'mother' and 'father' refer to whether the female or male involved in the cross were either *D. sechellia* or *D. mauritiana* as indicated in column "Cross".

		Mean number of ovarioles per ovary		
Cross	**ºC**	**mother ×** ***D. simulans*** **PE+**	**mother ×** ***D. simulans*** **PE-**	**Linear model**
D. sechellia × *D. simulans*	23	10.863 (3.573)	14.250 (1.506)	1094.3, $P < 1 \times 10^{-10}$
D. sechellia × *D. simulans*	29	11.484 (3.328)	7.653 (3.138)	1639.6, $P < 1 \times 10^{-10}$
D. mauritiana × *D. simulans*	23	8.629 (3.286)	16.711 (3.134)	2944.5, $P < 1 \times 10^{-10}$
D. mauritiana × *D. simulans*	29	13.151 (3.578)	5.603 (3.205)	1639.6, $P < 1 \times 10^{-10}$
Cross	**ºC**	***D. simulans*** **PE+ × father**	***D. simulans*** **PE- × father**	**Linear model**
D. simulans × *D. sechellia*	23	15.702 (1.584)	16.807 (2.073)	12.148, $P = 4.913 \times 10^{-4}$
D. simulans × *D. sechellia*	29	17.861 (2.845)	15.512 (2.777)	50.108, $P < 1 \times 10^{-10}$
D. simulans ×	23	16.586 (2.489)	17.686 (3.269)	9.109,

D. mauritiana				P = 2.544× 10^{-3}
D. simulans × *D. mauritiana*	29	13.248 (3.043)	13.116 (3.003)	1.139, P = 0.286

TABLE 5. PE copy number in the paternal genome has a negative effect on the number of mean ovarioles per ovary in heterospecific crosses at the two assayed temperatures (ºT). Each row shows a linear model that compares the effect of PE copy number in heterospecific crosses with conspecific crosses (*D. simulans* × *D. simulans*) of the same type (i.e., PE+ × PE-, or PE- × PE+).

		Linear model		
Cross	**ºT**	**PE copies**	**cross**	**PE copies × cross**
D. sechellia × *D. simulans*	23	581.82, $P < 1 \times 10^{-10}$	905.15, $P < 1 \times 10^{-10}$	135.67, $P < 1 \times 10^{-10}$
D. sechellia × *D. simulans*	29	779.30, $P < 1 \times 10^{-10}$	1081.28, $P < 1 \times 10^{-10}$	140.05, $P < 1 \times 10^{-10}$
D. mauritiana × *D. simulans*	23	2018.43, $P < 1 \times 10^{-10}$	399.97, $P < 1 \times 10^{-10}$	800.15, $P < 1 \times 10^{-10}$
D. mauritiana × *D. simulans*	29	2018.43, $P < 1 \times 10^{-10}$	399.97, $P < 1 \times 10^{-10}$	848.34, $P < 1 \times 10^{-10}$
D. simulans × *D. sechellia*	23	25.760, $P = 3.867 \times 10^{-7}$	639.975, $P < 1 \times 10^{-10}$	17.965, $P = 2.249 \times 10^{-5}$
D. simulans × *D. sechellia*	29	15.941, $P = 6.533 \times 10^{-5}$	121.698, $P < 1 \times 10^{-10}$	17.965, $P = 2.249 \times 10^{-5}$
D. simulans × *D. mauritiana*	23	0.202, P =0.653	221.761, $P < 1 \times 10^{-10}$	0.048, P = 0.827

D. simulans × *D. mauritiana*	29	0.529, P = 0.467	336.511, $P < 1 \times 10^{-10}$	0.396, P = 0.529

FIGURE 1. The number of ovaries in F1 hybrid females in the *simulans* complex is affected by the status of PE-infection of the *D. simulans* parent but only in PE- × PE+ crosses. Black lines show conspecific crosses within *D. simulans*, blue lines show crosses between *D. sechellia* and *D. simulans* and red lines show crosses between *D. simulans* and *D. mauritiana*. **A.** Number of ovaries of F1 PE- × PE+ females raised at 23ºC. **B.** Number of ovaries of F1 PE- × PE+ females raised at 29ºC. **C.** Number of ovaries of F1 PE+ × PE- females raised at 23ºC. **D.** Number of ovaries of F1 PE+ × PE- females raised at 29ºC.

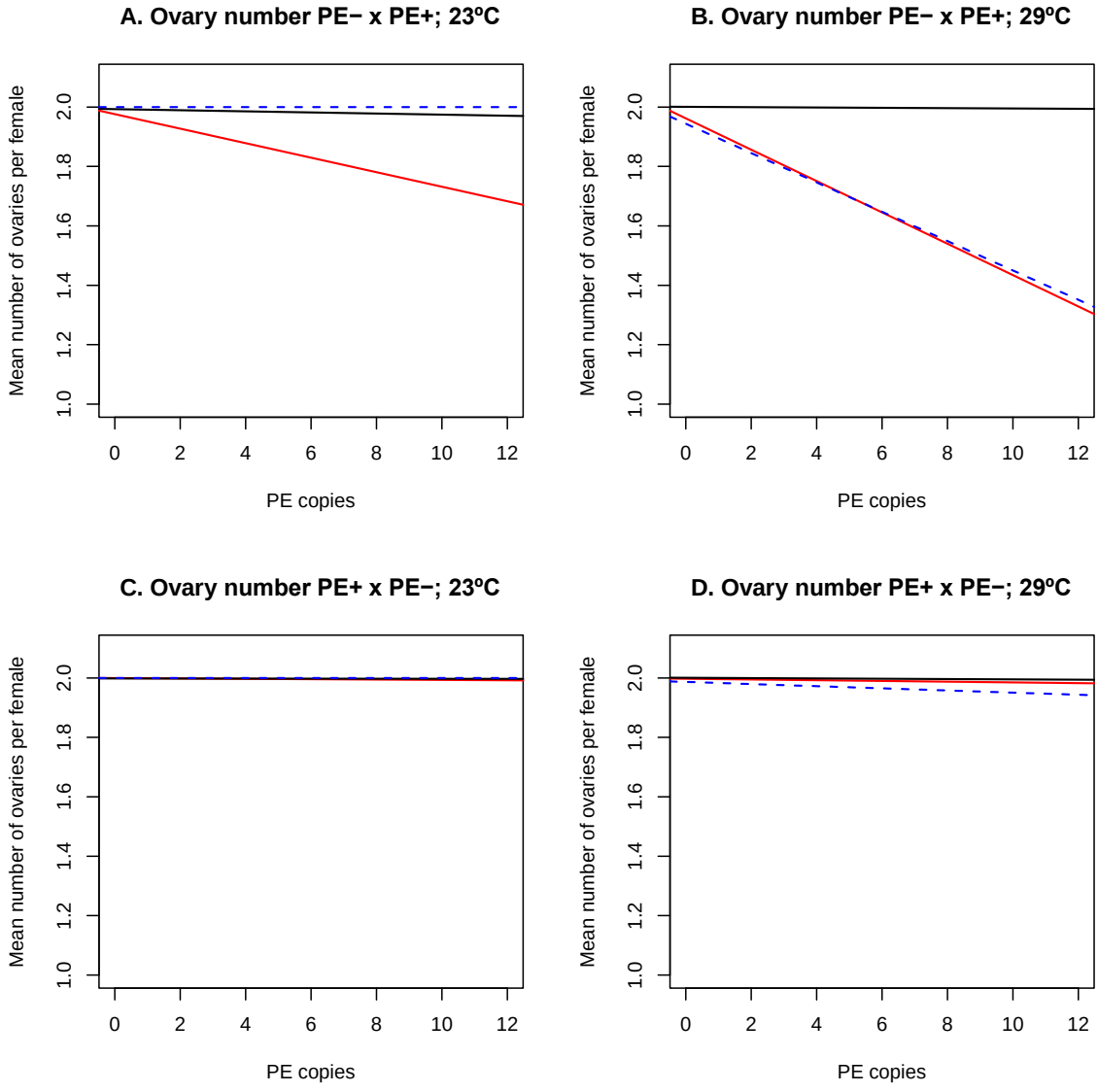


FIGURE 2. The mean number of ovarioles in F1 hybrid females in the crosses between *D. mauritiana* and *D. simulans* is affected by the status of PE-infection of the *D. simulans* parent but only when *D. simulans* is the father. Regressions show the mean number of ovarioles per female in conspecific and heterospecific crosses at 23ºC and 29ºC. Black points and lines correspond to conspecific crosses, red points and lines correspond to heterospecific crosses. **A.** *D. mauritiana* × *D. simulans* at 23ºC. **B.** *D. mauritiana* × *D. simulans* at 29ºC. **C.** *D. simulans* × *D. mauritiana* at 23ºC. **D.** *D. simulans* × *D. mauritiana* at 29ºC.

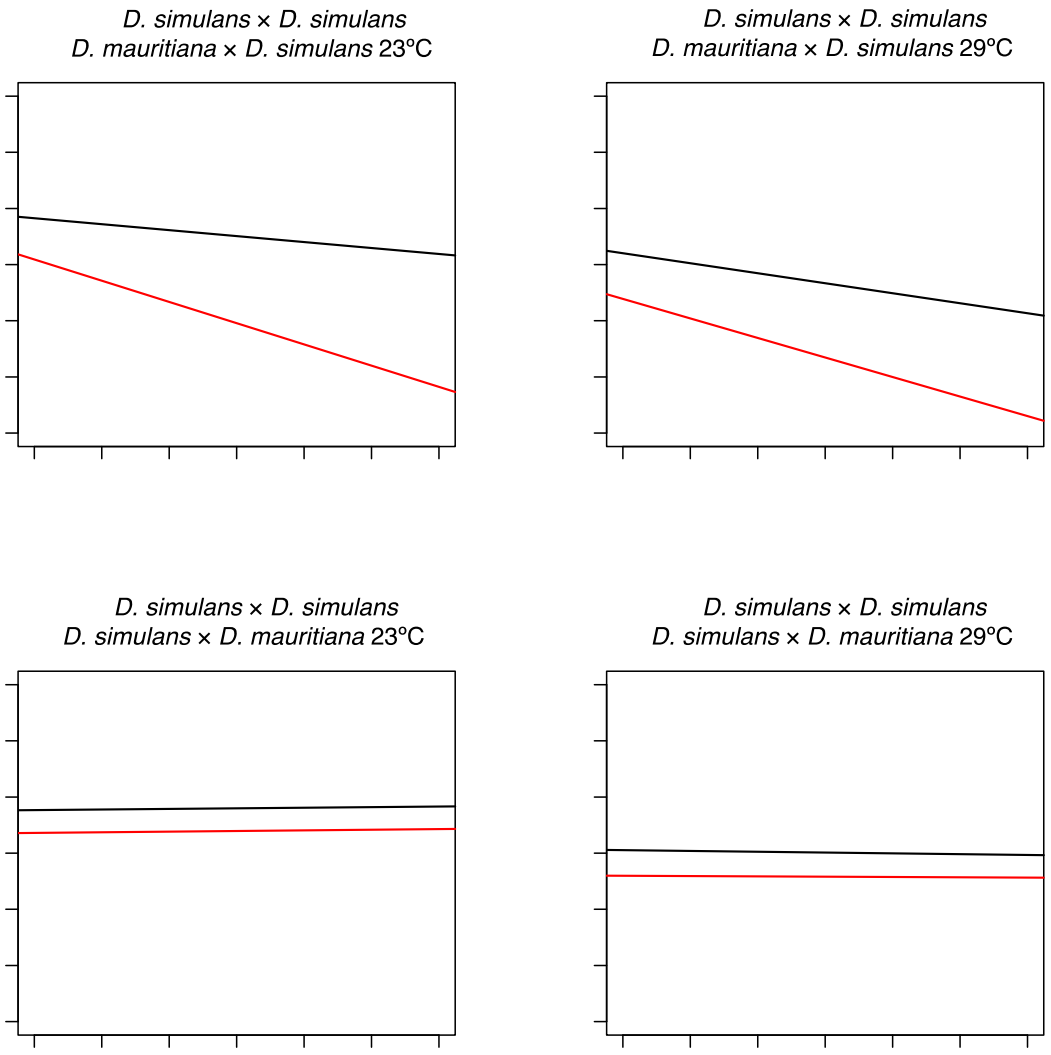
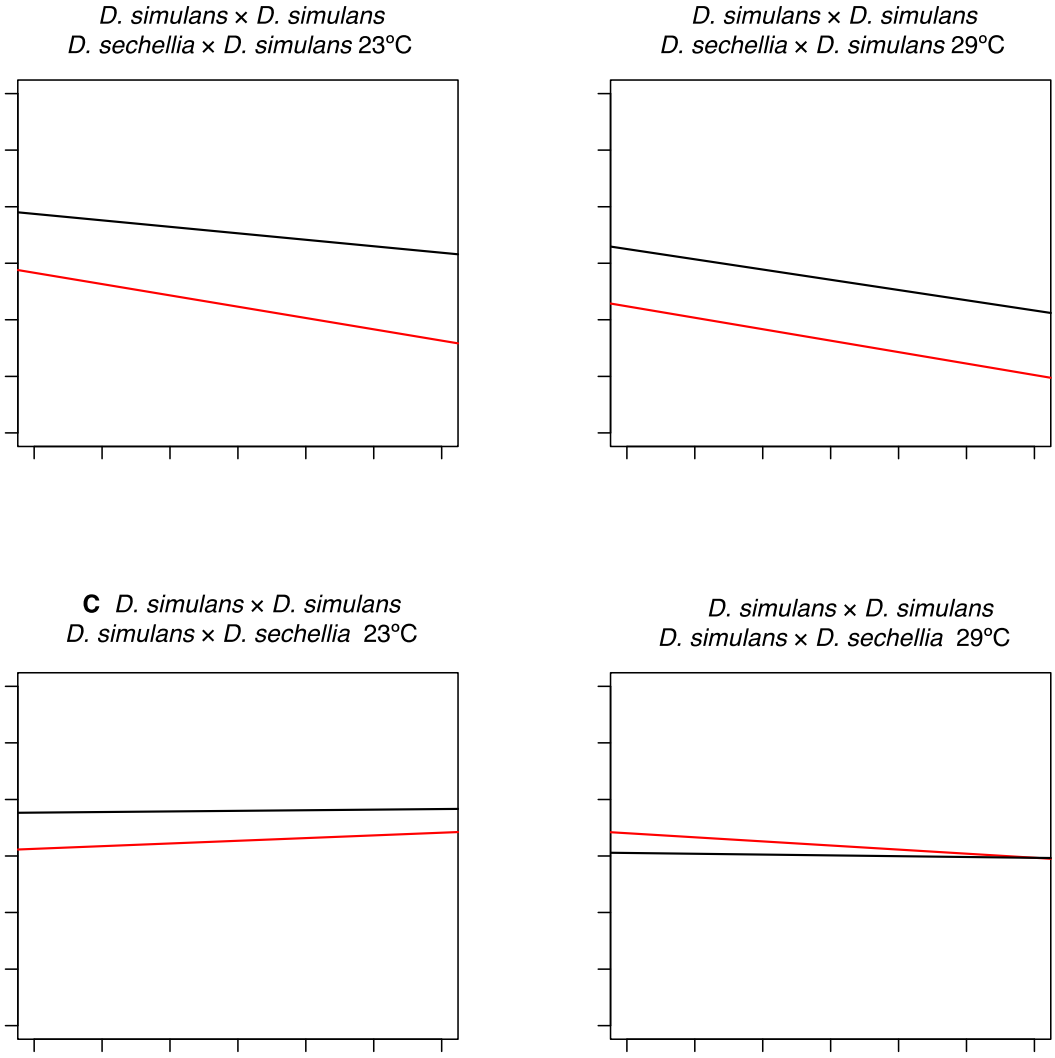


FIGURE 3. The mean number of ovarioles in F1 hybrid females in the crosses between *D. sechellia* and *D. simulans* is affected by the status of PE-infection of the *D. simulans* parent but only when *D. simulans* is the father. Regressions show the mean number of ovarioles per female in conspecific and heterospecific crosses at 23ºC and 29ºC. Black points and lines correspond to conspecific crosses, red points and lines correspond to heterospecific crosses. **A.** *D. sechellia* × *D. simulans* at 23ºC. **B.** *D. sechellia* × *D. simulans* at 29ºC. **C.** *D. simulans* × *D. sechellia* at 23ºC. **D.** *D. simulans* × *D. sechellia* at 29ºC.



REFERENCES

1. Coyne JA, Orr HA. Speciation. Sunderland, MA. 2004.
2. Nosil P. *Ecological Speciation*.; 2012. doi:10.1093/acprof:osobl/9780199587100.001.0001
3. Sobel JM, Chen GF, Watt LR, Schemske DW. The biology of speciation. *Evolution (N Y)*. 2010;64(2):295-315. doi:10.1111/j.1558-5646.2009.00877.x
4. Theodosius Dobzhansky. *Genetics and the Origin of Species*. Columbia University Press; 1937.
5. Orr HA. The genetic basis of reproductive isolation: Insights from Drosophila. *Proc Natl Acad Sci*. 2005;102(Supplement 1):6522-6526. doi:10.1073/pnas.0501893102
6. Orr HA, Presgraves DC. Speciation by postzygotic isolation: Forces, genes and molecules. *BioEssays*. 2000;22(12):1085-1094. doi:10.1002/1521-1878(200012)22:12<1085::AID-BIES6>3.0.CO;2-G
7. Turissini DA, Comeault AA, Liu G, Lee YCG, Matute DR. The ability of Drosophila hybrids to locate food declines with parental divergence. *Evolution*. 2017;71(4):960-973.
8. McQuillan MA, Roth TC, Huynh A V., Rice AM. Hybrid chickadees are deficient in learning and memory. *Evolution (N Y)*. 2018. doi:10.1111/evo.13470
9. Coyne JA, Orr HA. Patterns of Speciation in *Drosophila*. *Evolution (N Y)*. 1989;43(2):362-381. doi:10.2307/2410984
10. Coyne JA, Orr HA. "Patterns of Speciation in Drosophila" Revisited. *Evolution (N Y)*. 1997;51(1):295. doi:10.2307/2410984
11. Turissini DA, McGirr JA, Patel SS, David JR, Matute DR. The Rate Of Evolution Of Postmating-Prezygotic Reproductive Isolation In *Drosophila*. *bioRxiv*. May 2017:142059.
12. Rabosky DL, McCune AR. Reinventing species selection with molecular phylogenies. *Trends Ecol Evol*. 2010;25(2):68-74. doi:10.1016/j.tree.2009.07.002
13. Gilman RT, Behm JE. Hybridization, species collapse, and species reemergence after disturbance to premating mechanisms of reproductive isolation. *Evolution (N Y)*. 2011;65(9):2592-2605. doi:10.1111/j.1558-5646.2011.01320.x
14. Nosil P, Schluter D. The genes underlying the process of speciation. *Trends Ecol Evol*. 2011;26(4):160-167. doi:10.1016/j.tree.2011.01.001
15. Maheshwari S, Barbash DA. The Genetics of Hybrid Incompatibilities. *Annu Rev Genet*. 2011;45(1):331-355. doi:10.1146/annurev-genet-110410-132514
16. Barton NH. *Natural Selection and Random Genetic Drift as Causes of Evolution on Islands*.; 1996. https://royalsocietypublishing.org/doi/pdf/10.1098/rstb.1996.0073?casa_token=w0duBmZQF9EAAAAA:bXEz-X1-dyK8F2S6-RV6n2cjyROoIaUp-qsoT3hNYZH9DIj3S-6S1MJcPp5R-dGTr_RRwTICvb9YGA. Accessed April 20, 2019.
17. Phadnis N, Baker EP, Cooper JC, et al. An essential cell cycle regulation gene causes hybrid inviability in Drosophila. *Science*. 2015;350(6267):1552-1555.

doi:10.1126/science.aac7504

18. Meiklejohn CD, Landeen EL, Gordon KE, et al. Gene flow mediates the role of sex chromosome meiotic drive during complex speciation. *bioRxiv*. 2018. doi:10.1101/024711

19. Swanson WJ, Clark AG, Waldrip-Dail HM, Wolfner MF, Aquadro CF. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in Drosophila. *Proc Natl Acad Sci*. 2001;98(13):7375-7379. doi:10.1073/pnas.131568198

20. Presgraves DC, Balagopalan L, Abmayr SM, Orr HA. Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. *Nature*. 2003;423(6941):715-719. doi:10.1038/nature01679

21. Bono JM, Matzkin LM, Hoang K, Brandsmeier L. Molecular evolution of candidate genes involved in post-mating-prezygotic reproductive isolation. *J Evol Biol*. 2015;28(2):403-414. doi:10.1111/jeb.12574

22. Werth CR, Windham MD. A Model for Divergent, Allopatric Speciation of Polyploid Pteridophytes Resulting from Silencing of Duplicate-Gene Expression. *Am Nat*. 2002;137(4):515-526. doi:10.1086/285180

23. Lynch M, Force A. <451.Full.Pdf>. 2000. doi:10.1093/eurheartj/ehn602\r10.1093/eurheartj/ehp030

24. Moyle LC, Nakazato T. Hybrid Incompatibility {\textquotedblleft}Snowballs{\textquotedblright} Between Solanum Species. *Science (80-)*. 2010;329(5998):1521-1523.

25. Zuellig MP, Sweigart AL. Gene duplicates cause hybrid lethality between sympatric species of Mimulus. Malik HS, ed. *PLOS Genet*. 2018;14(4):e1007130. doi:10.1371/journal.pgen.1007130

26. Oka H-I. *ANALYSIS OF GENES CONTROLLING F, STERILITY IN RICE BY THE USE OF ISOGENIC LINES*. Vol 5.; 1974. https://www.genetics.org/content/genetics/77/3/521.full.pdf. Accessed July 1, 2019.

27. Bikard D, Patel D, Le Metté C, et al. Divergent Evolution of Duplicate Genes Leads to Genetic Incompatibilities Within A. thaliana. *Geochem Geophys Geosys*. 2002;296:4301. doi:10.1126/science.1167112

28. Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature*. 2006;440(7082):341-345. doi:10.1038/nature04562

29. Nolte AW, Renaut S, Bernatchez L. Divergence in gene regulation at young life history stages of whitefish (Coregonus sp.) and the emergence of genomic isolation. *BMC Evol Biol*. 2009;9(1):59. doi:10.1186/1471-2148-9-59

30. Dion-Côté AM, Barbash DA. Beyond speciation genes: an overview of genome stability in evolution and speciation. *Curr Opin Genet Dev*. 2017. doi:10.1016/j.gde.2017.07.014

31. Dion-Côté A-M, Renaut S, Normandeau E, Bernatchez L. RNA-seq Reveals Transcriptomic Shock Involving Transposable Elements Reactivation in Hybrids of Young Lake Whitefish Species. *Mol Biol Evol*. 2014;31(5):1188-1199.

doi:10.1093/molbev/msu069

32. McClintock B. Induction of Instability at Selected Loci in Maize. *Genetics*. 1953;38(6):579-599. http://www.genetics.org/content/genetics/38/6/579.full.pdf. Accessed April 30, 2018.

33. Werren JH. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc Natl Acad Sci U S A*. 2011;108 Suppl:10863-10870. doi:10.1073/pnas.1102343108

34. Trizzino M, Park YS, Holsbach-Beltrame M, et al. Transposable elements are the primary source of novelty in primate gene regulation. *Genome Res*. 2017;27(10):1623-1633. doi:10.1101/gr.218149.116

35. McClintock B. The significance of responses of the genome to challenge. *Science (80-).* 1984;226(4676):792-801. doi:10.1126/science.15739260

36. Shan X, Liu Z, Dong Z, et al. Mobilization of the active MITE transposons mPing and Pong in rice by introgression from wild rice (Zizania latifolia Griseb.). *Mol Biol Evol*. 2005;22(4):976-990. doi:10.1093/molbev/msi082

37. Ungerer MC, Strakosh SC, Zhen Y. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Curr Biol*. 2006;16(20):872-873. doi:10.1016/j.cub.2006.09.020

38. Michalak P. Epigenetic, transposon and small RNA determinants of hybrid dysfunctions. *Heredity (Edinb)*. 2009;102(1):45-50. doi:10.1038/hdy.2008.48

39. Metcalfe CJ, Bulazel K V., Ferreri GC, et al. Genomic instability within centromeres of interspecific marsupial hybrids. *Genetics*. 2007;177(4):2507-2517. doi:10.1534/genetics.107.082313

40. Waugh O'Neill RJ, O'Neill MJ, Marshall Graves JA. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature*. 1998;393(6680):68-72. doi:10.1038/29985

41. Serrato-Capuchina A, Matute DR. The role of transposable elements in speciation. *Genes (Basel)*. 2018;9(5):254. doi:10.3390/genes9050254

42. Van't Hof AE, Campagne P, Rigden DJ, et al. The industrial melanism mutation in British peppered moths is a transposable element. *Nature*. 2016;534(7605):102-105. doi:10.1038/nature17951

43. Ding Y, Berrocal A, Morita T, Longden KD, Stern DL. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature*. 2016:1-15. doi:10.1038/nature19093

44. Lafon-Placette C, Köhler C. Epigenetic mechanisms of postzygotic reproductive isolation in plants. *Curr Opin Plant Biol*. 2014;23:39-44. doi:10.1016/j.pbi.2014.10.006

45. Kidwell MG, Kidwell JF, Sved JA. Hybrid Dysgenesis in Drosophila melanogaster: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination. *Genetics*. 1977;86(4):813-833. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1213713&tool=pmcentrez&rendertype=abstract.

46. Engels WR, Preston CR. Components of hybrid dysgenesis in a wild population of Drosophila melanogaster. *Genetics*. 1980;95(1):111-128. http://www.genetics.org/content/genetics/95/1/111.full.pdf. Accessed April 30, 2018.

47. Schaefer RE, Kidwell MG, Fausto-Sterling A. Hybrid dysgenesis in Drosophila melanogaster: Morphological and cytological studies of ovarian dysgenesis. *Genetics*. 1979;92(4):1141-1152.

48. Kidwell MG, Kidwell JF. Cytoplasm-chromosome interactions in drosophila melanogaster. *Nature*. 1975. doi:10.1038/253755a0

49. Kidwell MG. Evolution of hybrid dysgenesis determinants in Drosophila melanogaster. *Proc Natl Acad Sci U S A*. 1983;80(6):1655-1659. doi:10.1073/pnas.80.6.1655

50. Rosenblum EB, Sarver BAJ, Brown JW, et al. Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model Explains Patterns of Diversification Across Time Scales. *Evol Biol*. 2012. doi:10.1007/s11692-012-9171-x

51. Davis AW, Roote J, Morley T, Sawamura K. Rescue of hybrid sterility in crosses between D. melanogaster and D. simulans. *Nature*. 1996.

52. Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang X, Barbash DA. Two Dobzhansky-Muller Genes interact to cause hybrid lethality in Drosophila. *Science (80-)*. 2006;314(5803):1292-1295. doi:10.1126/science.1133953

53. Barbash DA, Ashburner M. A novel system of fertility rescue in Drosophila hybrids reveals a link between hybrid lethality and female sterility. *Genetics*. 2003.

54. Kofler R, Hill T, Nolte V, Betancourt AJ, Schlötterer C. The recent invasion of natural Drosophila simulans populations by the P-element. *Proc Natl Acad Sci*. 2015;2015(28):201500758. doi:10.1073/pnas.1500758112

55. Mizrokhi LJ, Mazo AM. Evidence for horizontal transmission of the mobile element jockey between distant Drosophila species. *Proc Natl Acad Sci*. 1990;87(23):9216-9220. doi:10.1073/pnas.87.23.9216

56. Brookfield JF, Badge RM. Population genetics models of transposable elements. *Genetica*. 1997;100(1-3):281-294. doi:10.1007/978-94-011-4898-6_28

57. Kofler R, Hill T, Nolte V, Betancourt AJ, Schlötterer C. The recent invasion of natural *Drosophila simulans* populations by the P-element. *Proc Natl Acad Sci*. 2015;112(21):6659-6663. doi:10.1073/pnas.1500758112

58. Lachaise D, David JR, Lemeunier F, Tsacas L, Ashburner M. THE REPRODUCTIVE RELATIONSHIPS OF DROSOPHILA SECHELLIA WITH D. MAURITIANA, D. SIMULANS, AND D. MELANOGASTER FROM THE AFROTROPICAL REGION. *Evolution (N Y)*. 2018;40(2):262-271. doi:10.1111/j.1558-5646.1986.tb00468.x

59. Turissini DA, McGirr JA, Patel SS, David JR, Matute DR. The rate of evolution of postmating-prezygotic reproductive isolation in drosophila. *Mol Biol Evol*. 2018;35(2):312-334. doi:10.1093/molbev/msx271

60. Matute DR, Ayroles JF. Hybridization occurs between Drosophila simulans and D. sechellia in the Seychelles archipelago. *J Evol Biol*. 2014;27(6):1057-1068.

doi:10.1111/jeb.12391

61. Garrigan D, Kingan SB, Geneva AJ, et al. Genome sequencing reveals complex speciation in the Drosophila simulans clade. *Genome Res*. 2012;22(8):1499-1511. doi:10.1101/gr.130922.111

62. Schrider DR, Ayroles J, Matute DR, Kern AD. Supervised machine learning reveals introgressed loci in the genomes of Drosophila simulans and D. sechellia. *PLoS Genet*. 2018. doi:10.1371/journal.pgen.1007341

63. Brand CL, Kingan SB, Wu L, Garrigan D. A selective sweep across species boundaries in Drosophila. *Mol Biol Evol*. 2013. doi:10.1093/molbev/mst123

64. Hill T, Schlotterer C, Betancourt AJ. Hybrid dysgenesis in Drosophila simulans due to a rapid global invasion of the P-element. *PLoS Genet*. 2016:1-17. doi:10.1101/031781

65. Wong LC, Schedl P. Dissection of Drosophila ovaries. *J Vis Exp*. 2006;8(1):52. doi:10.3791/52

66. Hothorn T, Hornik K, Van De Wiel MA, Zeileis A. A lego system for conditional inference. *Am Stat*. 2006;60(3):257-263. doi:10.1198/000313006X118430

67. R-Core-Team. R: A language and environment for statistical computing. *JComputGraphStat*. 2013;5(3). http://www.r-project.org/.

68. Hill T, Schlötterer C, Betancourt AJ. Hybrid Dysgenesis in Drosophila simulans Associated with a Rapid Invasion of the P-Element. *PLoS Genet*. 2016;12(3):1-17. doi:10.1371/journal.pgen.1005920

69. Hudson EJ, Price TD. Pervasive reinforcement and the role of sexual selection in biological speciation. In: *Journal of Heredity*. Vol 105. ; 2014:821-833. doi:10.1093/jhered/esu041

70. Matute DR. Reinforcement Can Overcome Gene Flow during Speciation in Drosophila. *Curr Biol*. 2010;20(24):2229-2233.

71. Lachaise D, David JR, Lemeunier F, Tsacas L, Ashburner M. THE REPRODUCTIVE RELATIONSHIPS OF *DROSOPHILA SECHELLIA* WITH *D. MAURITIANA, D. SIMULANS* , AND *D. MELANOGASTER* FROM THE AFROTROPICAL REGION. *Evolution (N Y)*. 1986;40(2):262-271. doi:10.1111/j.1558-5646.1986.tb00468.x

72. Dion-Côté AM, Renaut S, Normandeau E, Bernatchez L. RNA-seq reveals transcriptomic shock involving transposable elements reactivation in hybrids of young lake whitefish species. *Mol Biol Evol*. 2014;31(5):1188-1199. doi:10.1093/molbev/msu069

73. Brand CL, Kingan SB, Wu L, Garrigan D. A selective sweep across species boundaries in Drosophila. *Mol Biol Evol*. 2013.

74. Turissini DA, Matute DR. Fine scale mapping of genomic introgressions within the Drosophila yakuba clade. Begun DJ, ed. *PLOS Genet*. 2017;13(9):e1006971. doi:10.1371/journal.pgen.1006971