

Review Article

Evolutionary Implications of Mechanistic Models of TE-Mediated Hybrid Incompatibility

Dean M. Castillo and Leonie C. Moyle

Department of Biology, Indiana University, 1003 Third Street, Bloomington, IN 47405, USA

Correspondence should be addressed to Leonie C. Moyle, lmoyle@indiana.edu

Received 16 July 2011; Accepted 16 November 2011

Academic Editor: Chau-Ti Ting

Copyright © 2012 D. M. Castillo and L. C. Moyle. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

New models of TE repression in plants (specifically *Arabidopsis*) have suggested specific mechanisms by which TE misregulation in hybrids might result in the expression of hybrid inviability. If true, these models suggest as yet undescribed consequences for (1) mechanistic connections between hybrid problems expressed at different postzygotic stages (e.g., inviability versus sterility), (2) the predicted strength, stage, and direction of isolation between diverging lineages that differ in TE activity, and (3) the association between species attributes that influence TE dynamics (e.g., mode of reproduction, geographical structure) and the rate at which they could accumulate incompatibilities. In this paper, we explore these implications and outline future empirical directions for generating data necessary to evaluate them.

1. Introduction

In many plants and animals, hybrid inviability (lethality) or hybrid sterility act as postfertilization barriers to hybridization [1]. These incompatibilities can be explained by negative genetic interactions between two or more loci under the Dobzhansky-Muller model of hybrid incompatibility [2, 3]. The Dobzhansky-Muller model does not specify the nature of the genetic elements that can lead to hybrid incompatibilities, and to date there are few cases where the loci responsible have been directly identified [4, 5]. The idea that transposable elements (TEs) can influence hybrid sterility is well established in *Drosophila* and has been highlighted in studies of intraspecific crosses resulting in hybrid dysgenesis [6–8]. However, since TE mobilization in hybrids between animal species is less frequently observed than in plants [9–11], skepticism has arisen surrounding the role that TE movement plays in the evolution of reproductive isolation [1, 11]. Recent epigenetic studies in plants, however, are providing evidence that a common genetic pathway, involving siRNA regulation of transposable elements, might lead to both hybrid inviability and hybrid sterility [12–14]. In this paper, our goal is to examine implications of these mecha-

nistic models for the genetics and evolution of reproductive isolation due to TE misregulation.

Transposable element suppression in somatic cells is generally conserved among all organisms and requires three steps [15], see Figure 1. (1) Transposon transcripts are detected by complementary small RNAs derived from previously transcribed transposons (typically siRNAs in plants and piRNAs in animals). (2) These transcripts are posttranscriptionally cleaved by small RNA-protein complexes creating small RNAs that are amplified through an RNA-dependent RNA polymerase. (3) These newly derived small RNAs are used to target transposon transcripts as in step (1) or to target the transposon sequences in the genome to induce DNA methylation and repressive chromatin modifications. Slight modifications to the pathway occur during regulation of TEs in germ line cells.

Small RNA pathways have been proposed to play a role in regulating imprinted genes and genome-wide methylation patterns; this role may influence gene expression in hybrid individuals. In this review, however, we will focus on TE regulation during the development of male and female plant gametophytes, and the consequences this could have for the expression of postzygotic hybrid incompatibilities. First, we

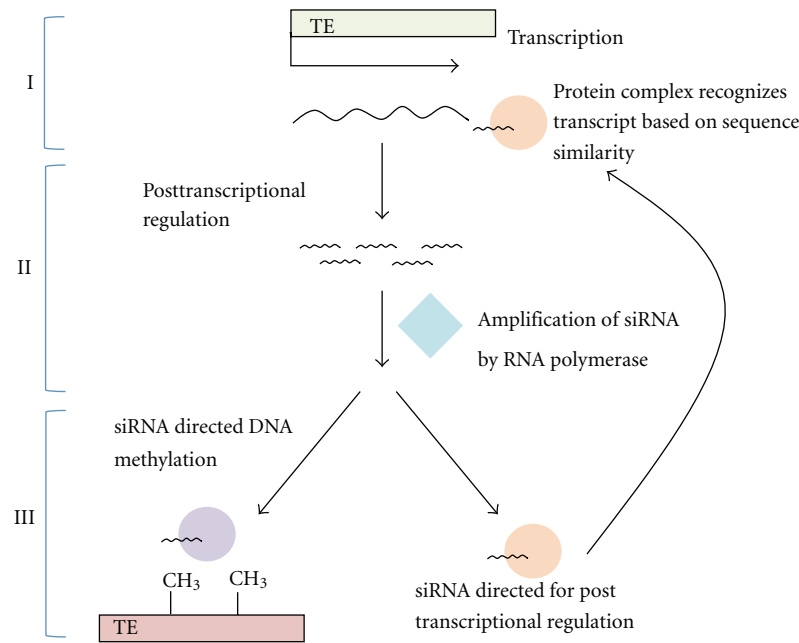


FIGURE 1: A general model of the siRNA pathway demonstrating the three processes observed in both plants and animals. (1) Transposon transcripts are detected by complementary small RNAs derived from previously transcribed transposons. (2) These transcripts are post-transcriptionally cleaved by small RNA-protein complexes creating small RNAs that are amplified through an RNA-dependent RNA polymerase. (3) These newly derived small RNAs are used to target transposon transcripts as in step (1) or target the transposon sequences in the genome to induce DNA methylation and repressive chromatin modifications.

briefly describe the current mechanistic models of TE regulation in gametogenesis, including prior connections that have been drawn between TE misregulation and the expression of hybrid incompatibilities, and indirect evidence for this association in plants (i.e., *Arabidopsis*). Second, we discuss unexplored consequences of these emerging models for the genetics and evolution of species barriers based on this mechanism. Based on this, we identify several implications of evolutionary importance: (1) TE-misregulation is a mechanism that can produce nonindependent accumulation of hybrid problems at different developmental stages of isolation between species (e.g., hybrid inviability and hybrid sterility). (2) Where this mechanism is an important contributor to reproductive isolation, the direction and strength of hybrid problems should be predictable based on lineage differences in TE abundance (“load”) and sequence identity. (3) Factors influencing the divergence of TEs among lineages should also influence when and where TE-mediated hybrid incompatibility is most likely to be observed. We identify empirical data that will be essential to assess these inferences in the future, and potential approaches for generating these data. Finally, we sketch some similarities and differences between plants and animals that could affect hybrid problems related to TEs in these groups. One of our goals is to anticipate patterns of reproductive isolation that might operate at different stages or under different ecological or evolutionary scenarios, if TE misregulation (as described by these mechanistic models) is an important contributor to reproductive isolation.

2. Mechanistic Models of TE Suppression during Male and Female Gametogenesis

Recent findings have prompted Slotkin et al. [16] to propose a companion cell model for TE regulation in male (pollen) gametogenesis that is similar to a model of piRNA accumulation in *Drosophila* egg development [17]. In angiosperm male gametes, the growing (postpollination) pollen tube contains three nuclei; two are sperm nuclei that will be involved in fertilization (see below), while one is a vegetative nucleus that does not fuse with the maternal gametophyte [18]. In the Slotkin et al. [16] model, the vegetative nucleus of pollen acts as a companion cell that “sacrifices” itself so that the sperm cells can maintain correct TE suppression. In this nucleus, TEs are demethylated, increasing transcription of elements (most that are normally somatically silenced). These transcripts are posttranscriptionally processed, creating a pool of small RNAs. These small RNAs are transported to the sperm nuclei where they direct methylation of TEs, effectively “resetting” the appropriate suppression of TEs in the germline nuclei prior to fertilization. The transcription of TEs in the pollen vegetative nucleus is initiated after the loss of heterochromatin and MET1 (cytosine-DNA-methyltransferase) and DDM1 (decreased DNA methylation) proteins [16], indicating an actively regulated process that is specific to this nucleus only. The small RNAs produced in pollen are predominantly 21-nt long and, besides translocation to the sperm cell, these small RNAs are also delivered into the zygote and endosperm [16], leading to potential

consequences for regulation, and misregulation, at postfertilization stages.

Similar to pollen development, the TE-derived small RNAs necessary for TE regulation in the female germ line (gametophyte) are created in an adjoining cell that will not contribute DNA to the next generation (in this case, epidermal cells lining the ovary) [19]. In angiosperms, the mature female gametophyte is composed of seven cells containing eight genetically identical haploid nuclei (all the products of a single initial meiotic product), surrounded by maternal (sporophytic) tissue. Two of these cells will contribute tissues to the F1 following typical “double fertilization”: the haploid ovule (egg) fuses with one sperm cell to give rise to the developing embryo; the doubled-haploid ($2 \times 1N$) “central cell” fuses with the second sperm cell to give rise to the 3N endosperm—a nutritive tissue that accumulates resources postfertilization to support development of the embryo. During development of the female gametophyte, small RNAs are translocated from the “sacrificial” epidermal cells into the ovule and central cell; these act to silence TEs in the developing embryo and endosperm after fertilization [20]. Another specific component of TE regulation, the Argonaut protein AGO9, is also expressed in the maternal sporophytic cells, but not in germ cells, and interacts preferentially with the TE-derived 24-nt small RNAs to suppress TE activity in the female gametophyte [19].

3. Connections between TE Misregulation and Hybrid Incompatibility

Based on these emerging mechanistic details, several models have recently been proposed connecting TE regulation and hybrid failure. These models have focused primarily on seed failure via misregulation of development in the endosperm [12, 13, 21], partly because some of the most evident hybrid incompatibility phenotypes in *Arabidopsis* involve endosperm failure in interspecies and interploidy crosses. Both types of crosses give rise to seed collapse (inviability) due to either overproliferation and failure to cellularize, or under-proliferation and premature cellularization, in the hybrid endosperm [22–24]. Because these phenotypes are reminiscent of the examples of overgrowth in *Peromyscus* [25, 26], earlier models hypothesized that hybrid failure results from misregulation of imprinted genes. Recently, however, Martienssen [13] proposed an alternative model where mismatches in small RNAs contributed by each parent are responsible for endosperm failure, based on the evidence that TE sequences and TE abundance differ substantially between the *A. arenosa* and *A. thaliana* genomes [21, 27]. Martienssen [13] suggests that if TE sequences in the central cell do not match siRNA in the pollen or, reciprocally, if siRNA from the female germ line do not match TE from the sperm, then TEs will be active in the endosperm causing endosperm failure. Martienssen’s model does not specify whether small RNAs have to match in terms of sequence similarity or overall quantity, as the crosses on which his model is based (*A. thaliana* by *A. arenosa*) potentially involve both differences in quantity and sequence.

The idea of an interaction between male and female “factors” controlling endosperm development has been used previously in the *Arabidopsis* literature. For example, Josefsson et al. [21] proposed a model—“dosage-dependent induction”—to explain endosperm failure in terms of the interactions between maternal and paternal factors that they observed in *Arabidopsis* hybrids, though they did not implicate TE-derived small RNAs. In their model, it is the sole responsibility of the maternal parent to deliver the proper number of “repressors” to saturate target sites in the maternal genome as well as target sites contributed by the paternal genome. If the total amount of repressor is insufficient, both maternal and paternal target sites will escape silencing. Conversely, extra doses of maternal genome can “rescue” an otherwise incompatible cross by increasing the total number of repressors. This model is generally consistent with longstanding observations in numerous species crosses, where simply manipulating maternal dosage can alter the compatibility of a cross (e.g., [28, 29]). The model is fundamentally different from the Martienssen [13] model because the male only contributes targets that require silencing but is not able to silence TEs that it contributes or that the maternal parent contributes. Michalak [12] has reinterpreted the “dosage-dependent induction” model in terms of small RNAs. In his example, he assumes that if there is a difference in the deposition of maternally loaded small RNAs between species, or if these small RNAs differ in their capabilities to suppress their targets, TEs can become active in hybrid crosses. The activation of TEs can cause endosperm failure, thereby acting as a reproductive barrier.

In both models, hybrid incompatibility is a property of TEs being preferentially activated in the endosperm, with no focus on TE activation in the embryo. One compelling reason that TE activation might preferentially occur in the endosperm is the empirical observation that the endosperm is loaded with RNAs by both parents [13], and gene regulation in the endosperm is, therefore, potentially influenced by both parental genomes. The embryo, in contrast, does not play a large role in early seed development and is less influenced by parental provisioning as most gene expression beyond the first few cell divisions is regulated within the embryo itself [30].

Both models provide new testable predictions and have the potential to provide a mechanism for earlier models of endosperm failure (e.g., the “endosperm balance number” model; [28]), but these models currently have a limited scope as general explanations of hybrid failure between species. In particular, they do not address expectations or predictions with respect to (a) the expression of hybrid failure at several other stages of reproductive isolation; (b) the differential effects of lineage differences in TE/small RNA abundance (“load”) versus TE/small RNA sequence identity; (c) the evolutionary/ecological factors that can influence TE dynamics, and therefore the expected accumulation of isolation between lineages that are differentially affected by these factors. Here, we address these different factors and some of their implications for the evolution of isolating barriers.

3.1. Broader Applications: TE-Mediated Hybrid Incompatibility at Other Isolation Stages. Although current plant models have focused on seed failure (endosperm breakdown in early F1 seed), there is reason to believe that inappropriate TE mobilization could also influence other isolation stages, including F1 male and female sterility. Indeed, the two models we have discussed are very similar to the piRNA models of transposon silencing in animals because they assume that TEs are reactivated in specific accessory cells to enhance small RNA-triggered silencing of TEs in germ line cells [20]. The phenotype observed in *Drosophila* studies of TE-mediated “hybrid dysgenesis” is male sterility. Hybrid sterility (in the form of pollen sterility) has also been observed in crosses between diploid *Arabidopsis* species [31], and Martienssen [13] discusses how the dysgenesis model could also be applied to pollen sterility in *Arabidopsis*. In addition to F1 hybrid male sterility, hybrid female sterility is also a potential consequence of TE-misregulation in a hybrid genome. For example, in the female gametophyte, proper regulation of TEs has been associated with restriction of cell fate [19]; if TEs are not correctly suppressed in the megaspore mother cell, aberrant phenotypes (including two female gametophytes in a single ovule) can occur, leading to female sterility. Finally, mutant studies indicate that misregulation of small RNAs has the potential to prevent proper double fertilization [32], although this empirical example does not involve TE-derived small RNAs.

These examples suggest that TE-mediated hybrid incompatibilities could occur at a range of stages, from early F1 inviability to F1 sterility. Given this, are there consequences for our understanding of the evolution of hybrid incompatibilities? TE-based models suggest a direct mechanistic connection, and therefore nonindependence, between different stages of postzygotic isolation in a single cross. This is inconsistent with most current models of the evolution of hybrid incompatibility, which assume independence among fixations contributing to different stages of isolation [33]. This independence assumption is reasonable for many loci contributing to species barriers, whose hybrid incompatibility effects are thought to be incidental by-products of evolutionary divergence at loci from many different potential developmental or reproductive processes [1]. However, the tight mechanistic connection between regulation of TEs during different developmental stages indicates that this independence assumption is likely violated, and therefore that predictions about hybrid incompatibilities that rely on this assumption might not hold for TE-mediated hybrid incompatibility. For example, under the Dobzhansky-Muller model of hybrid incompatibilities, the number of reproductive isolation loci is predicted [34] and observed [35, 36] to “snowball” (i.e., increase faster than linearly with time) between diverging lineages. The snowball prediction emerges from the condition that each new fixation within a diverging lineage can potentially interact with every other evolutionary change that has preceded it during lineage divergence [33, 34]). When fixations producing reproductive isolation instead involve interactions between a limited, nonindependent set of loci, these predictions do not hold (e.g., [37]). We suggest, then, that TE-mediated incompatibilities might not follow

the predictions of such models. Instead, TE-mediated incompatibilities might behave similarly to other “conflict” driven hybrid incompatibilities [4], where reproductive isolation is due to interactions between a small subset of loci evolving according to antagonistic coevolutionary processes (see further below).

3.2. Differentiating the Contribution of Divergence in TE/Small RNA Quantity versus Sequence Identity. Current models of hybrid failure are coy about the range of mechanisms by which maternal/paternal “mismatch” in TE regulation can occur. In the *Arabidopsis* interspecific crosses, it is clear that the abundance of TEs—either TE copy number or size of the resulting small RNA pool contributed by the paternal and maternal genomes—could determine the outcome of the cross, as both are correlated with endosperm failure. In this case, stoichiometric mismatch is responsible for the inappropriate regulation of either gametogenesis or postfertilization development (Figure 2(a)). Alternatively, parental “mismatch” could be due to sequence divergence between TEs in different lineages, whereby small RNAs generated from one lineage may fail to recognize or to effectively interact with target sequences from the alternative lineage, due to base-pairing mismatches (Figure 2(b)). Although there is no direct empirical evidence, it has been suggested that sequence specificity plays an important role in TE suppression [38, 39] and the proteins that interact with small RNAs rely on sequence complementarity to target TEs for methylation and silencing [40]. Given this, transcripts from one TE copy may not be able to target slightly different TE copies, although currently it is not clear how much sequence divergence must occur before TE copies can no longer suppress one another.

Regardless, determining the relative effect of TE abundance (“load”) versus sequence divergence on hybrid incompatibilities might be important because, as we argue here, these two different forms of divergence are expected to be more influential at different stages of reproductive isolation. These expectations can be inferred from the current mechanistic models of TE suppression at different developmental stages in plants. For example, TE silencing during prefertilization gametogenesis and postfertilization endosperm development both rely on sequence complementarity between small RNAs and target sites [38, 39]. However, these two developmental stages differ in that silencing in the endosperm is dependent on parentally derived small RNAs whereas silencing during gametogenesis is regulated by self-generated small RNAs. This difference in the source of siRNA regulators creates the potential for different kinds of TE divergence to differentially affect early F1 viability following heterospecific fertilization, versus gametogenesis in an F1 hybrid individual and/or hybrid problems in later generation (e.g., F2) individuals.

First, consider the F1 offspring from an interspecific (hybrid) cross. In the developing hybrid endosperm, small RNA sequences from the maternal/paternal genome will not necessarily match target sites in the other (heterospecific) genome, but they will be capable of silencing TEs derived from their own (homospecific) genome. Some of this TE

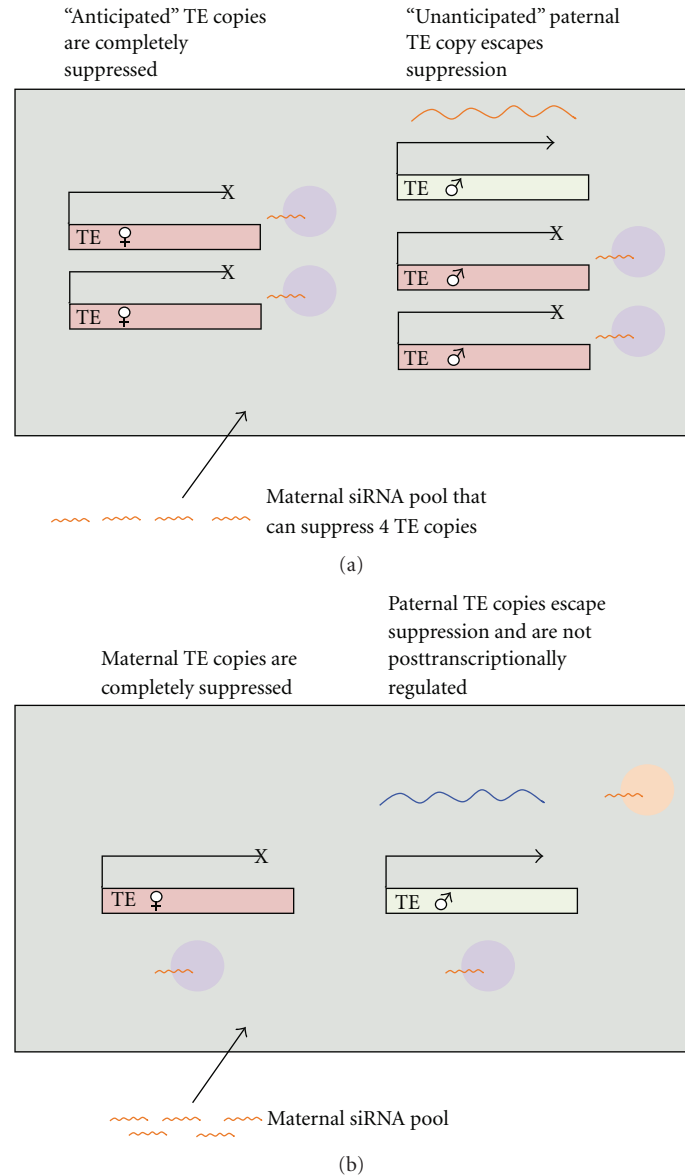


FIGURE 2: The misregulation of TEs due to a mismatch of maternal siRNA and paternal TE copies, consistent with the Martienssen and Josefsson models of endosperm failure. (a) TE load differences between parents: the maternal siRNA cannot suppress paternal TE copies due to excess of paternal TE copies. (b) TE sequence mismatch between parents: the maternal siRNA cannot suppress paternal TE copies due to differences in sequences of the TEs.

suppression takes place prior to fertilization, that is, in the sperm cells within pollen and in the female gametophyte (see above). Other regulatory control relies on postfertilization suppression, where it appears that the primary (but likely not sole; [16]) determinant of control is maternally loaded siRNAs. This "asymmetry" in control is important for the expression of hybrid problems in the early developing offspring. In particular, if the maternal parent contributes a quantity of small RNAs sufficient to regulate TEs in its own (diploid) genome, but not in excess to accommodate extra "unanticipated" copies from the other genome, TE suppression in the postfertilization offspring will be incomplete. This specifically occurs when a "low load" lineage acts as the

maternal parent in a cross with males from a "high load" lineage. Large differences in TE copies among parental genotypes could also be magnified during the production of small RNAs, if this is nonlinearly related to TE copy number.

Note that differences in sequence identity might exacerbate these load differences; when either parent is lacking copies of a specific TE, the endosperm might be overwhelmed with TE activity. Again, the expectation is that this will be observed when "naïve" lineages act as the maternal parent in crosses to males from lineages that contain the novel TE, but not necessarily in the reciprocal cross. Regardless, it appears that differences specifically in TE load are most likely to influence misregulation of TE suppression in early F1

(e.g., endosperm) development. Note that details of this prediction rely on the supremacy of maternally loaded siRNAs during early offspring development; if paternally loaded siRNAs are as important as maternally loaded siRNAs, some expectations might not hold. For example, in the Martienssen [13] model, either parental genome can be “overwhelmed” whereas in the Josefsson et al. [21] model only the maternal genome has the potential to be overwhelmed.

Second, consider the fertility of F1 offspring from an interspecific (hybrid) cross. For TEs to cause hybrid sterility, they must be improperly regulated during gametogenesis in the hybrid individual. This might occur if the small RNAs created during gametogenesis cannot fully silence TEs inherited from both parents. F1 hybrids will have all possible target sites and small RNA-producing loci, so how could incomplete suppression occur in F1s? One possible circumstance is if siRNA-mediated TE suppression is haploinsufficient (either due to siRNA production or TE targeting), for example, if there is a nonlinear relationship between siRNA suppression efficacy and TE copy number. Note that this is more likely to be due to differences in TE sequence identity among parental lineages; haploinsufficiency applies only to those loci for which the F1 is functionally haploid (i.e., loci missing in one parental lineage). In comparison, differences in load may not be important during hybrid gametogenesis because F1 individuals would have a haploid complement of TE copies from both parents; thus there are no “unanticipated” TE copies to be accounted for, unlike during endosperm TE suppression. Regardless, the expression of F1 sterility due to divergence in TE identity is expected to be “symmetric;” reciprocal F1s should show the same sterility effects.

Finally, consider the viability and fertility of recombinant later generation (e.g., F2) hybrids. In addition to the above effects, in these individuals the potentially independent segregation of TEs and their regulators is important. In animals, small RNAs used for TE suppression are primarily derived from TE clusters located in (sometimes distant) heterochromatic regions [41]. TEs in these clusters are generally no longer active and as a result their DNA sequence can change rapidly. If the hybrid progeny does not inherit these clusters (even if they have inherited other TE copies), or if sequences in the clusters no longer match the active copies outside of the cluster, then they might not generate sufficient small RNAs. If the small RNA cluster model is not appropriate for plants, variation in the efficiency of producing small RNAs may still exist between TE copies and could influence the production of small RNA pools.

Overall, on the basis of these mechanistic models, we infer that TE “load” (copy number) differences are likely more important at early (embryonic) stages of hybrid formation, when parentally (especially maternally) loaded siRNAs are critical for TE regulation. In comparison, differences in TE identity are likely more important for the expression of hybrid sterility. In addition, the symmetry of isolation between reciprocal crosses is expected to differ between these stages. Given this, depending on the factors that differentially affect these modes of TE divergence, different expression of TE-mediated hybrid incompatibility might

be expected under different ecological and evolutionary scenarios.

3.3. *Evolutionary Expectations from Predicted TE Dynamics.*

Evolutionary models have already been used to describe and predict TE dynamics within and between populations. Understanding the forces affecting TE outbreaks, proliferation, spread, and suppression might therefore help in predicting conditions under which lineages can differ in TEs, and therefore when TEs are most likely to be involved in species barriers between them.

3.3.1. *Factors Influencing Evolutionary Divergence in TE*

Abundance and Identity. For TE misregulation to influence the expression of hybrid incompatibilities, diverging lineages must differ in the activity and/or identity of their TEs. What conditions influence the accumulation of different TE copy number (load) or novel TE sequences among lineages? TE dynamics have been examined extensively using population genetic models, especially to understand conditions that produce a stable transposition-selection balance [42]. Some models have incorporated host responses in the form of alleles that suppress activity of new TE invaders [43], although we are aware of only one model that specifically incorporates the possibility of siRNA-mediated TE suppression [44]. Drawing from these models, Blumenstiel [45] has summarized the four phases of TE invasion in host populations (and the mechanisms/factors that influence these four phases): (i) invasion of a new TE (either via mutation of an existing TE, or via horizontal transmission); (ii) TE proliferation, polymorphism, and fixation in the host population (where the rate and extent of proliferation depends on transposition rates, selection against the negative effects of TEs in hosts, migration rates in the host species, and drift); (iii) the origin of a repressor locus (via a new insertion (mutation) with repressive effects); (iv) fixation of the repressor allele, and decay of the new TE family. The completion of phase IV can depend on the degree of linkage between the repressor and its targets, and whether there are other collateral effects of silencing on host function; under some conditions, the repressor allele is not expected to fix (see [45], and references therein).

Perhaps unsurprisingly, then, the factors affecting population differences in the proliferation and/or divergence of TEs are the standard evolutionary forces of mutation, migration, drift, and selection. Clearly, however, the relative influence of these factors could be affected by whether hosts can respond to new TE outbreaks by fine-tuning existing siRNA mechanisms of control [45]. For example, the (mutational) origin of a new suppressor allele becomes less important if hosts have a preexisting mechanism for suppressing the negative effects of novel TEs they encounter. Preexisting small RNA pathways might also influence the strength of selection against potential deleterious effects of TE proliferation. Indeed, models indicate that TE insertions that are the target of RNA-mediated silencing are more likely to drift to higher frequency (because their deleterious effects are moderated via this host control) than insertions that are not targets of RNA-mediated silencing [44].

Notwithstanding such observations, some factors seem more predictably associated with TE proliferation and spread regardless of, for example, the mechanism of TE control. In particular, migration can directly influence the rate at which TEs spread between populations. When there is gene flow, populations with different TE copy number can become homogenized because TEs have the potential to invade “empty” populations. In highly structured populations, however, the homogenization process is slow; as a result, heterogeneous TE copy number distributions (and, presumably, TE identities) can be maintained over long periods of time when TE proliferation is different between populations [46].

Migration or population subdivision will be influenced by life history and demographic characteristics of the host species. For example, on average, species with active dispersal mechanisms are expected (and observed; e.g., [47]) to be less subdivided than passively dispersing species. The magnitude of population subdivision can also be strongly influenced by the mode of reproduction. In plants, for example, selfing species are more strongly genetically subdivided than mixed maters or outcrossers (e.g., [47]). The mode of reproduction also influences effective population size (i.e., N_e is smaller in selfing species) and therefore the relative influence of drift on TE dynamics. Drift can act to magnify population differences in TEs by allowing the stochastic accumulation of TE load [48, 49].

These observations suggest that TE accumulation might be expected to be greater among selfing lineages. However, some models indicate that the influence of mating system on TE copy number is dependent on the specific mode of selection acting against TEs [50, 51]: where selection acts on the deleterious effects of TE insertions, both the probability that a TE is lost when its initial frequency is low and the TE copy number increase as self-fertilization increases; when selection acts on the effects of ectopic recombination between TE copies, the exact opposite effects are observed. Interestingly, several studies that have compared TE insertion frequency in selfing and outcrossing lineages produce inconsistent results as to whether selfing lineages have lower or higher TE copy number [52–55]. Also, there is no clear evidence that TE copy number and self-fertilization rate are correlated [51]. This might suggest that the most pertinent factors for TE accumulation and differentiation are N_e and migration, parameters that are only imperfectly correlated with mode of reproduction.

3.3.2. Evolutionary Expectations about the Accumulation of TE Differences. Given these complexities, based on the current models it is difficult to draw many strong generalizations about the factors that can influence the origin and spread of TEs, especially because it remains unclear how selection acts on TEs [56, 57], including those that are subject to siRNA-mediated repression [44] (and see below). Further work will be necessary to clarify this in the future. Nonetheless, it seems likely that factors that influence the frequency with which new TE outbreaks occur (such as low N_e that reduces the efficiency with which TEs are suppressed) and the degree to which individual TE outbreaks are locally “quarantined” (such as the extent of population subdivision/migration) can

influence the rates at which TE differences are able to build up between lineages. Similarly, factors that influence the rates of sequence evolution of TEs should also influence TE divergence. For example, antagonistic selection, including intergenomic conflict, has the potential to cause rapid divergence between populations [58]; thus conditions that contribute to the efficacy of antagonistic selection (such as the ability of parasites to escape from their own deleterious effects, via horizontal transmission) could also elevate the rates of lineage differentiation in TEs.

By influencing the accumulation of TE differences, such factors could influence the relative contribution of TE misregulation to the expression of species barriers among lineages. This is because these factors are not necessarily equally favorable for the fixation of other (non-TE) genetic differences that can also contribute to reproductive isolation. For example, as genomic parasites, TEs are able to “invade” empty populations (even when they have some fitness costs for their hosts), making quarantine fundamentally important in controlling TE homogenization. In comparison, complete population subdivision is not essential for adaptive genetic differentiation between populations, which can be more dependent upon the strength of local selection against immigrants. Similarly, conditions favoring antagonistic coevolution need not be the same as those favoring other adaptive fixations. Synergism among such factors might be especially favorable to TE differentiation. For example, under a geographic mosaic model, population subdivision and local (antagonistic) selection could jointly contribute to divergence in TE load and sequence identity [59, 60], although this is also true of any other potential isolating factors that are also subject to antagonistic coevolution. Regardless, given the historical attention to TE dynamics, it is surprisingly difficult to make strong predictions about factors expected to promote the accumulation of TEs between species. In future, perhaps the most clarity will not come from theoretical approaches, but rather from more taxonomically diverse and replicated empirical comparisons of TE differentiation between lineages that differ in key biological features.

3.3.3. Possible Targets of Evolutionary Change and Causes of RI. Even with limited clarity about evolutionary conditions favoring TE differences, there are some suggestive evolutionary targets that might be responsible for lineage differentiation in TE proliferation and control. For example, under the companion cell model, TEs must move from somatic cells to the germ cells to ensure they are passed to the next generation. It is at this stage of development that the interaction between TEs, suppressor alleles, and the small RNA pathway can be subject to strong selection [44] and therefore potentially rapid lineage divergence. Active TEs can only escape suppression if their sequence is divergent from small RNA producing loci or, if they produce small RNAs, when these RNAs do not interact efficiently with proteins in small RNA pathways. If small RNA pathways are under strong selection to suppress TEs in their genome, antagonistic selection from TEs trying to escape suppression can cause arms race dynamics. The piRNA machinery in some *Drosophila*

has been shown to be under strong positive selection, implying a potential arms race (but see [61]), although it has yet to be shown that selection is acting on TEs to escape suppression [62]. It is known in animals that Argonaute and Piwi proteins rely on sequence complementarity to direct TE methylation and suppression. If these proteins are fine-tuned to target specific TEs, then they may be less efficient at interacting with small RNAs from novel TEs and therefore at silencing these TEs.

Other potential targets of selective differentiation are host regulators of small RNA pools. For example, a general step in the small RNA directed TE-silencing pathway is amplification of small RNAs by an RNA-dependent polymerase. In *Arabidopsis*, the main polymerase involved in RNAi activities, including TE suppression, is Pol-IV [63]. As Pol-IV is not necessary for survival, selection acting on Pol-IV might be similar to forces affecting piRNA machinery in animals, including specialization for the TEs that occur in a specific genome; in this case, Pol-IV in hybrids may not be able to efficiently process novel small RNAs. Interestingly, changes in these targets might also result in more “standard” Dobzhansky-Muller type incompatibilities. RNA polymerases typically consist of 12 subunits, and specialization in different lineages might render subunits incompatible with their counterparts from different lineages when brought together in hybrids. In this scenario, incorrect enzyme assembly might disrupt protein function, similar to the description of PcG complex dynamics by Ishikawa and Kinoshita [64].

4. Future Empirical Approaches Linking TE Misregulation to Hybrid Incompatibility

To date, models inferring the involvement of TE misregulation in hybrid incompatibility are more suggestive than definitive, and unambiguous evidence connecting these phenomena has yet to be shown. Indeed, the specific molecular mechanisms by which TE derepression might cause cell death (and therefore hybrid problems) remains unknown, and there is currently no published evidence that mechanistically links inactivation of pathways that regulate TE suppression with hybrid sterility or lethality. In some cases, there is evidence that TE derepression has no significant fitness effect in hybrids. Chen et al. [27] have demonstrated that RNAi knockouts for *met1* in *Arabidopsis suecica* do exhibit increased levels of TE expression but sterility is not observed in these individuals and the increase of TE expression is not seen in resynthesized allotetraploids. Similarly, maize lineages appear to have large differences in TE content and abundance [65] but crosses among them do not show evidence of deleterious incompatibility phenotypes (e.g., [66]). Indeed, it is unclear whether mobilization *per se* should be predominantly responsible for hybrid incompatibility phenotypes; for example, an alternative mechanism is that TEs indirectly affect incompatibility through the collateral misregulation of genes that have acquired TE regulatory elements or of genes adjacent to these loci [67]. Ultimately, these are questions that must be addressed if the aim is to truly assess the evolutionary importance of TE-mediated isolation mechanisms.

Given this lack of empirical data, to demonstrate that TE misregulation is responsible for hybrid incompatibilities in any given system, several pieces of empirical evidence will be essential. First, lineages must differ in TE and small RNA load and/or identity, and in the parental contributions of these to offspring; if they do not, divergence in TEs clearly cannot be responsible for hybrid problems. Second, lineage divergence in TE load or identity must be consistently associated with the phenotypic expression of postzygotic isolation, such as, reproductive and developmental problems in hybrids. Finally, the expression of specific reproductive isolating barriers must be functionally linked to TE derepression. Fortunately, in combination with classical genetics, several emerging empirical approaches now make these pieces of evidence attainable now or in the near future. These data will also be useful in evaluating some of the expectations we have identified above.

4.1. Characterizing Lineage Differences in TEs and Parental Contributions of Small RNAs. Traditionally, quantifying genotype differences in TEs required approaches such as “transposon display,” where selective PCR produces a fingerprint of TE insertions for each genotype [68]. This approach can be used to roughly quantify relative copy number and describe differences between host lineages, for a known TE family [52, 53, 69, 70]. Increasingly, however, whole genome sequencing is being used to describe the entire suite of TEs within a given genome [71, 72]. This approach is not limited to known TE families and therefore can, in principle, detect TEs with novel sequences in addition to copy number differences between different genomes.

In addition to describing TE load and sequence differences between lineages, next generation deep sequencing of small RNA pools is also now routinely used to quantify tissue-specific small RNAs and changes in small RNA production in mutant lines [73–76]. An understanding of each parental contribution of both TE copy number (as a measure of load) and small RNAs would be necessary to examine the predictions of the Josefsson et al. [21] and Martienssen [13] models. For example, the main prediction of the Martienssen model is that endosperm failure will occur when maternal and paternal small RNAs do not match the TE sequences contributed by the other parent. Similarly, in the Josefsson et al. [21] model, differences in maternal small RNA contributions will determine endosperm development. Deep sequencing of small RNA pools in pollen, female gametophyte, and endosperm will allow researchers to identify which small RNAs are contributed by each parent. Small RNAs could be mapped to genomic or EST sequences to determine the identity of small RNA producing loci in the maternal and paternal genomes [41, 77]. As assembly algorithms improve for short sequences, small RNAs could be assembled *de novo*. Copy number could be assayed using inverse PCR techniques or qPCR based on sequences uncovered by deep sequencing.

Next generation sequencing, therefore, provides the technology necessary for describing lineage and parental differences in TEs and small RNAs, and for comparing hybrid small RNA pools to equivalent parental pools. Nonetheless, some limitations remain. As genome sequencing technology

improves, sequencing small RNAs and whole genomes will become easier but alignment and assembly is still a limiting step for repetitive elements. In addition, quantification of small RNAs can give an estimate of the total size of the small RNA pool, but determining the quantity contributed by a single TE is not yet achievable. Moreover, if the TE family is young and different insertions have identical DNA sequences, quantifying and identifying small RNAs to a specific TE locus is not possible. In plants, it is thought that all TE copies contribute to small RNAs, but the model emerging in animals is that clusters of mostly nonactive TEs contribute to the small RNA pool to target active TEs [41]. These challenges will need to be tackled by technological and bioinformatic advances in the future.

4.2. Demonstrating the Association between Isolating Barriers and TE Differentiation. To date, an association between differences in TE/small RNAs and endosperm failure has only been noted in crosses between *A. thaliana* and *A. arenosa*. However, in these specific crosses, several potentially contributing factors are difficult to disentangle: these species differ in TE sequence divergence, TE abundance, and ploidy. To more directly associate differences in TE/small RNAs with the expression of postzygotic barriers, it would be preferable to exclude factors, such as ploidy, that might also contribute to isolation via effects unrelated to TE differences. One such approach would be to compare the expression of isolating barriers among groups of closely related (homoploid) species that differed in known TE load and/or sequence identity. Contrasting the mean strength, identity, and direction of reproductive isolation between TE divergent lineages with that between TE-similar lineages would provide a statistical assessment of the potential contribution of TE divergence (and misregulation) to the expression of species barriers and identify which isolation stages are differentially affected. These associations might also allow a finer resolution analysis as, depending on the species group, some species might only differ in TE copy number or sequence identity rather than differing in both. This would permit the effects of TE load and sequence identity to be evaluated independently. Finally, this approach would identify specific taxon pairs for which crossing evidence associating TE differences and isolating barriers is strongest. Such pairs could be targets for further detailed functional analysis.

A complementary strategy to understand the association between TE divergence and isolating barriers is to examine previously identified incompatibility QTL for the presence of TEs (or TE regulatory sequences). Fluorescent *in situ* hybridization (FISH) could be used to evaluate colocalization of TEs and incompatibility QTL, but this would rely on existing libraries of known TEs from the organism (or a closely related species) and *a priori* hypotheses of the importance of specific TEs. Positional cloning of QTL regions—itself a challenging empirical goal—could also facilitate examination of QTL for TE sequences.

4.3. Evaluating Repression of TEs in Hybrids. Finally, to unambiguously link reproductive isolating barriers with TE misregulation, it will be necessary to functionally connect

specific developmental or reproductive problems in hybrids to TE derepression. To achieve this, a combined technique of methylation profiling and qPCR could be used to determine if TE transcripts (and specifically those known to differ between lineages) are active in the affected tissues. Methylation profiling can be achieved by using methyl-sensitive restriction enzymes on PCR products from DNA, but this requires a known TE sequence to develop appropriate primers. Bisulfite sequencing of DNA could provide an alternative approach, but a genomic DNA sequence is still needed to see which nucleotides have been converted after bisulfite treatment.

In this respect, profiling the endosperm appears to be the most straightforward initial step, as this tissue is relatively easily accessible via dissection from within the developing seed. For example, laser microdissection has already been used to specifically profile gene expression individually in the endosperm and embryo [78]; identical techniques could be used to isolate relevant tissues for siRNA and TE profiling. In comparison, to profile TE misregulation in the pollen, individual sperm cells must be dissected from their surrounding tissue, including the adjacent vegetative nucleus which is known to be demethylated (see above). Regardless, it is clear that established next generation and microdissection techniques can be used to address the specific question of whether divergent TEs are derepressed in developmentally abnormal hybrid tissues.

5. Conclusions

The idea that some reproductive isolating barriers (and therefore speciation processes) are the outcome of evolutionary conflicts has a long history but has been, until recently, mostly lacking in direct empirical support [4]. One such conflict is that between genomic parasites such as TEs and their hosts. Emerging mechanistic models of TE regulation via siRNA-mediated pathways provide renewed support for the hypothesis that TEs might be involved in the expression of interspecific reproductive barriers. Here, we have outlined several evolutionary corollaries that emerge from these mechanistic models, including the likely nonindependence of reproductive isolation acting at different stages, and the possible influence of particular demographic and life history factors on the relative susceptibility of lineages to hybrid problems based on TE misregulation. Some of these conditions themselves suggest further implications. For example, our inferences suggest that TE-mediated hybrid problems might be strongest at early stages of F1 development (e.g., during endosperm development) where lineage differences in TE load and identity can both contribute to hybrid problems. In comparison, F1 sterility is likely only influenced by differences in TE identity (see above) and requires additional conditions that seem restrictive, such as haploinsufficiency of siRNA-mediated TE suppression. (We know of no current evidence for or against this condition.) Other expectations about when and where TE misexpression might be most important are more challenging to predict at present. For example, differentiation of TEs is arguably more contingent on population subdivision than are some other potential hybrid incompatibilities; however, more theoretical and empirical

attention will be necessary to address such questions conclusively.

Drawing from specific molecular models, in this review we have focused on mechanisms described in plants. Similar inferences could be tailored to RNA-mediated TE silencing mechanisms in animals. Indeed, both plants and animals appear to have similar small RNA silencing pathways; for example, while PIWI components have not been found in plants, divergent RNA polymerases that are responsible for small RNA amplification could replace PIWI proteins [63]. Still, as noted above, the lack of TE transposition in animal hybrids has created speculation as to whether TEs could play a role in the evolution of animal hybrid incompatibilities. Hybrid inviability associated with TE misregulation has not been observed in *Drosophila*, although the *Drosophila* studies that describe hybrid dysgenesis used lines from the same species that differ in the presence of relatively few TEs [6–8]. The involvement of TEs in *Drosophila* hybrid problems remains, therefore, to be established. We note, however, that the detection ability of TE-mediated hybrid problems might be greater in the endosperm because embryos can be “rescued” from failed endosperm (via independent culture), but *Drosophila* larvae cannot be rescued from failed eggs. In addition, the most detailed studies of hybrid problems in *Drosophila* (in the melanogaster-simulans group) are all crosses involving species with small genomes, and relatively few TEs. (Some *Drosophila* appear to have large genomes due to TEs, but these are not the species used to study speciation.) Still, it is possible that TEs are more relevant to plants; on the whole, plant genomes have many TEs, and even closely related species can differ substantially in their TE complement. As yet, it remains an open question whether TEs will be a “specialist” speciation mechanism or can apply broadly across sexually reproducing organisms.

Finally, we argue that, in order to more convincingly connect TE activity with postzygotic isolating barriers, three pieces of empirical evidence will be essential: lineage differences in TE load and/or identity; clear associations between this differentiation and the expression of hybrid incompatibilities; functional evidence for a mechanistic link between these two phenomena. These data can be obtained through current or rapidly emerging approaches in genetics and genomics. As such, the field is now well positioned to evaluate the connection, if any, between the misregulation of TEs and the expression of hybrid problems—a long-held, intriguing, but poorly assessed mechanism of speciation.

References

- [1] J. A. Coyne and H. A. Orr, *Speciation*, Sinauer Associates, Sunderland, Mass, USA, 2004.
- [2] T. Dobzhansky, *Genetics and the Origin of Species*, Columbia University Press, West Sussex, UK, 1937.
- [3] H. J. Muller, “Isolation mechanisms, evolution and temperature,” *Biological Symposia*, vol. 6, pp. 71–125, 1942.
- [4] N. A. Johnson, “Hybrid incompatibility genes: remnants of a genomic battlefield?” *Trends in Genetics*, vol. 26, no. 7, pp. 317–325, 2010.
- [5] D. C. Presgraves, “The molecular evolutionary basis of species formation,” *Nature Reviews Genetics*, vol. 11, no. 3, pp. 175–180, 2010.
- [6] M. G. Kidwell, “Hybrid dysgenesis in *Drosophila melanogaster*: nature and inheritance of P element regulation,” *Genetics*, vol. 111, no. 2, pp. 337–350, 1985.
- [7] G. Yannopoulos, N. Stamatis, M. Monastirioti, P. Hatzopoulos, and C. Louis, “hobo is responsible for the induction of hybrid dysgenesis by strains of *Drosophila melanogaster* bearing the male recombination factor 23.5 MRF,” *Cell*, vol. 49, no. 4, pp. 487–495, 1987.
- [8] J. P. Blumenstiel and D. L. Hartl, “Evidence for maternally transmitted small interfering RNA in the repression of transposition in *Drosophila virilis*,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 44, pp. 15965–15970, 2005.
- [9] J. A. Coyne, “Meiotic segregation and male recombination in interspecific hybrids of *Drosophila*,” *Genetics*, vol. 114, no. 2, pp. 485–494, 1986.
- [10] J. Hey, “Speciation via hybrid dysgenesis: negative evidence from the *Drosophila* affinis subgroup,” *Genetica*, vol. 78, no. 2, pp. 97–104, 1988.
- [11] J. A. Coyne, “Mutation rates in hybrids between sibling species of *Drosophila*,” *Heredity*, vol. 63, pp. 155–162, 1989.
- [12] P. Michalak, “Epigenetic, transposon and small RNA determinants of hybrid dysfunctions,” *Heredity*, vol. 102, no. 1, pp. 45–50, 2009.
- [13] R. A. Martienssen, “Heterochromatin, small RNA and post-fertilization dysgenesis in allopolyploid and interloid hybrids of *Arabidopsis*,” *New Phytologist*, vol. 186, no. 1, pp. 46–53, 2010.
- [14] J. P. Calarco and R. A. Martienssen, “Genome reprogramming and small interfering RNA in the *Arabidopsis* germline,” *Current Opinion in Genetics and Development*, vol. 21, no. 2, pp. 134–139, 2011.
- [15] A. Girard and G. J. Hannon, “Conserved themes in small-RNA-mediated transposon control,” *Trends in Cell Biology*, vol. 18, no. 3, pp. 136–148, 2008.
- [16] R. K. Slotkin, M. Vaughn, F. Borges et al., “Epigenetic reprogramming and small RNA silencing of transposable elements in pollen,” *Cell*, vol. 136, no. 3, pp. 461–472, 2009.
- [17] S. Chambeyron, A. Popkova, G. Payen-Groschène et al., “piRNA-mediated nuclear accumulation of retrotransposon transcripts in the *Drosophila* female germline,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 39, pp. 14964–14969, 2008.
- [18] S. McCormick, “Male gametophyte development,” *Plant Cell*, vol. 5, no. 10, pp. 1265–1275, 1993.
- [19] V. Olmedo-Monfil, N. Durán-Figueroa, M. Arteaga-Vázquez et al., “Control of female gamete formation by a small RNA pathway in *Arabidopsis*,” *Nature*, vol. 464, no. 7288, pp. 628–632, 2010.
- [20] R. A. Mosher and C. Melnyk, “siRNAs and DNA methylation: seedy epigenetics,” *Trends in Plant Science*, vol. 15, no. 4, pp. 204–210, 2010.
- [21] C. Josefsson, B. Dilkes, and L. Comai, “Parent-dependent loss of gene silencing during interspecies hybridization,” *Current Biology*, vol. 16, no. 13, pp. 1322–1328, 2006.
- [22] S. Adams, R. Vinkenoog, M. Spielman, H. G. Dickinson, and R. J. Scott, “Parent-of-origin effects on seed development in *Arabidopsis thaliana* require DNA methylation,” *Development*, vol. 127, no. 11, pp. 2493–2502, 2000.

- [23] L. Comai, A. P. Tyagi, K. Winter et al., "Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids," *Plant Cell*, vol. 12, no. 9, pp. 1551–1568, 2000.
- [24] C. Bushell, M. Spielman, and R. J. Scott, "The basis of natural and artificial postzygotic hybridization barriers in *Arabidopsis* species," *Plant Cell*, vol. 15, no. 6, pp. 1430–1442, 2003.
- [25] W. D. Dawson, "Fertility and size inheritance in a *Peromyscus* species cross," *Evolution*, vol. 19, no. 1, pp. 44–55, 1965.
- [26] P. B. Vrana, J. A. Fossella, P. Matteson, T. del Rio, M. J. O'Neill, and S. M. Tilghman, "Genetic and epigenetic incompatibilities underlie hybrid dysgenesis in *peromyscus*," *Nature Genetics*, vol. 25, no. 1, pp. 120–124, 2000.
- [27] M. Chen, M. Ha, E. Lackey, J. Wang, and Z. J. Chen, "RNAi of met1 reduces DNA methylation and induces genome-specific changes in gene expression and centromeric small RNA accumulation in *Arabidopsis* allopolyploids," *Genetics*, vol. 178, no. 4, pp. 1845–1858, 2008.
- [28] S. A. Johnston, T. P. M. den Nijs, S. J. Peloquin, and R. E. Hanneman, "The significance of genic balance to endosperm development in interspecific crosses," *Theoretical and Applied Genetics*, vol. 57, no. 1, pp. 5–9, 1980.
- [29] A. Katsiotis, R. E. Hanneman, and R. A. Forsberg, "Endosperm balance number and the polar-nuclei activation hypotheses for endosperm development in interspecific crosses of Solanaceae and Gramineae, respectively," *Theoretical and Applied Genetics*, vol. 91, no. 6–7, pp. 848–855, 1995.
- [30] G. C. Ingram, "Family life at close quarters: communication and constraint in angiosperm seed development," *Protoplasma*, vol. 247, no. 3, pp. 195–214, 2010.
- [31] M. E. Nasrallah, K. Yogeewaran, S. Snyder, and J. B. Nasrallah, "*Arabidopsis* species hybrids in the study of species differences and evolution of amphiploidy in plants," *Plant Physiology*, vol. 124, no. 4, pp. 1605–1614, 2000.
- [32] M. Ron, M. A. Saez, L. E. Williams, J. C. Fletcher, and S. McCormick, "Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in *Arabidopsis*," *Genes and Development*, vol. 24, no. 10, pp. 1010–1021, 2010.
- [33] H. A. Orr and M. Turelli, "The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities," *Evolution*, vol. 55, no. 6, pp. 1085–1094, 2001.
- [34] H. A. Orr, "The population genetics of speciation: the evolution of hybrid incompatibilities," *Genetics*, vol. 139, no. 4, pp. 1805–1813, 1995.
- [35] L. C. Moyle and T. Nakazato, "Hybrid incompatibility 'snowballs' between *Solanum* species," *Science*, vol. 329, no. 5998, pp. 1521–1523, 2010.
- [36] D. R. Matute, I. A. Butler, D. A. Turissini, and J. A. Coyne, "A test of the snowball theory for the rate of evolution of hybrid incompatibilities," *Science*, vol. 329, no. 5998, pp. 1518–1521, 2010.
- [37] A. S. Kondrashov, "Accumulation of Dobzhansky-Muller incompatibilities within a spatially structured population," *Evolution*, vol. 57, no. 1, pp. 151–153, 2003.
- [38] R. K. Slotkin and R. Martienssen, "Transposable elements and the epigenetic regulation of the genome," *Nature Reviews Genetics*, vol. 8, no. 4, pp. 272–285, 2007.
- [39] M. A. Johnson and J. Bender, "Reprogramming the epigenome during germline and seed development," *Genome Biology*, vol. 10, no. 8, p. 232, 2009.
- [40] H. Okamoto and H. Hirochika, "Silencing of transposable elements in plants," *Trends in Plant Science*, vol. 6, no. 11, pp. 527–534, 2001.
- [41] J. Brennecke, A. A. Aravin, A. Stark et al., "Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*," *Cell*, vol. 128, no. 6, pp. 1089–1103, 2007.
- [42] B. Charlesworth, P. Sniegowski, and W. Stephan, "The evolutionary dynamics of repetitive DNA in eukaryotes," *Nature*, vol. 371, no. 6494, pp. 215–220, 1994.
- [43] B. Charlesworth and C. H. Langley, "The evolution of self-regulated transposition of transposable elements," *Genetics*, vol. 112, no. 2, pp. 359–383, 1986.
- [44] J. Lu and A. G. Clark, "Population dynamics of PIWI-interacting RNAs (piRNAs) and their targets in *Drosophila*," *Genome Research*, vol. 20, no. 2, pp. 212–227, 2010.
- [45] J. P. Blumenstiel, "Evolutionary dynamics of transposable elements in a small RNA world," *Trends in Genetics*, vol. 27, no. 1, pp. 23–31, 2011.
- [46] G. Decelie, S. Charles, and C. Biémont, "The dynamics of transposable elements in structured populations," *Genetics*, vol. 169, no. 1, pp. 467–474, 2005.
- [47] J. L. Hamrick and M. J. W. Godt, "Effects of life history traits on genetic diversity in plant species," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 351, no. 1345, pp. 1291–1298, 1996.
- [48] B. Charlesworth and D. Charlesworth, "The population dynamics of transposable elements," *Genetical Research*, vol. 42, no. 1, pp. 1–27, 1983.
- [49] C. H. Langley, J. F. Y. Brookfield, and N. Kaplan, "Transposable elements in Mendelian populations. I. A theory," *Genetics*, vol. 104, no. 3, pp. 457–471, 1983.
- [50] S. I. Wright and D. J. Schoen, "Transposon dynamics and the breeding system," *Genetica*, vol. 107, no. 1–3, pp. 139–148, 1999.
- [51] M. T. Morgan, "Transposable element number in mixed mating populations," *Genetical Research*, vol. 77, no. 3, pp. 261–275, 2001.
- [52] S. I. Wright, Q. H. Le, D. J. Schoen, and T. E. Bureau, "Population dynamics of an Ac-like transposable element in self- and cross-pollinating *Arabidopsis*," *Genetics*, vol. 158, no. 3, pp. 1279–1288, 2001.
- [53] S. M. Tam, M. Causse, C. Garchery, H. Burck, C. Mhiri, and M. A. Grandbastien, "The distribution of copia-type retrotransposons and the evolutionary history of tomato and related wild species," *Journal of Evolutionary Biology*, vol. 20, no. 3, pp. 1056–1072, 2007.
- [54] E. S. Dolgin, B. Charlesworth, and A. D. Cutter, "Population frequencies of transposable elements in selfing and outcrossing *Caenorhabditis* nematodes," *Genetics Research*, vol. 90, no. 4, pp. 317–329, 2008.
- [55] S. I. Wright, R. W. Ness, J. P. Foxe, and S. C. H. Barretty, "Genomic consequences of outcrossing and selfing in plants," *International Journal of Plant Sciences*, vol. 169, no. 1, pp. 105–118, 2008.
- [56] A. le Rouzic and G. Decelie, "Models of the population genetics of transposable elements," *Genetical Research*, vol. 85, no. 3, pp. 171–181, 2005.
- [57] Y. C. G. Lee and C. H. Langley, "Transposable elements in natural populations of *Drosophila melanogaster*," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1544, pp. 1219–1228, 2010.
- [58] W. R. Rice and B. Holland, "The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen," *Behavioral Ecology and Sociobiology*, vol. 41, no. 1, pp. 1–10, 1997.

- [59] S. L. Nuismer, J. N. Thompson, and R. Gomulkiewicz, "Gene flow and geographically structured coevolution," *Proceedings of the Royal Society B: Biological Sciences*, vol. 266, no. 1419, pp. 605–609, 1999.
- [60] J. N. Thompson, "Coevolution: the geographic mosaic of co-evolutionary arms races," *Current Biology*, vol. 15, no. 24, pp. R992–R994, 2005.
- [61] D. M. Castillo, J. C. Mell, K. S. Box, and J. P. Blumenstiel, "Molecular evolution under increasing transposable element burden in *Drosophila*: a speed limit on the evolutionary arms race," *BMC Evolutionary Biology*, vol. 11, no. 1, Article ID 258, 2011.
- [62] D. J. Obbard, K. H. J. Gordon, A. H. Buck, and F. M. Jiggins, "The evolution of RNAi as a defence against viruses and transposable elements," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1513, pp. 99–115, 2009.
- [63] R. A. Mosher, "Maternal control of Pol IV-dependent siRNAs in *Arabidopsis* endosperm," *New Phytologist*, vol. 186, no. 2, pp. 358–364, 2010.
- [64] R. Ishikawa and T. Kinoshita, "Epigenetic programming: the challenge to species hybridization," *Molecular Plant*, vol. 2, no. 4, pp. 589–599, 2009.
- [65] N. M. Springer, K. Ying, Y. Fu et al., "Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content," *PLoS Genetics*, vol. 5, no. 11, Article ID e1000734, 2009.
- [66] L. Rayburn, D. P. Biradar, D. G. Bullock, and L. M. McMurphy, "Nuclear DNA content in F1 hybrids of maize," *Heredity*, vol. 70, pp. 294–300, 1993.
- [67] C. Feschotte, "Transposable elements and the evolution of regulatory networks," *Nature Reviews Genetics*, vol. 9, no. 5, pp. 397–405, 2008.
- [68] D. van den Broeck, T. Maes, M. Sauer et al., "Transposon display identifies individual transposable elements in high copy number lines," *The Plant Journal*, vol. 13, no. 1, pp. 121–129, 1998.
- [69] N. de Setta, A. P. P. Costa, F. R. Lopes, M. A. van Sluys, and C. M. A. Carareto, "Transposon display supports transpositional activity of P elements in species of the *saltans* group of *Drosophila*," *Journal of Genetics*, vol. 86, no. 1, pp. 37–43, 2007.
- [70] S. Schaack, E. J. Pritham, A. Wolf, and M. Lynch, "DNA transposon dynamics in populations of *Daphnia pulex* with and without sex," *Proceedings of the Royal Society B: Biological Sciences*, vol. 277, no. 1692, pp. 2381–2387, 2010.
- [71] Q. H. Le, S. I. Wright, Z. Yu, and T. E. Bureau, "Transposon diversity in *Arabidopsis thaliana*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 13, pp. 7376–7381, 2000.
- [72] M. I. Tenaillon, M. B. Hufford, B. S. Gaut, and J. Ross-Ibarra, "Genome size and transposable element content as determined by high-throughput sequencing in maize and *Zea luxuri-ans*," *Genome Biology and Evolution*, vol. 3, no. 1, pp. 219–229, 2011.
- [73] C. Lu, S. S. Tej, S. Luo, C. D. Haudenschild, B. C. Meyers, and P. J. Green, "Elucidation of the small RNA component of the transcriptome," *Science*, vol. 309, no. 5740, pp. 1567–1569, 2005.
- [74] A. Girard, R. Sachidanandam, G. J. Hannon, and M. A. Carmell, "A germline-specific class of small RNAs binds mammalian Piwi proteins," *Nature*, vol. 442, no. 7099, pp. 199–202, 2006.
- [75] I. R. Henderson, X. Y. Zhang, C. Lu et al., "Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning," *Nature Genetics*, vol. 38, no. 6, pp. 721–725, 2006.
- [76] K. D. Kasschau, N. Fahlgren, E. J. Chapman et al., "Genome-wide profiling and analysis of *Arabidopsis* siRNAs," *PLoS Biology*, vol. 5, no. 3, pp. 479–493, 2007.
- [77] H. Yin and H. Lin, "An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*," *Nature*, vol. 450, no. 7167, pp. 304–308, 2007.
- [78] R. C. Day, U. Grossniklaus, and R. C. Macknight, "Be more specific! laser-assisted microdissection of plant cells," *Trends in Plant Science*, vol. 10, no. 8, pp. 397–406, 2005.

