Title
The spliceosome as a transposon sensor

Journal
RNA Biology, 10(11)

ISSN
1547-6286

Authors
Dumesic, Phillip A
Madhani, Hiten D

Publication Date
2013-11-01

DOI
10.4161/rna.26800

Peer reviewed

RNA Biology 10:11, 1653–1660; November 2013; © 2013 Landes Bioscience

The spliceosome as a transposon sensor

Phillip A Dumesic and Hiten D Madhani*
Department of Biochemistry and Biophysics; University of California; San Francisco, CA USA

The ability to distinguish self from non-self nucleic acids enables eukaryotes to suppress mobile elements and maintain genome integrity. In organisms from protist to human, this function is performed by RNA silencing pathways. There have been major advances in our understanding of the RNA silencing machinery, but the mechanisms by which these pathways distinguish self from non-self remain unclear. Recent studies in the yeast *C. neoformans* indicate that transposon-derived transcripts encode suboptimal introns and tend to stall in spliceosomes, which promotes the biogenesis of siRNA that targets these transcripts. These findings identify gene expression signal strength as a metric by which a foreign element can be distinguished from a host gene, and reveal a new function for introns and the spliceosome in genome defense. Anticipating that these principles may apply to RNA silencing in other systems, we discuss strong hints in the literature suggesting that the spliceosome may guide small RNA biogenesis in the siRNA and piRNA pathways of plants and animals.

Keywords: transposon, genome defense, small RNA, RNA interference, pre-mRNA splicing, spliceosome, RNA processing

*Correspondence to: Hiten D Madhani; Email: hitenmadhani@gmail.com

Submitted: 07/24/2013; Revised: 10/09/2013; Accepted: 10/14/2013

http://dx.doi.org/10.4161/rna.26800

RNA Silencing Pathways Defend Against Mobile Genetic Elements

Eukaryotes are generally colonized by transposons and other parasitic nucleic acids.[1] If not suppressed, transposons hijack host expression machinery to mobilize throughout the genome, which can disrupt host genes,[2] promote chromosomal rearrangements,[3] and cause disease.[4,5] The human genome carries scars of such events: over half of its sequence is derived from transposons.[6] Genome defense against transposons, however, is confounded by transposon diversity. Transposons exist in multiple families that bear little homology to each other, hindering sequence-based recognition strategies. Furthermore, they use diverse mobilization mechanisms, impeding identification based on a distinctive enzymology or intermediate.[7] Thus, genome defense systems must be adaptable enough to recognize myriad transposon types, yet specific enough to silence foreign elements without perturbing host genes.

RNAi-related RNA silencing pathways represent deeply conserved mechanisms by which eukaryotes from protist to human recognize and silence transposons.[8] In these pathways, small RNAs of ~20–30 nt are loaded into Argonaute family proteins, then bind complementary target RNAs to trigger one or more silencing mechanisms. In some contexts, mechanisms such as RNA endonucleolysis or translational repression act directly on the target transcript to silence protein expression.[9] In other contexts, repressive histone modification and DNA methylation silence the target transcript's corresponding DNA locus.[10]

Presumably, the specificity with which RNA silencing pathways target foreign genetic elements is achieved in large part by the appropriate selection of substrates from which small RNAs are produced. One major substrate is long double-stranded RNA (dsRNA), which is processed by RNaseIII-type Dicer enzymes to generate small interfering RNA (siRNA). Because Dicer acts in a sequence-independent manner,[11] its ability to target foreign elements depends on their ability to produce dsRNA. Cellular dsRNA in *A. thaliana*, *C. elegans*, and *Drosophila* is indeed enriched in transposon sequences,[12,13]

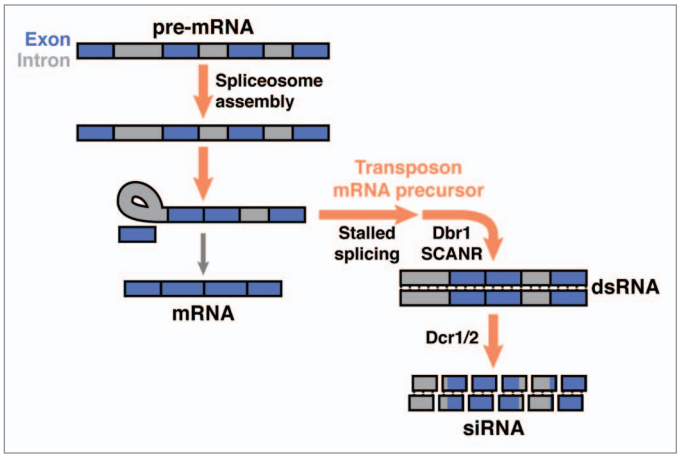

Figure 1. Kinetic competition model for siRNA biogenesis in *C. neoformans*. A kinetic competition between splicing and dsRNA synthesis contributes to the targeting of inefficiently spliced transcripts by siRNA. In this hypothetical example, splicing of a transcript's first intron stalls at the lariat intermediate stage. The intermediate is processed by the lariat debranching enzyme (Dbr1) and SCANR in order to generate dsRNA, which is converted to siRNA by Dcr1/2. The tendency of transposon-derived transcripts to encode suboptimal splicing features and accumulate in spliceosomes targets them for RNA silencing.

perhaps due to the inverted repeats and internal antisense promoters encoded by some transposon families, which promote the formation of intra- and intermolecular dsRNA, respectively.[14-16] In addition, the tendency of transposons to mobilize can place them adjacent to a host-encoded promoter, resulting in antisense transposon transcripts that may promote intermolecular dsRNA formation.[15,17] Thus, Dicer supports RNA silencing of many transposon types,[14,16-19] even potentially those of novel sequence, based simply on their tendency to form dsRNA.

However, some triggers of RNA silencing do not naturally produce dsRNA, raising the possibility that single-stranded transcripts can trigger small RNA production. How are such triggers identified, given their structural similarity to host transcripts? In some cases, RNA-dependent RNA polymerases (RdRPs) are required for the silencing of transposons and transgenes,[20-22] suggesting a model in which the targeted activity of RdRP produces substrates for Dicer.

PIWI-interacting RNA (piRNA) pathways, including those in mouse and *Drosophila*, represent another example in which single-stranded transcripts initiate RNA silencing. In these systems, piRNA precursor transcripts are processed in a Dicer-independent manner to generate small RNAs that function with Argonaute proteins of the PIWI clade to silence complementary sequences.[23] The specificity with which piRNAs target foreign elements is thought to be achieved by restricting piRNA biogenesis to precursors transcribed from particular genomic loci. These loci, termed piRNA clusters, are composed of transposons and transposon remnants, and thus appear to be a memory of foreign elements encountered by an organism.[24,25] The signals that identify piRNA precursor transcripts are unknown.

As described above, the specificity of RNA silencing depends on the selection of appropriate substrates for small RNA biogenesis. Elucidation of the signals that guide small RNA biogenesis is thus an important goal in the study of RNA silencing, and should illustrate the principles by which eukaryotes distinguish self from non-self nucleic acids. Here, we highlight the recent finding that transposon transcripts can be targeted as substrates for small RNA production owing to their tendency to stall in spliceosomes. Furthermore, we discuss hints in the literature that point to a general relevance of spliceosome-coupled small RNA biogenesis to other systems.

Stalled Spliceosomes are a Signal for RNAi-Mediated Genome Defense

The human pathogenic yeast *Cryptococcus neoformans* is a genetically tractable model system for the study of endogenous siRNA pathways. RNA silencing in this organism is performed by one Argonaute protein (Ago1), two Dicers (Dcr1/2), and an RdRP (Rdp1).[21] Null mutations in their corresponding genes cause no defect in vegetative growth, but elicit increased transposon expression and mobilization, pointing to genome defense as a key function for this pathway.[21,26] As *S. cerevisiae* lacks RNA silencing and *S. pombe* synthesizes small RNAs against repeats but not against transposons, *C. neoformans* offers a well-developed yeast model to study RNA silencing of transposable elements.

Deep sequencing of *C. neoformans* siRNAs revealed that they map predominantly to repetitive sequences derived from transposons, consistent with their known role in transposon suppression.[21,27] Surprisingly, however, siRNAs map not only to the exons but also to the introns of their corresponding transcripts, implicating incompletely spliced mRNA precursors as substrates for siRNA biogenesis. These initial observations suggested that transposon-derived transcripts may be recognized as triggers of RNA silencing while still undergoing splicing in the nucleus.

The next clue came from biochemical studies revealing that several *C. neoformans* RNA silencing factors, including Ago1 and Rdp1, are members of a spliceosome-coupled and nuclear RNAi complex (SCANR), which physically interacts with the spliceosome and is required for siRNA biogenesis.[27] These observations raised the possibility that a kinetic competition determines transcripts from which small RNAs are produced. In this model, rapidly spliced transcripts are not targeted for dsRNA synthesis, whereas transposon transcripts, by virtue of their poor splicing kinetics, accumulate in spliceosomes and are targeted for dsRNA synthesis by SCANR (**Fig. 1**).

Consistent with this model, transcripts targeted by siRNA in *C. neoformans* exhibit intron sequence features predictive of poor splicing. Furthermore, we observed that these transcripts accumulate abnormally on spliceosomes in vivo, indicating that they stall in the spliceosome during the splicing cycle. Importantly, we demonstrated a number of predictions of the kinetic competition model

by experimentally manipulating introns within transcripts that template siRNA biogenesis. First, intron elimination reduced a transcript's ability to template siRNA production. Second, introduction of a 3′ splice site mutation, which stalls splicing upon lariat intermediate formation, dramatically increased the accumulation of siRNA targeting the transcript. Third, a 5′ splice site mutation, which prevents intron engagement with the spliceosome, suppressed the ability of a 3′ splice site mutation in the same intron to promote siRNA biogenesis. Fourth, loss of the lariat debranching enzyme (Dbr1) also blocked siRNA production, further implicating stalled splicing intermediates (which require Dbr1 for their processing and degradation)[28,29] as substrates for siRNA production. Together, these findings establish stalled spliceosomes as a necessary signal for the specification of RNA silencing targets. As such, they reveal a new function for introns and the spliceosome in genome defense (**Fig. 1**).[27]

The findings described above support a kinetic competition model for siRNA template selection in *C. neoformans* and demonstrate that gene expression signal strength can influence RNA silencing specificity. In this example, the suboptimal splicing properties of transposons, relative to host genes, provide an opportunity for the evolution of genome defense. Given the universality of spliceosomal introns in eukaryotes, the capacity for splicing features to guide RNA silencing is not necessarily unique to *Cryptococcus*. Below we consider potential roles for the spliceosome in small RNA pathways that mediate genome defense in other organisms. We also suggest potential advantages of spliceosome-coupled genome defense, and speculate on evolutionary forces that may encourage suboptimal splicing of transposon transcripts.

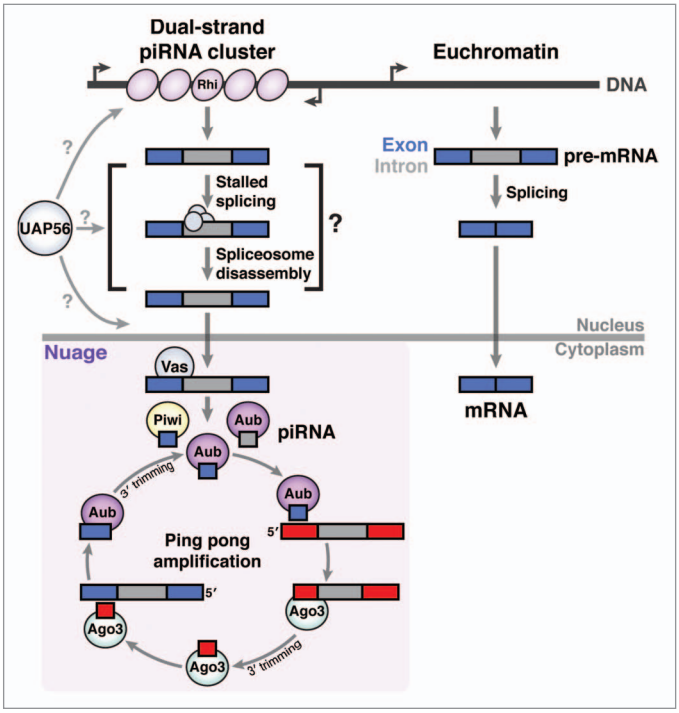

Figure 2. Speculative model for the utilization of incompletely spliced transcripts in *Drosophila* piRNA biosynthesis. The stalled splicing of transcripts originating from piRNA cluster loci may target them for piRNA biogenesis. Transgenes inserted into dual-strand piRNA clusters are spliced less efficiently than when expressed from euchromatic loci, and they give rise to piRNAs that correspond to both intronic and exonic regions.[58] We speculate that this effect may be caused by sequence features of piRNA clusters, or by their heterochromatin context, as indicated by the presence of Rhino (Rhi), an HP1 variant. The splicing and nuclear export factor UAP56 is required for piRNA production from dual-strand piRNA clusters.[59] UAP56 binds piRNA precursor transcripts and colocalizes with Rhi foci in the nucleus, suggesting that UAP56 targets piRNA precursors to Vasa (Vas), a protein that coordinates piRNA processing in the perinuclear nuage.[59] There are several potential mechanisms by which UAP56, a DEAD box protein, might act in the context of this model. First, it could promote heterochromatinization of piRNA cluster loci by Rhi, as suggested by the fact that nuclear Rhi foci require UAP56.[59] Second, UAP56 could bind to piRNA precursor transcripts in order to cause stalled splicing or to disassemble stalled spliceosomes for downstream processing of the precursors. Finally, UAP56 could mediate the nuclear export of precursor transcripts to Vas. Vas subsequently promotes primary piRNA processing as well as the ping pong amplification cycle, in which primary piRNA acts with the PIWI protein Aub to cleave complementary transcripts (indicated in red) that originate either from the dual-strand piRNA cluster itself or from transposons located elsewhere in the genome. Cleavage defines the 5′ end of a secondary piRNA, whose 3′ end is subsequently trimmed to proper length. Secondary piRNA acts with the distinct PIWI protein Ago3 to cleave piRNA cluster-derived precursor transcripts, thereby amplifying their conversion to mature piRNA.

Connections Between RNA Splicing and Plant siRNA Biogenesis

From the discovery of transposons[30] to the suggestion that small RNAs mediate RNA silencing,[31] plant systems have provided insights into the biology of transposons and the RNA silencing pathways that suppress them. In plants, two siRNA pathways contribute to genome defense, each of which requires distinct RdRP, Dicer, and Argonaute proteins. In the RNA-directed DNA methylation (RdDM) pathway, the specialized RNA polymerases Pol IV and Pol V act to trigger siRNA production and DNA methylation, respectively, at the loci they transcribe.[10,32] In the post-transcriptional gene silencing (PTGS) pathway, siRNA production results in target transcript destabilization. The ability of these pathways to broadly silence transposons and transgenes, including some that do not naturally generate dsRNA,[22,33] suggests that they recognize distinguishing features of foreign elements. That splicing signals could represent one of these features is suggested by two observations. First, mutations in splicing factors affect the production of siRNA for the RdDM pathway. Second, efficient splicing reduces the accumulation of transgene-specific siRNA for the PTGS pathway. The details of these observations are described below.

Functional studies have implicated the spliceosome in RdDM pathway targeting. Specifically, mutations of the *A. thaliana* splicing factors ZOP1, MAC3A, MOS4, MOS12, MOS14, and SR45 reduce accumulation of Pol IV/V-dependent siRNA.[34,35] MAC3A (yeast Prp19) and MOS4 associate with the evolutionarily conserved, spliceosomal Nineteen complex.[36] We note that Prp19 accumulates on transcripts in *C. neoformans* whose stalled splicing targets them for RNA silencing.[27] SR45 contains arginine/serine-rich domains, which are found in splicing factors of the SR protein family, whereas MOS14 is a nuclear import receptor for SR proteins.[37] A potential direct role for SR proteins in RdDM is intriguing given the observation that Srr1, a spliceosome-associated protein that contains an arginine/serine-rich domain, physically interacts with siRNA biogenesis factors of the SCANR complex in *C. neoformans*.[27] Important questions for future work include determining whether these proteins mediate splicing events at RdDM loci, whether they promote transcription by Pol IV/V, and whether they physically interact with known components of the RdDM pathway. The possibility of a role in transcription is made more likely by recent work in mammalian cells showing that SR proteins can license transcription by promoting RNA polymerase pause release.[38] Nevertheless, further mechanistic studies will be required to rule out the possibility that splicing factor mutants affect RNA silencing indirectly, via their global effects on splicing.

Other experiments have assessed connections between splicing and RNA silencing by manipulating intron features in transcripts that template siRNA production. In contrast to *C. neoformans*, where experiments of this type demonstrated that poorly spliced introns promote siRNA production, plant experiments described to date have highlighted a potentially distinct mechanism by which introns can influence RNA silencing, in which efficiently spliced introns protect transcripts from RNA silencing. For instance, a comparison of intron-containing and intronless GFP transgenes demonstrated that introns impede PTGS-mediated transgene silencing in *A. thaliana*.[39] Furthermore, better-spliced introns conferred greater protection against RNA silencing. It was suggested that this phenomenon explains why endogenous intronless genes template more siRNA production than do intron-containing genes, as assessed by genome-wide siRNA sequencing.[39] These findings echo the earlier observation that introns retard the RdRP-dependent phenomenon of secondary siRNA, in which a single-stranded target of siRNA becomes gradually itself a template for siRNA synthesis.[40] Although the mechanisms by which efficient splicing opposes RNA silencing are unclear, the requirement for RdRP in plant siRNA pathways raises the possibility that efficient splicing acts to prevent transcripts from exhibiting triggers for RdRP activity, which are thought to include absence of a 5′ cap or poly-A tail.[41,42]

To summarize, the data to date suggest that efficient splicing impedes RNA silencing in plants. In *C. neoformans*, stalled splicing promotes RNA silencing. Both systems may thus utilize a similar logic for transposon recognition, in which efficiently spliced transcripts avoid RNA silencing. Future studies will be required to determine whether efficiently spliced introns can actively protect against RNA silencing in *C. neoformans*, and whether stalled spliceosomes promote siRNA production in plants. In this regard, it is interesting to note that the insertion of ectopic 5′ splice sites into a plant transgene, which might mimic the effects of a spliceosome stalled at the first step of splicing, has been shown to cause improper transcript maturation and lead to siRNA production.[43]

Splicing Factors are Required for siRNA Production in *C. elegans*

Early studies of nematode RNAi triggered by exogenous dsRNA demonstrated that many of the factors required for this process are also required to silence endogenous transposons via siRNA production.[14,44,45] Since then, forward genetic screens have helped disentangle the myriad RNA silencing factors in *C. elegans*, whose genome encodes one Dicer, 27 Argonaute, and four RdRP proteins. Remarkably, dozens of splicing factors were found in these screens to be required for transgene or transposon suppression.[46-48] These splicing factors are not restricted to any particular spliceosome subunit, nor do they act at a common step in the splicing cycle. Thus, confirmation that these factors affect RNA silencing in a direct manner, and not simply by globally perturbing splicing, awaits determination of whether they associate with transposon transcripts, whether their mutant phenotypes can be recapitulated by intronic mutations in transposon transcripts, and whether they specifically affect particular siRNA types. One potential clue is the recent observation that a specific subset of 26G siRNA maps to both introns and exons of its target transcripts, suggesting that siRNA of this class can be triggered by incompletely spliced mRNA precursors.[49] A further motivation is the recent finding that nematode spliceosome components are enriched among proteins whose phylogenetic conservation correlates with that of RNA silencing factors.[48] Since proteins with similar conservation tend to act in the same pathways,[50] this finding suggests a conserved connection between splicing and small RNA function.

A Potential Function for Stalled Spliceosomes in piRNA Biogenesis

In animals such as *Drosophila* and mice, piRNAs act to suppress transposons in the germline, where their function is required for fertility.[51] The biogenesis of piRNAs appears to be restricted to the particular sequences in defined genomic loci. These loci typically comprise clusters of transposon remnants, consistent with their function as historical databases of transposon encounters.[25] Some piRNAs are also produced from protein-coding genes, but their functions are less clear.

piRNA biogenesis begins with primary processing, a poorly characterized conversion of single-stranded precursor transcripts into mature piRNAs, which act in Argonaute proteins of the PIWI clade to carry out transcriptional and post-transcriptional gene silencing.[23,52,53] For some piRNA loci, the piRNA pool is amplified by a ping-pong cleavage cycle, which is best understood in *Drosophila*. In this cycle, a primary piRNA directs cleavage of a complementary transposon transcript,

thereby defining the 5′ end of a secondary piRNA, whose 3′ end is subsequently trimmed to proper length. These secondary piRNAs, acting in a distinct PIWI protein, then bind and cleave piRNA precursor transcripts, thereby amplifying precursor conversion into mature piRNA (**Fig. 2**).[25] Effective piRNA pathway function depends on the production of piRNAs specifically from piRNA precursor transcripts, but the signals that enforce this specificity are unknown.

Recent observations hint at a potential functional coupling between splicing and piRNA biogenesis. First, piRNAs in *Drosophila* and mouse derive from both exons and introns of their precursors,[24,54-57] suggesting that precursor transcripts are specified while still undergoing splicing. Second, when transgenes are inserted into *Drosophila* piRNA clusters, they are spliced less efficiently than when they are inserted into euchromatic regions.[58] Third, the *Drosophila* splicing and nuclear export factor UAP56 binds to piRNA precursor transcripts and is required for precursor processing specificity.[59] In a viable UAP56 allele, piRNA production from genomic clusters is reduced, whereas piRNA production from protein-coding genes is increased.[59] The subnuclear localization of UAP56 suggests a model in which UAP56 identifies precursors transcribed from heterochromatic piRNA clusters, then shuttles them to Vasa, a protein that coordinates piRNA processing in the perinuclear nuage.[59]

The above observations suggest a speculative model in which the inefficient splicing of piRNA precursors—caused either by their intrinsic sequence features or by the influence of local heterochromatin at piRNA clusters—contributes to their selection by UAP56 (**Fig. 2**). Intriguingly, the *C. elegans* UAP56 ortholog acts in cooperation with exon junction complex components, some of which are deposited on transcripts prior to the completion of splicing,[60] in order to divert unspliced transcripts away from the canonical nuclear export pathway.[61] This finding demonstrates the potential for splicing efficiency to dictate a transcript's fate in a UAP56-dependent manner, and may inform studies of the piRNA processing mechanism, which itself requires exon junction complex components.[62]

Whether piRNA loci that do not require UAP56 for their processing are also marked by the suppression of splicing signals has not yet been examined. These loci, which include protein-coding genes such as *traffic jam*,[59,63,64] do not universally contain annotated introns. Nevertheless, as cryptic splicing signals are found in many intronless transcripts, the spliceosome might still bind and influence the fate of these transcripts. Therefore, an assessment of spliceosome occupancy on these and other piRNA precursors will be critical to test the hypothesis that inefficient splicing identifies substrates for piRNA processing.

Disparities in the Strength of RNA Processing Signals May Generally Distinguish Transposons from Host Genes

While direct evidence that stalled spliceosomes signal for small RNA-mediated genome defense exists thus far only in the yeast *C. neoformans*, there are emerging hints that the spliceosome's role in RNA silencing may be conserved. This is in addition to roles for splicing factors in the biogenesis of small RNAs that do not defend the genome, such as *S. pombe* siRNA[65-67] and human miRNA.[68-70] What advantages might be gained by using splicing signal strength to determine templates for small RNA biogenesis?

In the context of genome defense, directing RNA silencing against inefficiently spliced transcripts might be advantageous for transposon targeting because the evolutionary history of transposons is distinct from that of host genes. Specifically, transposons that enter a host genome by horizontal transfer have limited time to adapt to the new host's splicing preferences, which include particular splice site sequences, intron size, and exonic splicing enhancer sequences.[71-73] Therefore, horizontal transfer from organisms with distinct optimal *cis*-acting splicing signals might contribute to the inefficient splicing of transposon-derived transcripts, as observed in *Cryptococcus*, and may mark these transcripts as targets for spliceosome-coupled small RNA biogenesis.[27] Although the spliceosome could potentially be evaded by intronless transposons, such a resistance mechanism would be hindered by the fact that introns are required for robust gene expression in many systems,[74] including *C. neoformans*,[75] which has a particularly intron-rich genome.[76]

Another advantage to the use of inefficient splicing as a guide for genome defense is that transposons may broadly utilize suboptimal splice sites in order to mitigate the negative fitness consequences of their mobilization. Multiple transposon families encode cryptic splicing signals at their termini that are only utilized upon insertion into active host genes.[77-81] These are thought to limit the negative impact of transposons on host gene expression by removing transposon sequences at the RNA level.[80,82] Importantly, for cases in which these splicing events are incompatible with transposon protein expression, inefficient splicing may be favored as a compromise between host gene expression and transposon mobility.[83] The unusual purpose of transposon splice sites may thus betray them to genome defense mechanisms that target stalled splicing.

Splicing signals are one of many gene expression signals that might exhibit disparity between transposons and host genes because of their distinct evolutionary histories and gene expression strategies. For instance, kinetic competitions could be applied not only to pre-mRNA splicing, but also to a nascent transcript's efficiency of transcription, 5′ capping, termination, polyadenylation, nuclear export, or translation. Hints that such mechanisms may exist are seen in plant systems where defects in transcript termination, polyadenylation, and capping have been associated with enhanced RNA silencing activity.[41,42] Anecdotal observations indicate that at least some transposon families tend to be capped and polyadenylated less efficiently than host genes, supporting the use of these gene expression signals to identify foreign elements.[84-86] Since many of the aforementioned steps in RNA processing are functionally coupled,[87] a kinetic competition, even if applied only to splicing, might remain sensitive to transposon-specific defects in other steps of gene expression.

Conclusions

The maintenance of eukaryotic genome stability necessitates the recognition and suppression of transposable elements. In organisms from protist to human, this function is performed by RNA silencing pathways. Our recent work has identified stalled splicing as a signal for small RNA biogenesis, thereby defining RNA processing efficiency as a metric by which transposons can be distinguished from host genes.[27] Transposons may be identified by a similar logic in other systems, as splicing factors have been found to broadly influence the function of RNA silencing pathways that carry out genome defense, including the plant and worm siRNA pathways and the *Drosophila* piRNA pathway. Although the mechanisms underlying these effects remain to be elucidated, recent studies suggest that efficient splicing protects host transcripts from RNA silencing in plants, whereas inefficient splicing is associated with the chromosomal clusters that encode piRNA sequences in *Drosophila*. We speculate that disparities in RNA processing signal strength between transposons and host genes has driven the evolution of genome defense strategies, including some that are coupled to the spliceosome.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank members of the Madhani lab for helpful discussions, and apologize to authors whose work we could not cite owing to space constraints. Our work is supported by the National Institutes of Health.

References

1. Huang CR, Burns KH, Boeke JD. Active transposition in genomes. Annu Rev Genet 2012; 46:651-75; PMID:23145912; http://dx.doi.org/10.1146/annurev-genet-110711-155616
2. Burns KH, Boeke JD. Human transposon tectonics. Cell 2012; 149:740-52; PMID:22579280; http://dx.doi.org/10.1016/j.cell.2012.04.019
3. Hedges DJ, Deininger PL. Inviting instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity. Mutat Res 2007; 616:46-59; PMID:17157332; http://dx.doi.org/10.1016/j.mrfmmm.2006.11.021
4. Shukla R, Upton KR, Muñoz-Lopez M, Gerhardt DJ, Fisher ME, Nguyen T, Brennan PM, Baillie JK, Collino A, Ghisletti S, et al. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell 2013; 153:101-11; PMID:23540693; http://dx.doi.org/10.1016/j.cell.2013.02.032
5. O'Donnell KA, Burns KH. Mobilizing diversity: transposable element insertions in genetic variation and disease. Mob DNA 2010; 1:21; PMID:20813032; http://dx.doi.org/10.1186/1759-8753-1-21
6. de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet 2011; 7:e1002384; PMID:22144907; http://dx.doi.org/10.1371/journal.pgen.1002384
7. Malone CD, Hannon GJ. Small RNAs as guardians of the genome. Cell 2009; 136:656-68; PMID:19239887; http://dx.doi.org/10.1016/j.cell.2009.01.045
8. Shabalina SA, Koonin EV. Origins and evolution of eukaryotic RNA interference. Trends Ecol Evol 2008; 23:578-87; PMID:18715673; http://dx.doi.org/10.1016/j.tree.2008.06.005
9. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet 2009; 10:94-108; PMID:19148191; http://dx.doi.org/10.1038/nrg2504
10. Castel SE, Martienssen RA. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet 2013; 14:100-12; PMID:23329111; http://dx.doi.org/10.1038/nrg3355
11. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 2001; 409:363-6; PMID:11201747; http://dx.doi.org/10.1038/35053110
12. Zheng Q, Ryvkin P, Li F, Dragomir I, Valladares O, Yang J, Cao K, Wang LS, Gregory BD. Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in Arabidopsis. PLoS Genet 2010; 6:e1001141; PMID:20941385; http://dx.doi.org/10.1371/journal.pgen.1001141
13. Li F, Zheng Q, Ryvkin P, Dragomir I, Desai Y, Aiyer S, Valladares O, Yang J, Bambina S, Sabin LR, et al. Global analysis of RNA secondary structure in two metazoans. Cell Rep 2012; 1:69-82; PMID:22832108; http://dx.doi.org/10.1016/j.celrep.2011.10.002
14. Sijen T, Plasterk RH. Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. Nature 2003; 426:310-4; PMID:14628056; http://dx.doi.org/10.1038/nature02107
15. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 2007; 8:272-85; PMID:17363976; http://dx.doi.org/10.1038/nrg2072
16. Yang N, Kazazian HH Jr. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol 2006; 13:763-71; PMID:16936727; http://dx.doi.org/10.1038/nsmb1141
17. Conley AB, Miller WJ, Jordan IK. Human cis natural antisense transcripts initiated by transposable elements. Trends Genet 2008; 24:53-6; PMID:18192066; http://dx.doi.org/10.1016/j.tig.2007.11.008
18. Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP. RNAi in budding yeast. Science 2009; 326:544-50; PMID:19745116; http://dx.doi.org/10.1126/science.1176945
19. Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler EL, Zapp ML, Weng Z, et al. Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science 2008; 320:1077-81; PMID:18403677; http://dx.doi.org/10.1126/science.1157396
20. Cogoni C, Macino G. Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature 1999; 399:166-9; PMID:10335848; http://dx.doi.org/10.1038/20215
21. Wang X, Hsueh YP, Li W, Floyd A, Skalsky R, Heitman J. Sex-induced silencing defends the genome of Cryptococcus neoformans via RNAi. Genes Dev 2010; 24:2566-82; PMID:21078820; http://dx.doi.org/10.1101/gad.1970910
22. Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC. An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell 2000; 101:543-53; PMID:10850496; http://dx.doi.org/10.1016/S0092-8674(00)80864-8
23. Ishizu H, Siomi H, Siomi MC. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes Dev 2012; 26:2361-73; PMID:23124062; http://dx.doi.org/10.1101/gad.203786.112
24. Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ. Developmentally regulated piRNA clusters implicate MILI in transposon control. Science 2007; 316:744-7; PMID:17446352; http://dx.doi.org/10.1126/science.1142612
25. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 2007; 128:1089-103; PMID:17346786; http://dx.doi.org/10.1016/j.cell.2007.01.043
26. Janbon G, Maeng S, Yang DH, Ko YJ, Jung KW, Moyrand F, Floyd A, Heitman J, Bahn YS. Characterizing the role of RNA silencing components in Cryptococcus neoformans. Fungal Genet Biol 2010; 47:1070-80; PMID:21067947; http://dx.doi.org/10.1016/j.fgb.2010.10.005
27. Dumesic PA, Natarajan P, Chen C, Drinnenberg IA, Schiller BJ, Thompson J, Moresco JJ, Yates JR 3rd, Bartel DP, Madhani HD. Stalled spliceosomes are a signal for RNAi-mediated genome defense. Cell 2013; 152:957-68; PMID:23415457; http://dx.doi.org/10.1016/j.cell.2013.01.046
28. Chapman KB, Boeke JD. Isolation and characterization of the gene encoding yeast debranching enzyme. Cell 1991; 65:483-92; PMID:1850323; http://dx.doi.org/10.1016/0092-8674(91)90466-C
29. Mayas RM, Maita H, Semlow DR, Staley JP. Spliceosome discards intermediates via the DEAH box ATPase Prp43p. Proc Natl Acad Sci U S A 2010; 107:10020-5; PMID:20463285; http://dx.doi.org/10.1073/pnas.0906022107
30. McClintock B. The significance of responses of the genome to challenge. Science 1984; 226:792-801; PMID:15739260; http://dx.doi.org/10.1126/science.15739260
31. Hamilton AJ, Baulcombe DC. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 1999; 286:950-2; PMID:10542148; http://dx.doi.org/10.1126/science.286.5441.950
32. Panda K, Slotkin RK. Proposed mechanism for the initiation of transposable element silencing by the RDR6-directed DNA methylation pathway. Plant Signal Behav 2013; 8: (Forthcoming); PMID:23759554; http://dx.doi.org/10.4161/psb.25206.
33. Luo Z, Chen Z. Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in Arabidopsis. Plant Cell 2007; 19:943-58; PMID:17384170; http://dx.doi.org/10.1105/tpc.106.045724
34. Ausin I, Greenberg MV, Li CF, Jacobsen SE. The splicing factor SR45 affects the RNA-directed DNA methylation pathway in Arabidopsis. Epigenetics 2012; 7:29-33; PMID:22274613; http://dx.doi.org/10.4161/epi.7.1.18782

35. Zhang C-J, Zhou J-X, Liu J, Ma Z-Y, Zhang S-W, Dou K, Huang HW, Cai T, Liu R, Zhu JK, et al. The splicing machinery promotes RNA-directed DNA methylation and transcriptional silencing in Arabidopsis. EMBO J 2013; 32:1128-40; PMID:23524848; http://dx.doi.org/10.1038/emboj.2013.49
36. Monaghan J, Xu F, Gao M, Zhao Q, Palma K, Long C, Chen S, Zhang Y, Li X. Two Prp19-like U-box proteins in the MOS4-associated complex play redundant roles in plant innate immunity. PLoS Pathog 2009; 5:e1000526; PMID:19629177; http://dx.doi.org/10.1371/journal.ppat.1000526
37. Xu S, Zhang Z, Jing B, Gannon P, Ding J, Xu F, Li X, Zhang Y. Transportin-SR is required for proper splicing of resistance genes and plant immunity. PLoS Genet 2011; 7:e1002159; PMID:21738492; http://dx.doi.org/10.1371/journal.pgen.1002159
38. Ji X, Zhou Y, Pandit S, Huang J, Li H, Lin CY, Xiao R, Burge CB, Fu XD. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell 2013; 153:855-68; PMID:23663783; http://dx.doi.org/10.1016/j.cell.2013.04.028
39. Christie M, Croft LJ, Carroll BJ. Intron splicing suppresses RNA silencing in Arabidopsis. Plant J 2011; 68:159-67; PMID:21689169; http://dx.doi.org/10.1111/j.1365-313X.2011.04676.x
40. Vermeersch L, De Winne N, Depicker A. Introns reduce transitivity proportionally to their length, suggesting that silencing spreads along the pre-mRNA. Plant J 2010; 64:392-401; PMID:21049564; http://dx.doi.org/10.1111/j.1365-313X.2010.04335.x
41. Gazzani S, Lawrenson T, Woodward C, Headon D, Sablowski R. A link between mRNA turnover and RNA interference in Arabidopsis. Science 2004; 306:1046-8; PMID:15528448; http://dx.doi.org/10.1126/science.1101092
42. Herr AJ, Molnàr A, Jones A, Baulcombe DC. Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis. Proc Natl Acad Sci U S A 2006; 103:14994-5001; PMID:17008405; http://dx.doi.org/10.1073/pnas.0606536103
43. Wypijewski K, Hornyik C, Shaw JA, Stephens J, Goraczniak R, Gunderson SI, Lacomme C. Ectopic 5′ splice sites inhibit gene expression by engaging RNA surveillance and silencing pathways in plants. Plant Physiol 2009; 151:955-65; PMID:19666706; http://dx.doi.org/10.1104/pp.109.139733
44. Ketting RF, Haverkamp TH, van Luenen HG, Plasterk RH. Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell 1999; 99:133-41; PMID:10535732; http://dx.doi.org/10.1016/S0092-8674(00)81645-1
45. Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC. The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell 1999; 99:123-32; PMID:10535731; http://dx.doi.org/10.1016/S0092-8674(00)81644-X
46. Kim JK, Gabel HW, Kamath RS, Tewari M, Pasquinelli A, Rual J-F, Kennedy S, Dybbs M, Bertin N, Kaplan JM, et al. Functional genomic analysis of RNA interference in C. elegans. Science 2005; 308:1164-7; PMID:15790806; http://dx.doi.org/10.1126/science.1109267
47. Robert VJP, Sijen T, van Wolfswinkel J, Plasterk RHA. Chromatin and RNAi factors protect the C. elegans germline against repetitive sequences. Genes Dev 2005; 19:782-7; PMID:15774721; http://dx.doi.org/10.1101/gad.332305
48. Tabach Y, Billi AC, Hayes GD, Newman MA, Zuk O, Gabel H, Kamath R, Yacoby K, Chapman B, Garcia SM, et al. Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. Nature 2013; 493:694-8; PMID:23364702; http://dx.doi.org/10.1038/nature11779
49. Warf MB, Shepherd BA, Johnson WE, Bass BL. Effects of ADARs on small RNA processing pathways in C. elegans. Genome Res 2012; 22:1488-98; PMID:22673872; http://dx.doi.org/10.1101/gr.134841.111
50. Aravind L, Watanabe H, Lipman DJ, Koonin EV. Lineage-specific loss and divergence of functionally linked genes in eukaryotes. Proc Natl Acad Sci U S A 2000; 97:11319-24; PMID:11016957; http://dx.doi.org/10.1073/pnas.200346997
51. Khurana JS, Theurkauf W. piRNAs, transposon silencing, and Drosophila germline development. J Cell Biol 2010; 191:905-13; PMID:21115802; http://dx.doi.org/10.1083/jcb.201006034
52. Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Tóth KF. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev 2013; 27:390-9; PMID:23392610; http://dx.doi.org/10.1101/gad.209841.112
53. Rozhkov NV, Hammell M, Hannon GJ. Multiple roles for Piwi in silencing Drosophila transposons. Genes Dev 2013; 27:400-12; PMID:23392609; http://dx.doi.org/10.1101/gad.209767.112
54. Nordstrand LM, Furu K, Paulsen J, Rognes T, Klungland A. Alkbh1 and Tzfp repress a non-repeat piRNA cluster in pachytene spermatocytes. Nucleic Acids Res 2012; 40:10950-63; PMID:22965116; http://dx.doi.org/10.1093/nar/gks839
55. Gan H, Lin X, Zhang Z, Zhang W, Liao S, Wang L, Han C. piRNA profiling during specific stages of mouse spermatogenesis. RNA 2011; 17:1191-203; PMID:21602304; http://dx.doi.org/10.1261/rna.2648411
56. Beyret E, Liu N, Lin H. piRNA biogenesis during adult spermatogenesis in mice is independent of the ping-pong mechanism. Cell Res 2012; 22:1429-39; PMID:22907665; http://dx.doi.org/10.1038/cr.2012.120
57. Yin H, Lin H. An epigenetic activation role of Piwi and a Piwi-associated piRNA in Drosophila melanogaster. Nature 2007; 450:304-8; PMID:17952056; http://dx.doi.org/10.1038/nature06263
58. Muerdter F, Olovnikov I, Molaro A, Rozhkov NV, Czech B, Gordon A, Hannon GJ, Aravin AA. Production of artificial piRNAs in flies and mice. RNA 2012; 18:42-52; PMID:22096018; http://dx.doi.org/10.1261/rna.029769.111
59. Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD, et al. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. Cell 2012; 151:871-84; PMID:23141543; http://dx.doi.org/10.1016/j.cell.2012.09.040
60. Gehring NH, Lamprinaki S, Hentze MW, Kulozik AE. The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. PLoS Biol 2009; 7:e1000120; PMID:19478851; http://dx.doi.org/10.1371/journal.pbio.1000120
61. Shiimori M, Inoue K, Sakamoto H. A specific set of exon junction complex subunits is required for the nuclear retention of unspliced RNAs in Caenorhabditis elegans. Mol Cell Biol 2013; 33:444-56; PMID:23149939; http://dx.doi.org/10.1128/MCB.01298-12
62. Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS, Brennecke J. The genetic makeup of the Drosophila piRNA pathway. Mol Cell 2013; 50:762-77; PMID:23665231; http://dx.doi.org/10.1016/j.molcel.2013.04.031
63. Robine N, Lau NC, Balla S, Jin Z, Okamura K, Kuramochi-Miyagawa S, Blower MD, Lai EC. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr Biol 2009; 19:2066-76; PMID:20022248; http://dx.doi.org/10.1016/j.cub.2009.11.064
64. Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi MC. A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature 2009; 461:1296-9; PMID:19812547; http://dx.doi.org/10.1038/nature08501
65. Bayne EH, Portoso M, Kagansky A, Kos-Braun IC, Urano T, Ekwall K, Alves F, Rappsilber J, Allshire RC. Splicing factors facilitate RNAi-directed silencing in fission yeast. Science 2008; 322:602-6; PMID:18948543; http://dx.doi.org/10.1126/science.1164029
66. Bernard P, Drogat J, Dheur S, Genier S, Javerzat J-P. Splicing factor Spf30 assists exosome-mediated gene silencing in fission yeast. Mol Cell Biol 2010; 30:1145-57; PMID:20028739; http://dx.doi.org/10.1128/MCB.01317-09
67. Chinen M, Morita M, Fukumura K, Tani T. Involvement of the spliceosomal U4 small nuclear RNA in heterochromatic gene silencing at fission yeast centromeres. J Biol Chem 2010; 285:5630-8; PMID:20018856; http://dx.doi.org/10.1074/jbc.M109.074393
68. Janas MM, Khaled M, Schubert S, Bernstein JG, Golan D, Veguilla RA, Fisher DE, Shomron N, Levy C, Novina CD. Feed-forward microprocessing and splicing activities at a microRNA-containing intron. PLoS Genet 2011; 7:e1002330; PMID:22028668; http://dx.doi.org/10.1371/journal.pgen.1002330
69. Trabucchi M, Briata P, Garcia-Mayoral M, Haase AD, Filipowicz W, Ramos A, Gherzi R, Rosenfeld MG. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature 2009; 459:1010-4; PMID:19458619; http://dx.doi.org/10.1038/nature08025
70. Wu H, Sun S, Tu K, Gao Y, Xie B, Krainer AR, Zhu J. A splicing-independent function of SF2/ASF in microRNA processing. Mol Cell 2010; 38:67-77; PMID:20385090; http://dx.doi.org/10.1016/j.molcel.2010.02.021
71. Ast G. How did alternative splicing evolve? Nat Rev Genet 2004; 5:773-82; PMID:15510168; http://dx.doi.org/10.1038/nrg1451
72. Warnecke T, Parmley JL, Hurst LD. Finding exonic islands in a sea of non-coding sequence: splicing related constraints on protein composition and evolution are common in intron-rich genomes. Genome Biol 2008; 9:R29; PMID:18257921; http://dx.doi.org/10.1186/gb-2008-9-2-r29
73. Kupfer DM, Drabenstot SD, Buchanan KL, Lai H, Zhu H, Dyer DW, Roe BA, Murphy JW. Introns and splicing elements of five diverse fungi. Eukaryot Cell 2004; 3:1088-100; PMID:15470237; http://dx.doi.org/10.1128/EC.3.5.1088-1100.2004
74. Le Hir H, Nott A, Moore MJ. How introns influence and enhance eukaryotic gene expression. Trends Biochem Sci 2003; 28:215-20; PMID:12713906; http://dx.doi.org/10.1016/S0968-0004(03)00052-5
75. Goebels C, Thonn A, Gonzalez-Hilarion S, Rolland O, Moyrand F, Beilharz TH, Janbon G. Introns regulate gene expression in *Cryptococcus neoformans* in a Pab2p dependent pathway. PLoS Genet 2013; 9:e1003686; PMID:23966870; http://dx.doi.org/10.1371/journal.pgen.1003686
76. Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, Vamathevan J, Miranda M, Anderson IJ, Fraser JA, et al. The genome of the basidiomycetous yeast and human pathogen Cryptococcus neoformans. Science 2005; 307:1321-4; PMID:15653466; http://dx.doi.org/10.1126/science.1103773
77. Belancio VP, Roy-Engel AM, Deininger P. The impact of multiple splice sites in human L1 elements. Gene 2008; 411:38-45; PMID:18261861; http://dx.doi.org/10.1016/j.gene.2007.12.022

78. Menssen A, Höhmann S, Martin W, Schnable PS, Peterson PA, Saedler H, Gierl A. The En/Spm transposable element of Zea mays contains splice sites at the termini generating a novel intron from a dSpm element in the A2 gene. EMBO J 1990; 9:3051-7; PMID:2170105
79. Purugganan M, Wessler S. The splicing of transposable elements and its role in intron evolution. Genetica 1992; 86:295-303; PMID:1334914; http://dx.doi.org/10.1007/BF00133728
80. Purugganan MD. The Splicing of Transposable Elements: Evolution of a Nuclear Defense Against Genomic Invaders? In: Syvanen MaK, C.I., ed. Horizontal Gene Transfer. San Diego, CA: Academic Press, 2002:187-95.
81. Rushforth AM, Anderson P. Splicing removes the Caenorhabditis elegans transposon Tc1 from most mutant pre-mRNAs. Mol Cell Biol 1996; 16:422-9; PMID:8524324
82. Wessler SR. The splicing of maize transposable elements from pre-mRNA--a minireview. Gene 1989; 82:127-33; PMID:2555263; http://dx.doi.org/10.1016/0378-1119(89)90037-1
83. Purugganan MD. Transposable elements as introns: evolutionary connections. Trends Ecol Evol 1993; 8:239-43; PMID:21236158; http://dx.doi.org/10.1016/0169-5347(93)90198-X
84. Jiao Y, Riechmann JL, Meyerowitz EM. Transcriptome-wide analysis of uncapped mRNAs in Arabidopsis reveals regulation of mRNA degradation. Plant Cell 2008; 20:2571-85; PMID:18952771; http://dx.doi.org/10.1105/tpc.108.062786
85. Dewannieux M, Heidmann T. Role of poly(A) tail length in Alu retrotransposition. Genomics 2005; 86:378-81; PMID:15993034; http://dx.doi.org/10.1016/j.ygeno.2005.05.009
86. Chang W, Schulman AH. BARE retrotransposons produce multiple groups of rarely polyadenylated transcripts from two differentially regulated promoters. Plant J 2008; 56:40-50; PMID:18547398; http://dx.doi.org/10.1111/j.1365-313X.2008.03572.x
87. Neugebauer KM. On the importance of being co-transcriptional. J Cell Sci 2002; 115:3865-71; PMID:12244124; http://dx.doi.org/10.1242/jcs.00073