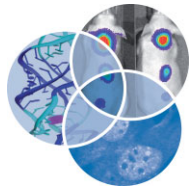
Advanced Review



Lives that introns lead after splicing

Jay R. Hesselberth*

After transcription of a eukaryotic pre-mRNA, its introns are removed by the spliceosome, joining exons for translation. The intron products of splicing have long been considered 'junk' and destined only for destruction. But because they are large in size and under weak selection constraints, many introns have been evolutionarily repurposed to serve roles after splicing. Some spliced introns are precursors for further processing of other encoded RNAs such as small nucleolar RNAs, microRNAs, and long noncoding RNAs. Other intron products have long half-lives and can be exported to the cytoplasm, suggesting that they have roles in translation. Some viruses encode introns that accumulate after splicing and play important but mysterious roles in viral latency. Turnover of most lariat-introns is initiated by cleavage of their internal 2′-5′ phosphodiester bonds by a unique debranching endonuclease, and the linear products are further degraded by exoribonucleases. However, several introns appear to evade this turnover pathway and the determinants of their stability are largely unknown. Whereas many stable intron products were discovered serendipitously, new experimental and computational tools will enable their direct identification and study. Finally, the origins and mechanisms of mobility of eukaryotic introns are mysterious, and mechanistic studies of the intron life cycle may yield new insights into how they arose and became widespread.

How to cite this article:
WIREs RNA 2013, 4:677–691. doi: 10.1002/wrna.1187

LIFE AND DEATH OF INTRONIC RNA

Biogenesis of Intronic RNA Derived from Splicing

Eukaryotic genomes contain numerous introns that divide their genes into discontinuous pieces. Cells exert tremendous energy and regulatory control to transcribe these genes as large pre-mRNAs, but only a fraction of the nucleotides polymerized by RNA polymerase II end up as functional spliced mRNAs. The bodies of many primary transcripts tend to be intron-rich,[1] and the entire locus must be transcribed to access the protein-coding exons and untranslated regions of mRNAs. Once introns are spliced out, what other purpose might these large and abundant RNAs serve? Many intron products of splicing serve as precursors for processing of small RNAs. In addition, a number of introns accumulate in cells after splicing that have no ascribed functional roles. Thus, key questions about the birth, death, stabilization, and function of these introns after splicing remain to be answered.

Introns are removed from primary transcripts by the spliceosome—a large ribonucleoprotein complex that precisely excises each intron in a pre-mRNA and joins its flanking exons. Spliceosomal removal of introns occurs in two steps. In the first step, the 2′-hydroxyl of an adenosine in the intron (the future 'branch point') serves as a nucleophile to attack the phosphodiester at the 5′ splice site, creating a new 3′-hydroxyl on the 5′ exon, and a lariat-intermediate with an internal 2′-5′ phosphodiester bond. In the second step, the 3′-hydroxyl of the cleaved 5′-exon attacks the 3′ splice site, ligating the exons and generating the lariat-intron product.

Both the lariat-intermediate and lariat-intron are 'branched RNAs'—they contain internal 2′-5′ phosphodiester bonds (Figure 1). Throughout this review, we refer to specific branched RNAs in

*Correspondence to: jay.hesselberth@ucdenver.edu
Department of Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical School, Aurora, CO, USA

Conflict of interest: The author has declared no conflicts of interest for this article.

context, but the broader principles of branched RNA production, turnover, and function are likely applicable to each member of this class, and possibly other types of RNA as well.

Branched RNA products were identified in early studies of pre-mRNA splicing as anomalously migrating intermediates upon incubation of splicing extracts with pre-mRNA substrates.[2–5] These molecules were shown to contain 2′-5′ phosphodiester linkages at internal positions. In addition to the lariat-intron products of splicing, a variety of other branched RNA forms have been observed during splicing *in vitro* and *in vivo*. *In vitro* splicing of β-globin pre-mRNA with human splicing extract produced three branched RNA forms: lariat-intermediate, lariat-intron, and a shorter form of the lariat-intron lacking a portion of its 3′ tail (i.e., a nearly circular RNA)[6] (Figures 1 and 2(a)). These 'trimmed' lariat-introns were also observed during purifications of debranching activity from human splicing extracts, and their appearance was attributed to an additional exoribonuclease that was biochemically separable from debranching activity[7,8] (discussed later).

Branched RNA is a product of SL *trans*-splicing, a type of RNA splicing found in diverse eukaryotes that uses a different splicing machinery to join two independently transcribed RNAs.[9] In this reaction, a noncoding exon derived from a small, independently transcribed spliced-leader RNA (the SL RNA) is ligated to a target pre-mRNA *in trans* by spliceosomal catalysis, forming the mature 5′ end of the SL-mRNA and branched (Y-form) products[10] (Figures 1 and 2(b)).

Trans-splicing has also been documented between independently transcribed protein-coding pre-mRNAs. In this genic *trans*-splicing, two separate pre-mRNAs undergo splicing to join portions of the precursor mRNA into a single transcript. The *Drosophila mod* gene undergoes genic *trans*-splicing, joining exons transcribed from opposite strands at the same locus into a single mRNA transcript.[11] Other examples of genic *trans*-splicing have been documented in silkworms,[12] Giardia[13] and humans.[14] One study also identified putative regions of base-pairing in the introns of *trans*-spliced gene pairs, suggesting that these interactions provide the specificity for genic *trans*-splicing.[15] Spurred by these anecdotal observations, a larger-scale search for other examples of genic *trans*-splicing in *Drosophila* found only a few *bona fide trans*-splicing events, and cautioned that many of these events arise from artifacts attributable to library preparation and suggested that careful validation of putative individual genic *trans*-splicing events must be performed.[16] *In vitro*, the spliceosome is capable of ligating two exons together *in trans* under the appropriate conditions,[17,18] suggesting a mechanism by which genic *trans*-spliced mRNAs are produced. The intron products of specific *genic* trans-splicing events have not been detected, but are presumably subject to the same turnover pathways that degrade RNA products after normal pre-mRNA splicing.

Group II self-splicing introns are a class of catalytic RNA found in the organelle genomes of fungi, plants, and protists, and encode all of the necessary components to mobilize and reintroduce themselves into genomic DNA.[19] During their mobilization, group II introns generate a lariat-intron using chemistry that is identical to that of the spliceosome.[19] Once mobilized, the group II lariat-intron remains associated with protein factors, which subsequently facilitate its reintegration into other genomic loci. Whereas there is abundant evolutionary evidence of mobilization and reintegration of group

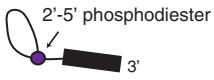
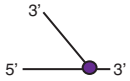
Branched RNA class		Production	Turnover	Functions after splicing
Lariat-intermediate		1st step of splicing	Dbr1, then cytoplasmic exosome	Translation
Lariat-intron		2nd step of splicing	Dbr1	Precursor for small RNA processing, viral latency factor
Circular-intron		Processing of lariat-intron or lariat-intermediate by unknown mechanism	Not characterized	Unknown
Y-form		SL *trans*-splicing	Not characterized, presumably Dbr1-mediated	Unknown, contains transcription start site

FIGURE 1 | Classification of branched RNA. Summary of branched RNA classes and their biogenesis, turnover, and function.

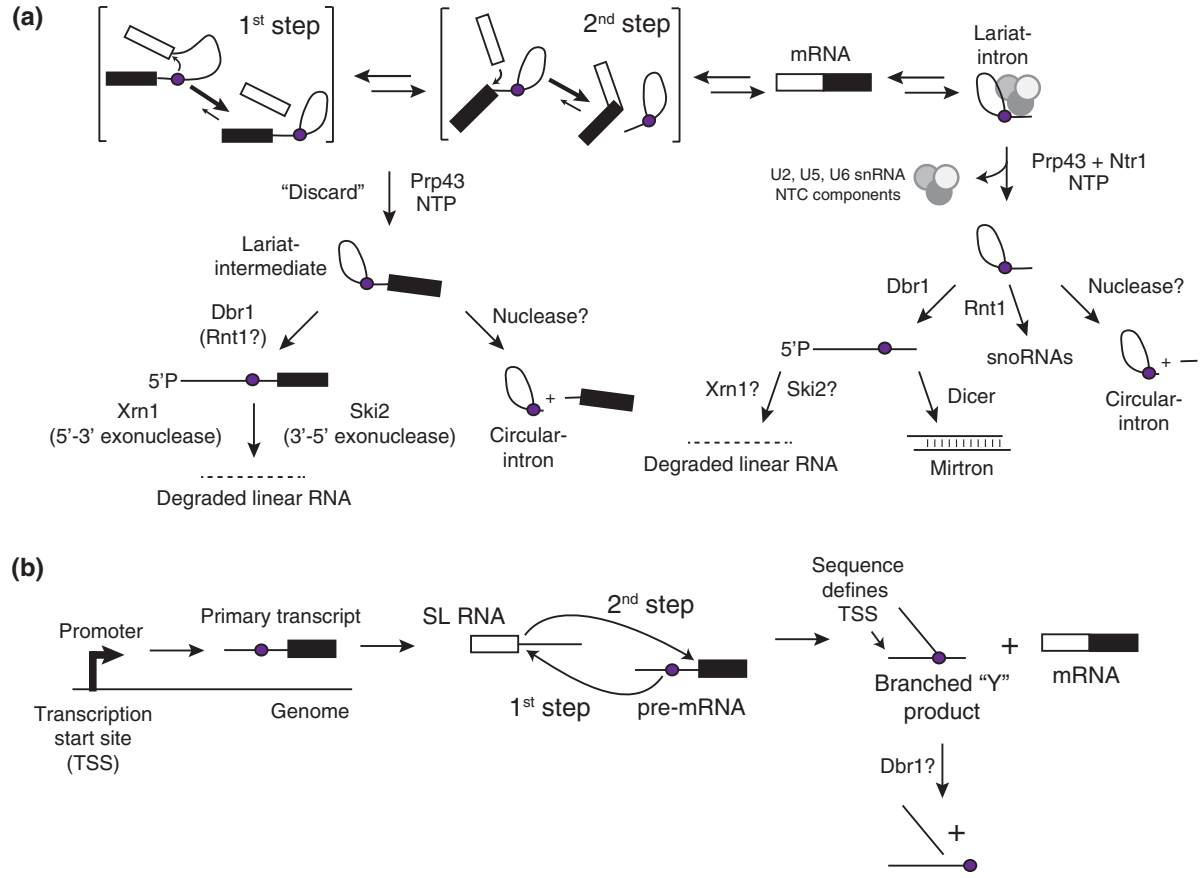


FIGURE 2 | Creation and destruction of branched RNA. (a) The first two steps of genic splicing yield a lariat-intermediate and lariat-intron, respectively. The lariat-intermediate can be discarded by Prp43, and is subject to further turnover by Dbr1, or possibly an unidentified nuclease activity. Debranched lariat-intermediates can subsequently be degraded by Xrn1 (a 5′-3′ exonuclease) or Ski2 (a component of the cytoplasmic 3′-5′ exoribonuclease). Lariat-introns are disassembled from spliceosomal complexes by Prp43, and can be further processed by debranching, endonucleolytic cleavage by Rnt1, or additional nucleolytic pathways. Some lariat-introns are precursors for small nucleolar RNA (snoRNA) or mirtrons processing. (b) In SL *trans*-splicing, primary transcripts are transcribed and are spliced to the SL RNA, yielding a messenger RNA (mRNA) and a Y-form branched product, which is presumably degraded by Dbr1. The sequence of the 'outron' portion of the Y-form branched RNA represents the site of transcription initiation, and can be used to annotation transcription start sites.

II introns,[20] little is known about how these lariat-introns interact with host machinery to promote or prevent their mobility. However, the ability of lariat-intron-containing ribonucleoproteins to integrate into arbitrary genomic loci has been exploited to create genome-engineering tools for targeted and untargeted disruptions.[21,22]

Turnover of the Intron Products of Splicing

A unique feature of the lariat-intermediate and lariat-intron products of splicing is their internal 2′-5′ phosphodiester bond, which protects the bulk of the structure from turnover by cellular exoribonucleases. Debranching of lariat-introns to linear form is catalyzed by a 2′-5′ phosphodiesterase called the lariat-debranching enzyme. In budding yeast, the lariat-debranching enzyme Dbr1 was identified in a genetic screen for modifiers of retrotransposition (Box 1),[23] and was shown to have 2′-5′ phosphodiesterase activity.[24] Biochemical studies of *Saccharomyces cerevisiae* Dbr1 activity showed it is a metallophosphodiesterase with catalytic residues that coordinate two metal ions, with preference for manganese.[25] Dbr1 is not essential for viability in budding yeast, but *dbr1*Δ mutants of the fission yeast *Schizosaccharomyces pombe* grow slowly and adopt an elongated shape[26] and *DBR1* is required for embryogenesis in the mustard plant *Arabidopsis thaliana*[27]; thus it seems likely that Dbr1 function is required for viability in metazoans. In organisms with SL *trans*-splicing, turnover of the Y-form products is presumably initiated by the lariat-debranching enzyme (Figures 1 and 2(b)), although this has not been tested

experimentally. Dbr1 homologs from fission yeast,[25] mouse,[25] and human[28] can complement a *dbr1*Δ budding yeast strain, showing conservation of the role of Dbr1 in the turnover of lariat-introns.

BOX 1

FUNCTIONS OF DBR1 OUTSIDE OF CANONICAL LARIAT-INTRON TURNOVER

Besides its role in lariat-intron turnover, Dbr1 plays a mysterious—and controversial—role in retrotransposition. Recent studies have suggested that a branched RNA intermediate is involved in retroelement transposition.[29] This followed from the initial discovery that Dbr1 is required for efficient retrotransposition in *S. cerevisiae*.[23,30] Consistent with this, Dbr1 is required for efficient replication of the HIV-1 retrovirus in human cells.[31] However, other studies have questioned the validity of these putative branched intermediates[32] and their role in reverse transcriptase-mediated cDNA conversion.[33] In addition, a proposed role for 2′-5′ phosphodiester bonds in the initiation of Ty1 retrotransposon replication was not Dbr1-dependent.[34] Thus, the precise role for Dbr1—and RNA debranching *per se*—in retroelement mobility remains elusive, but an understanding of this mechanism could open up new avenues for mechanistic studies of retrotransposition and may improve our understanding of an important step in the retrovirus life cycle.

Recently, loss of Dbr1 in *S. cerevisiae* was shown to suppress the toxicity of heterologously expressed TDP-43, a protein which aggregates in cytoplasmic foci in cell lines from patients with amyotrophic lateral sclerosis (ALS, or Lou Gehrig's disease).[35] In yeast, *dbr1*Δ causes the accumulation of lariat-introns in the cytoplasm, where they sequester TDP-43 protein in cytoplasmic foci, preventing its pathologic association with another predicted RNA target. Currently it is unclear whether branched RNA *per se* has a specific role in abrogating the deleterious effects of TDP-43 toxicity, or whether *dbr1*Δ is simply the easiest way to increase RNA to superphysiologic levels, enabling the sequestration of TDP-43 and suppression of its toxicity. Notably, this phenotype was unique to Dbr1: mutation of other key RNA turnover factors failed to suppress TDP-43 toxicity. Finally, the study demonstrated that inhibition of Dbr1 function in mammalian cells can suppress TDP-43 toxicity,[35] offering a plausible therapeutic angle for treatment of ALS patients.

Debranching of lariat-introns is the initiating step of turnover for most introns.[23] The 2′-5′ phosphodiesterase activity of Dbr1 converts lariat and branched RNAs to a linear RNA molecule with a 5′ phosphate, and disruption of *DBR1* causes a dramatic accumulation of lariat RNA *in vivo*.[23,25] Once lariat-introns are debranched to their linear forms, they can be further degraded by cellular exonucleases. In addition, given that splicing is a nuclear process, it is reasonable to assume that most lariat-intron turnover occurs in the nucleus.

In addition to lariat-introns, the lariat-intermediate products of the first step of splicing are released by the spliceosome—following unproductive splicing—and degraded.[36] While there have been no direct studies of the turnover of lariat-introns, studies of the turnover of lariat-intermediates have informed our understanding of branched RNA turnover in the cell. After debranching, turnover of linearized lariat-intermediates is mediated by the 5′-3′ exoribonuclease Xrn1[37] (Figure 2 (a)). This discovery was surprising insofar as Xrn1 is a cytoplasmic 5′-3′ exoribonuclease; disruption of the nuclear 5′-3′ exoribonuclease (Rat1 in budding yeast) had minimal effect on levels of lariat-intermediates *in vivo*.[37] Genomewide analysis of 5′-phosphate intermediates that accumulate in the absence of both Xrn1 and a decapping enzyme subunit identified abundant debranched lariat-intermediates from a number of pre-mRNAs,[38] suggesting extensive turnover of these intermediates in wild type cells.

Other RNA turnover pathways have been associated with the turnover of intronic RNAs. Some lariat-introns in *S. cerevisiae* are turned over via cleavage of internal stem loops by the ribonuclease III homolog Rnt1.[39] Ribonuclease III activity has also been associated with the degradation of spliced group II lariat-introns in plants,[40] suggesting that this mode of intron turnover is widely used. In *S. cerevisiae*, Trf4 and Trf5 are noncanonical poly(A) polymerases that interact with the nuclear exosome to promote the turnover of short-lived transcripts.[41] Trf4 and Trf5 are required for the turnover of a subset of lariat-introns, but surprisingly this turnover does not appear to require the polyadenylation activity of Trf4, insofar as a catalytically inactive version of Trf4 can complement the intron-accumulation phenotype of *trf4*Δ cells.[42]

During the identification of lariat-introns during *in vitro* splicing, an activity was identified that converted lariat-introns to nearly circular form.[7,8,43] The identity of this putative nuclease is still unknown, but its activity could serve to facilitate the turnover of specific lariat-introns by removing a large portion

of its tail, possibly improving accessibility of the 2′-5′ phosphodiester to the lariat-debranching enzyme. While these early biochemical studies suggested that this molecule derives from lariat-intron,[8] in principle, such an activity might also cleave downstream of the branch point of lariat-intermediates, liberating a circular-intron and linear exon (Figure 2(a)) and precluding reversal of the second catalytic step of splicing and regeneration of the pre-mRNA substrate.[44]

ROLES OF STABLE BRANCHED RNA *IN VIVO*

Determinants of Intron Stability

Stable lariat-introns have been identified that appear to evade normal turnover pathways (Figures 2(a) and 3). One of these is spliced from the TCR-β gene and accumulates in mouse and human T-cells.[46] Systematic studies of the rates of production and turnover of this TCR-β intron showed that it has a half-life of 6 min in cells, nearly twice as long as the half-life of spliced mRNA transcript.[51] These studies also revealed a correlation between the cellular localization of spliced introns and their physical location in the primary transcript: subcellular fractionation showed that introns near the 5′ end of the gene—and likely co-transcriptionally spliced—were found in the cytoplasm, whereas introns near the 3′ end of the gene—likely spliced after transcription—were mainly localized to the nucleus.[52] Thus, the cellular localization of spliced lariat-introns seems linked to co-transcriptional nuclear export events that are coupled to splicing. Other anecdotal examples of stable introns in animal cells exist from highly expressed genes (e.g., β-globin and immunoglobulin κ).[4,53] In the future, new genomic approaches for studying splicing—and particularly the intron products of splicing—will likely turn up more examples of stable lariat-intron products (discussed later).

Lariat-introns are protected from degradation by spliceosomal factors that remain associated with them after splicing (Figure 2(a)). Lariat-introns are released from the spliceosome complexed with a number of spliceosomal components, including the U2, U5, and U6 snRNAs, and specific protein components of the Nineteen Complex.[54,55] Prior to their release from these spliceosomal complexes by the DEAH-box ATPase Prp43, lariat-introns are protected from debranching by Dbr1.[56] In addition, there is some evidence supporting a role for the DEAH-box ATPase Brr2 and GTPase Snu114 in the disassembly of lariat-intron complexes.[55,57]

Lariat-intermediates are released from the spliceosome following nonproductive splicing events via the discard pathway.[36,45] Discarded lariat-intermediates are created by inefficient splicing of pre-mRNAs, and represent a possible regulatory step that limits productive splicing under specific conditions. Indeed, a slow second step of splicing for the *S. pombe* telomerase RNA subunit triggers discard, releasing the 5′ exon as a functional telomerase RNA subunit.[58] In *S. cerevisiae*, splicing of many pre-mRNAs is altered in response to changes in the environment,[59] suggesting extensive regulation at both the first and second step of splicing.[60] In addition, lariat-intermediates released from the spliceosome prior to the second step of splicing can be exported to the cytoplasm and can be made competent for translation by the addition of an internal ribosomal entry site[45] (Figure 3). The biological roles of cytoplasmic lariat-intermediates are not known, but suggest that regulated discard during splicing may create a variety of RNA forms from a single transcript, even beyond canonical alternatively spliced exons. A recent study showed that the spliceosome itself is capable of debranching lariat-intermediates in competition with reversal of the first step of splicing.[61] Similar to lariat-intermediates, lariat-introns have been found in the cytoplasm,[50,52] suggesting that they may also have roles in translation or RNA stability.

Lariat RNAs as RNA Processing Intermediates

Once they are removed from their parent pre-mRNA, some lariat-introns are further processed to liberate smaller RNAs encoded within them (Figure 2(a)). These RNAs include microRNAs, snoRNAs, and long-noncoding RNAs (lncRNAs), as well as a number of small RNAs with no known function. Processing of these small RNAs is accomplished by cellular ribonucleases to liberate the RNA products.

The lariat-debranching enzyme Dbr1 is required for the processing of intron-encoded small nucleolar RNAs (snoRNAs) in *S. cerevisiae*.[62] In humans, introns of the *UHG* (U22 host) gene encode eight snoRNAs, whereas the spliced exons have no known functional role. Notably, the introns of the *UHG* gene exhibit greater evolutionary conservation than the exons.[63] A recently discovered class of lncRNAs derive from processing of snoRNA-containing introns.[64] These 'sno-lncRNAs' derive from processing of introns with two flanking snoRNAs, such that the sequence between the snoRNAs remains stable, presumably protected from degradation by RNP complexes that remain at the site of snoRNA

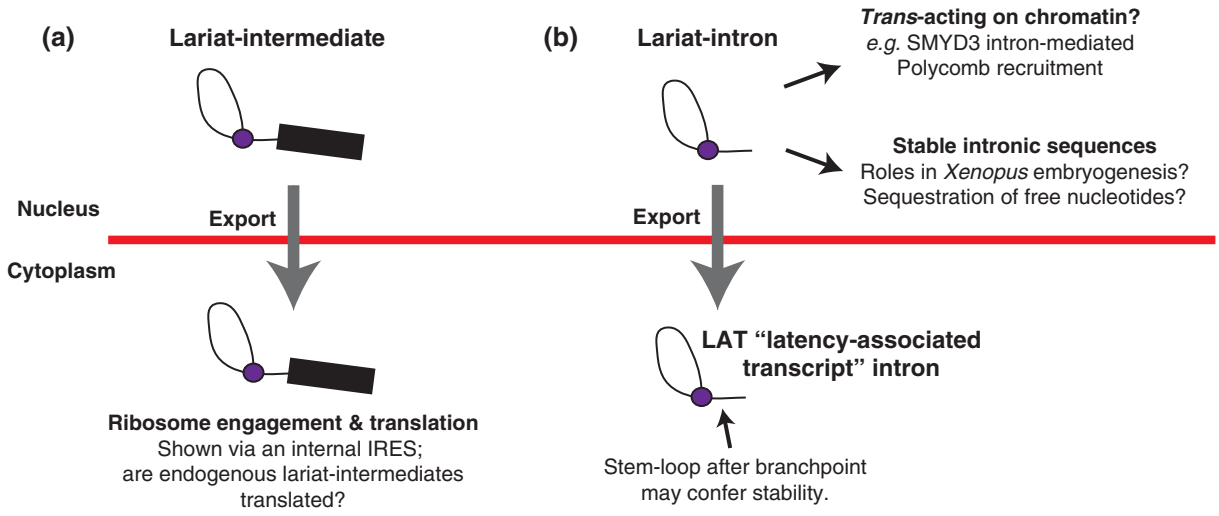


FIGURE 3 | Alternative fates of lariat-intermediates and lariat-introns after splicing. (a) After splicing, lariat-intermediates can be exported to the cytoplasm and translated via an IRES,[45] but it is not known whether endogenous lariat-intermediates are translated. (b) Lariat-introns can accumulate in the nucleus,[46] where they have been associated with repressive chromatin complexes,[47] and are restricted from export in *Xenopus* oocytes.[48] Stable lariat-introns might sequester free nucleotides for further rounds of transcription. Latency-associated introns from some viruses are stable in cells, where they can promote viral latency[49] and associate with ribosomes.[50]

processing. Thus the cellular components of snoRNA processing appear to protect intervening RNAs from unintended degradation—a conceptually novel idea that may be widely used during small RNA biogenesis.

RNA debranching by Dbr1 is also required for the processing of microRNA precursors encoded within introns called 'mirtrons'[65,66] (Figure 2(a)). Canonical miRNA processing pathways begin with linear precursor transcripts that adopt a stem-loop structure, enabling their recognition and processing by Drosha and then Dicer.[67] In contrast, mirtron biogenesis begins with lariat-introns derived from spliceosomal processing of a pre-mRNA. One processing pathway for mirtrons involves the debranching of lariat precursors to linear form, followed by endonucleolytic cleavage by Dicer, bypassing Drosha cleavage.[65,66] A second processing pathway uses exonucleolytic trimming to process 'tailed' mirtrons.[68] Systematic searches for products of mirtron processing have identified numerous examples of mirtron-like miRNA products.[69] There are only a handful of studies demonstrating functional significance of specific mirtrons, but new genomic approaches will likely yield new roles for mirtrons.[70]

The branched 'Y-form' products of SL *trans*-splicing contain the site of transcription initiation of the pre-mRNA substrate (Figure 2(b)). These branched products are degraded rapidly, and have been difficult to identify systematically due to their short half-lives. However, recent global approaches in *C. elegans* have identified these branched RNA products, which are useful to determine the site of transcription initiation of the original pre-mRNA prior *trans*-splicing[71]; transcription start sites have also been mapped by identifying unspliced pre-mRNAs directly.[72] Mapping of 5′ ends of specific *C. elegans* pre-mRNAs showed that transcriptional start sites can be hundreds of base pairs upstream of the site of spliced-leader attachment,[73] creating the potential for additional RNAs to be encoded within the intervening sequence.

Virus-Encoded Latency-Associated Intron RNAs

Members of the herpesvirus family encode intronic RNAs that remain stable in host cells after splicing[74] (Figure 3(b)). Despite their discovery over two decades ago, key questions about virus-encoded stable lariat-introns remain unanswered. A 2-kilobase intron spliced from the latency-associated transcript (LAT) of herpes simplex virus-1 (HSV-1; an α-herpesvirus) accumulates in nonlinear (i.e., lariat) form,[75,76] suggesting that the branched structure of the lariat protects the LAT intron from turnover. Stable LAT introns from HSV-1 and a similar intron RNA from cytomegalovirus (a β-herpesvirus) promote persistent infection of these viruses, underscoring their importance in the viral life cycle.[77] Expression of the HSV-1 stable LAT intron protects neural cells from apoptosis, which may help the virus remain

latent in a population of cells by an unknown mechanism.[49,78] The stable LAT intron has also been associated with polysomes,[50,79] suggesting that it can be exported to the cytoplasm and translated, similar to nonproductively spliced lariat-intermediates released by the spliceosomal discard pathway.[45] However, despite its co-migration with ribosomal subunits, translation of the LAT intron is inefficient and the protein reading frame is not well conserved, raising questions about the relevance of the putative protein product *in vivo*.[50] In addition to its role in translation, the stable LAT intron contributes to the establishment or modulation of heterochromatin at the promoters of viral lytic genes, silencing viral gene expression and promoting viral latency.[80–82] The mechanism by which the stable LAT intron establishes heterochromatin is currently unclear, but may involve its recruitment of Polycomb Repressive Complexes to lytic promoters,[82] analogous to the recruitment of Polycomb Repressive Complexes by lncRNAs like HOTAIR.[83]

The molecular determinants that promote the stabilization of virus introns have been elusive. Some *cis*-sequence elements in the HSV-1 LAT intron required for stabilization have been identified: mutation of a stem-loop downstream of the preferred branch point causes the selection of sub-optimal branch point sequences, and reduces stability of the product lariat-intron.[84] Thus, it is possible that this sequence promotes the retention of specific protein complexes (e.g., U2 snRNP components or polypyrimidine tract-binding proteins) that protect the branch point sequence, preventing turnover by debranching. Alternatively, the branch point itself might be an inefficient substrate for debranching enzyme.[85] What is the functional role of the stable LAT intron and how does this RNA affect so many aspects of the virus life cycle? Does it interact with specific host RNA binding proteins to mediate these effects? As has been the case for other virus-encoded RNAs that co-opt host machinery, eg., translation mediated by internal ribosomal entry sites (IRESs),[86] a molecular understanding of the strategies underlying the stability of the LAT intron may reveal a host mechanism exploited by the virus to achieve latency, and may provide more general insights into the mechanisms responsible for the stability of other host-encoded lariat-introns.

Evolutionary Impact of Intron Products of Splicing

The evolutionary origins of spliceosomal introns are mysterious, and the timing of the expansion of introns in eukaryotic genomes is the subject of much debate. Two alternative models have been proposed for the timing of intron acquisition, termed 'introns-early' and 'introns-late'; 'introns-early' posits that introns are evolutionarily ancient and were present in both prokaryotes and eukaryotes, but were and were lost exclusively in prokaryotic lineages; 'introns–late' holds that the restriction of introns to eukaryotes is parsimonious with their acquisition later during evolution.[87]

Several theories for the mechanisms of intron gain (i.e., the acquisition of new introns in a genome) involve lariat-introns as intermediates (reviewed in Ref 87). One of the prevailing mechanisms, termed intron transposition, was proposed over two decades ago[88] to underlie the expansion of introns in genomes with existing introns, and is based on the known activities of the group II introns and the spliceosome (Figure 4). Several types of molecular evidence support intron transposition as a likely mechanism for intron mobility. For example, group II lariat-introns can be reintegrated into their original splice sites by reverse splicing, recreating the original *cis*-encoded enzyme-substrate construct.[19] Similarly, the spliceosome is capable of reversing both the first and the second catalytic steps of splicing.[44] Reverse splicing was proposed to account for spliceosomal introns in yeast U6 snRNA present near the branch site-forming ACAGAGA sequence,[89] as well as the prevalence of recent intron gains in spliceosomal protein-encoding genes in *Caenorhabditis* species.[90] The ability of lariat-introns to reintegrate into pre-mRNA transcripts illustrates the surprising versatility of group II introns and the spliceosome. However, specific examples of reverse splicing have not been identified during splicing *in vivo*; because reversal of the second step is precise and recreates the starting pre-mRNA, it is not possible to identify such events from sequence data alone.

Given current knowledge of spliceosome mechanisms, how might intron transposition be catalyzed? We consider two alternatives for spliceosome-mediated intron transposition. In the first model (Figure 4(a)), a spliceosome completes the second step of a splicing on an intron-containing pre-mRNA, and releases the spliced mRNA by the action of the DExH-box ATPase Prp22.[91,92] At this point, a new 'proto' pre-mRNA (i.e., an mRNA that did not previously have an intron) might be loaded (by multiple ATP-dependent rearrangements) into the spliceosome in the place of the previous mRNA: reversal of the second step of splicing would insert the lariat-intron into the mRNA, and reversal of the first step of splicing would create a novel pre-mRNA (Figure 4(a)). As a corollary

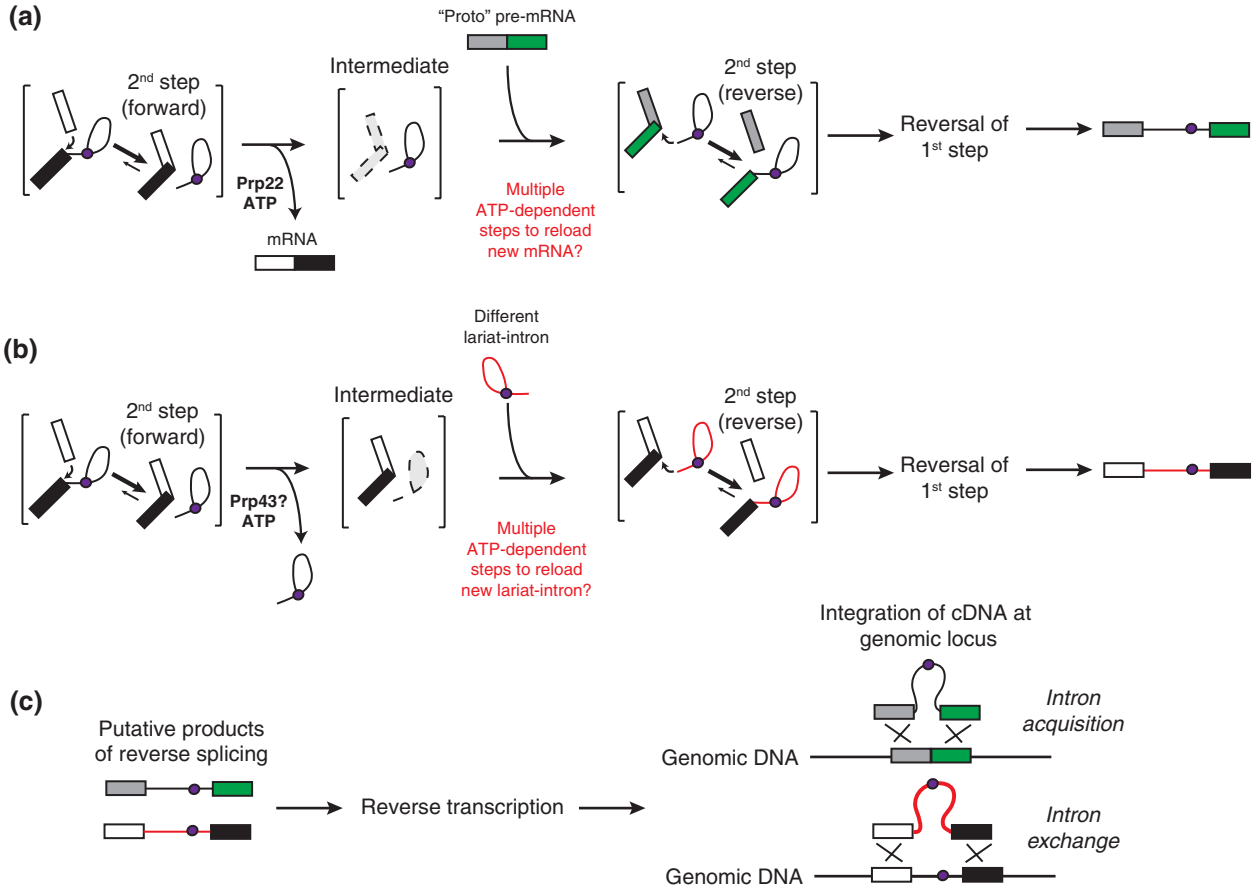


FIGURE 4 | Putative mechanisms of intron transposition. (a) Following the first step of splicing, Prp22 mediates release of the messenger RNA (mRNA) product. Prior to release of the lariat-intron, a new 'proto' pre-mRNA is loaded into this spliceosome, and reversal of the second and first catalytic steps yields a new pre-mRNA. (b) Following the first step of splicing, Prp43 mediates release of the lariat-intron without releasing the mRNA product. A second lariat-intron is loaded into this spliceosome, and reversal of the second and first catalytic steps yields a new pre-mRNA. (c) The putative products of both (a) and (b) are reverse transcribed and recombined into their original genomic locus, representing intron acquisition and intron exchange events, respectively.

to the first model, it is possible that some conditions may promote the release of a spliced lariat-intron prior to release of the mature mRNA (Figure 4(b)); this would likely involve the action of the DEAH-box ATPase Prp43, which is required for lariat-intron release.[56] This idea is incongruent with current knowledge about the ordering of Prp22 and Prp43 function (e.g., mRNA release necessarily precedes lariat-intron release, at least for some substrates[92]), but seems possible given considerable flexibility in spliceosomal structural transitions.[44] In this second model (Figure 4(b)), after release of a first lariat-intron, a second lariat-intron—previously spliced from a different mRNA—might be loaded into the spliceosome, and reversal of the second and first steps would create a novel pre-mRNA. In both of these cases, reverse transcription of the putative reverse spliced mRNAs and cDNA-mediated recombination the genomic locus of the receiving mRNA would complete an intron transposition: in the first case, intron acquisition produces a new intron in an encoded mRNA; in the second case, an intron exchange replaces one encoded intron with another (Figure 4(c)). Notably, many of these events would be detrimental to the fitness of an organism, as acquisition of introns at new sites would often yield nonfunctional mRNAs upon transcription and splicing. These mechanisms are speculative, but over evolutionary time-scales may create new mRNA forms. They also assume unknown capabilities of the spliceosome (e.g., release of lariat-intron prior to mRNA), and gloss over the enormous complexity of spliceosomal substrate recognition and loading (reviewed in[93]). Nonetheless, future study should focus on characterizing the mechanisms and regulation of reverse splicing, using experimental methods geared to selectively generate and identify reverse splicing products from *in*

vitro splicing reactions, or even identify these events *in vivo*, which would provide direct experimental support for intron transposition.

Speculative Roles for Intron Products of Splicing

Introns can directly affect splicing of pre-mRNAs in *cis*, but could also affect cellular physiology in less obvious ways. Stable introns might affect other nuclear and cytoplasmic processes by titrating available factors. For example, spliced introns remain associated with a number of hnRNP proteins, reducing the levels of these proteins available for other splicing events. The levels of available hnRNP proteins can regulate splicing,[94,95] and thus sequestration of hnRNP proteins by lariat-introns could indirectly affect further rounds of splicing of other genes. Similarly, turnover of large spliced introns might affect the availability of exoribonucleolytic activities available to act on other introns and mature mRNAs, indirectly increasing the half-lives of these molecules.

Because of their large size and abundance, intron stability could perturb nucleotide pools available for further rounds of transcription. As such, they might act as a reservoir for ribonucleotide pools, and their degradation might replenish nucleotide pools for new rounds of transcription. Along these lines, a recent study identified abundant stable intronic RNA in the nucleus of *Xenopus* oocytes, while spliced mRNAs were found predominantly in the cytoplasm[48] (Figure 3). This apparently new class of stable nuclear intron RNAs could affect a variety of nuclear processes, directly or indirectly, and their presence in oocytes and during development suggests that their biogenesis and putative roles are specific to embryogenesis.

Lariat-introns might act *in trans* after splicing to affect subsequent rounds of transcription at the locus they were transcribed from. Recently, the association of Polycomb Repressive Complex-examined member EZH2 was by crosslinked and immunoprecipitated RNAs (CLIP-seq) to identify its direct RNA interaction partners, which revealed several intronic targets.[47] Overexpression of one of these—an intron in the *SMYD3* gene—was sufficient to direct silencing of the *SMYD3* locus, suggesting that the intronic RNA targets Polycomb Repressive Complex to this locus *in trans*. Whereas this study did not address the state of these introns (e.g., either as spliced introns, or as separate transcripts), they raised the profile of intron-derived RNAs in actuating genome silencing via their association with repressive protein complexes. Given the ability of some lncRNAs to target regions of the genome other than their transcribed locus (e.g., Xist and HOTAIR), similar events may befall a number of specific spliced introns, giving them life after splicing.

New Tools for the Analysis of Branched RNA Production and Turnover

Given the serendipitous discovery of many stable lariat-introns, their systematic and global study may yield new classes of RNAs encoded within them, and may reveal new regulatory roles for introns after splicing. In addition, existing annotations of introns have relied on analysis of evolutionary conservation of protein-coding genes and the observation of specific expressed sequences.[96] Thus, new experimental and computational approaches for studying intron RNAs *in vivo* will improve our understanding of their biogenesis and turnover, and may provide clues into the mechanisms underlying their stability.

One way to study intron RNA products of splicing is to disrupt the turnover pathways that degrade branched RNAs. Because lariat-introns are stabilized in the absence of Dbr1, the accumulation of lariat-introns in *dbr1*Δ budding yeast has been analyzed by tiling microarrays to validate existing intron annotations and identify new ones.[97,98] This approach could be updated and expanded to take advantage of the depth and precision of next-generation DNA sequencing approaches, allowing for an unbiased survey of the abundance of lariat-introns that accumulate when debranching is inhibited. A caveat of these experiments is that certain systems may allow reduction of Dbr1 (e.g., by siRNA knockdown) but not its elimination, which may not be sufficient to stabilize all lariat-intron products. Nonetheless, these analyses would complement ongoing large-scale studies of the exon-containing mRNA products of splicing, providing additional validation of gene expression programs and mRNA isoforms. In addition, these approaches might yield new types of RNAs subject to regulation by branched RNA turnover pathways.

Pharmacological manipulation of transcription and splicing might also yield fresh insights into intron turnover. Inhibition of RNA polymerase II with the adenosine analog DRB was found to reduce intron degradation in HeLa cells by an unknown post-transcriptional mechanism.[99] Whereas these early studies focused on the characterization of a few introns by northern blotting and ribonuclease protection assay, interrogation of intron degradation products by next-generation DNA sequencing may yield important

new information on the rates of intron decay in mammalian cells.

As mentioned previously, the molecular determinants of lariat-intron stability after splicing are largely unknown. Some of the stability is likely due to proteins and RNA factors that remain associated with the lariat-intron after splicing. High-throughput sequencing of UV-crosslinked RNA:protein complexes (CLIP-seq) is a powerful method for identifying RNAs that interact directly with a protein of interest.[100] In this context, selection of proteins that are preferentially associated with lariat-introns may yield new insights into stability mechanisms. For example, the Ntr1 protein and components of the nineteen complex associate with lariat-introns in post-splicing complexes[54,101]; analysis of these components by CLIP-seq could provide a high resolution view of sequences in lariat-introns that are protected with a given post-splicing complex. A general approach to identifying protein-binding sites in RNA, global PAR-CLIP, could identify specific regions of lariat-introns bound by proteins, without identifying the protein factors themselves.[102] In addition, the turnover of specific introns could be studied more closely by purification with RNA affinity tags such as the RNA hairpin bound by the MS2 phage coat protein.[54] CLIP-seq of components of the nuclear exosome (e.g., *S. cerevisiae* Rrp44) showed interesting patterns of enrichment over intron regions, suggesting that they are direct targets of nuclear exosome-mediated turnover.[41] In the same vein, analysis of lariat-intron turnover pathway components such as Dbr1 by CLIP-seq would certainly identify new intron targets, and similar studies of cytoplasmic exosome components may provide further evidence of its role in turnover of released lariat-intermediates.[45]

A large-scale study used 4-thiouridine pulse-labeling followed by affinity purification and high-throughput sequencing to measure the decay rates of spliced introns throughout the human genome.[103] As suggested previously,[52] the study found considerable variation in intron half-lives. In addition, the large-scale study identified a strong correlation between an intron's size and its half-life: in general, shorter introns were more stable than long introns (above 10 kb), but there is a local maximum of stability for introns in the range of 200–500 nucleotides, suggesting that this class might be stabilized by a specific mechanism.

One difficulty in identifying post-catalytic introns using DNA sequencing approaches is distinguishing between intronic sequence that derives from pre-mRNAs, lariat-intermediates, lariat-introns and their debranched forms. To address this, a new computational method was developed to search for the presence of characteristic 'swapped' sequencing reads that arise via read-through of branched structures (e.g., in lariat-introns and lariat-intermediates) during reverse transcription *in vitro*,[104] enabling the reliable detection of lariat-introns (and, in principle, lariat-intermediates) in complex populations. This approach confirmed high fidelity of selection of commonly used branch points, but some genes exhibit considerable flexibility in branch point selection. These 'swapped' reads were rare in sequencing libraries (1 in 10^6 reads), but optimization of library preparation methods or inactivation of specific intron turnover pathways might enable more sensitive detection of these intermediates, facilitating more detailed studies of intron abundance and turnover by next-generation DNA sequencing.

The unique features of lariat RNAs enable their selective enrichment and purification from complex RNA populations. Ribonuclease R, a 3′-5′ exoribonuclease, can degrade linear RNA in total RNA preparations, but is inhibited by branched RNA structures; this activity was leveraged to enrich total RNA populations for lariat-introns that are protected from digestion by internal 2′-5′ phosphodiester bonds.[105] In addition, the altered mobility of branched RNA on two-dimensional polyacrylamide gels[35,43] allows the purification of lariat-introns from linear RNA in total RNA preparations, potentially enabling their large-scale analysis by next-generation DNA sequencing. Finally, the 2′-5′ phosphodiesterase activity of Dbr1[25] could be manipulated to identify other components involved in branched RNA processing by expressing mutant forms (e.g., *S. cerevisiae* Dbr1-H86A[25,35]) and identifying affected substrates.

Validation of newly identified stable branched RNAs should be accomplished biochemically, and new molecular and synthetic tools for creating branched RNAs will be useful for studying branched RNA metabolism *in vitro*. A suite of deoxyribozymes (i.e., enzymes made of DNA) have been evolved and engineered to convert linear RNA molecules to both branched and lariat forms.[106,107] In addition, these deoxyribozymes can be used to create branched RNAs with noncanonical bases at the branch point,[108] enabling the creation of molecules that are made inefficiently during *in vitro* splicing. As expected, branched RNA molecules created using these methods are substrates for debranching by Dbr1,[106] and provide a variety of substrates for surveying the specificity in debranching by Dbr1.[24,85] In parallel, chemical synthesis of branched RNAs enables production at the milligram scale, facilitating structural studies and the design of chemical inhibitors of debranching.[109] These synthetic methods also enable the incorporation of specific RNA modifications (e.g., phosphorothioate

bonds[110]), protecting them from degradation *in vivo* and adding additional functionality for biochemical studies of branched RNA metabolism.

CONCLUSIONS

Introns are pervasive in eukaryotic genomes and likely encode undiscovered RNA forms with unexpected functions (Box 1). The specialized turnover pathways for branched RNA turnover have been co-opted to facilitate processing of a variety of small and lncRNAs. A number of introns are stable in cells after splicing: several host-encoded introns are stable, but have unknown functions, and some virus-encoded introns are required for the virus life cycle, but have mysterious molecular functions. The unique structural properties of lariat-intron products provide unique functional and regulatory opportunities, some of which are still mysterious. Despite their categorization as 'junk' RNAs not long ago, new global experimental and computational methods promise to identify and define the roles of intron products of splicing, and may reveal new roles for this class of RNA cellular metabolism.

ACKNOWLEDGMENTS

We thank R. Davis, C. Kulesza, T. Blumenthal, and C. Query for discussions and comments on the manuscript. J.H. is supported by a Damon Runyon-Rachleff Innovation Award (DRR-17-12) and an American Cancer Society Research Scholar Grant (RSG-13-216-01-DMC).

REFERENCES

1. Lynch M, Richardson AO. The evolution of spliceosomal introns. *Curr Opin Genet Dev* 2002, 12:701–710.
2. Domdey H, Apostol B, Lin RJ, Newman A, Brody E, Abelson J. Lariat structures are *in vivo* intermediates in yeast pre-mRNA splicing. *Cell* 1984, 39:611–621.
3. Rodriguez JR, Pikielny CW, Rosbash M. *In vivo* characterization of yeast mRNA processing intermediates. *Cell* 1984, 39:603–610.
4. Zeitlin S, Efstratiadis A. *In vivo* splicing products of the rabbit β-globin pre-mRNA. *Cell* 1984, 39: 589–602.
5. Padgett RA, Konarska MM, Grabowski PJ, Hardy SF, Sharp PA. Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science* 1984, 225:898–903.
6. Ruskin B, Krainer AR, Maniatis T, Green MR. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell* 1984, 38: 317–331.
7. Arenas J, Hurwitz J. Purification of a RNA debranching activity from HeLa cells. *J Biol Chem* 1987, 262:4274–4279.
8. Krämer A, Keller W. Purification of a protein required for the splicing of pre-mRNA and its separation from the lariat debranching enzyme. *EMBO J* 1985, 4:3571–3581.
9. Lasda EL, Blumenthal T. Trans-splicing. *WIREs RNA* 2011, 2:417–434.
10. Pettitt J, Harrison N, Stansfield I, Connolly B, Müller B. The evolution of spliced leader trans-splicing in nematodes. *Biochem Soc Trans* 2010, 38:1125–1130.
11. Dorn R, Reuter G, Loewendorf A. Transgene analysis proves mRNA trans-splicing at the complex mod(mdg4) locus in Drosophila. *Proc Natl Acad Sci USA* 2001, 98:9724–9729.
12. Shao W, Zhao Q-Y, Wang X-Y, Xu X-Y, Tang Q, Li M, Li X, Xu Y-Z. Alternative splicing and trans-splicing events revealed by analysis of the Bombyx mori transcriptome. *RNA* 2012, 18:1395–1407.
13. Kamikawa R, Inagaki Y, Tokoro M, Roger AJ, Hashimoto T. Split introns in the genome of *Giardia intestinalis* are excised by spliceosome-mediated trans-splicing. *Curr Biol* 2011, 21:311–315.
14. Li H, Wang J, Mor G, Sklar J. A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. *Science* 2008, 321:1357–1361.
15. Roy SW, Hudson AJ, Joseph J, Yee J, Russell AG. Numerous fragmented spliceosomal introns, AT-AC splicing, and an unusual dynein gene expression pathway in Giardia lamblia. *Mol Biol Evol* 2012, 29:43–49.
16. McManus CJ, Duff MO, Eipper-Mains J, Graveley BR. Global analysis of trans-splicing in Drosophila. *Proc Natl Acad Sci USA* 2010, 107:12975–12979.
17. Anderson K, Moore MJ. Bimolecular exon ligation by the human spliceosome. *Science* 1997, 276:1712–1716.
18. Anderson K, Moore MJ. Bimolecular exon ligation by the human spliceosome bypasses early 3′ splice site AG recognition and requires NTP hydrolysis. *RNA* 2000, 6:16–25.

19. Lambowitz AM, Zimmerly S. Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb Perspect Biol* 2011, 3:a003616.
20. Koonin EV. The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate? *Biol Direct* 2006, 1:22.
21. Yao J, Zhong J, Fang Y, Geisinger E, Novick RP, Lambowitz AM. Use of targetrons to disrupt essential and nonessential genes in Staphylococcus aureus reveals temperature sensitivity of Ll.LtrB group II intron splicing. *RNA* 2006, 12:1271–1281.
22. Zhuang F, Mastroianni M, White TB, Lambowitz AM. Linear group II intron RNAs can retrohome in eukaryotes and may use nonhomologous end-joining for cDNA ligation. *Proc Natl Acad Sci USA* 2009, 106:18189–18194.
23. Chapman KB, Boeke JD. Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell* 1991, 65:483–492.
24. Nam K, Hudson RH, Chapman KB, Ganeshan K, Damha MJ, Boeke JD. Yeast lariat debranching enzyme. Substrate and sequence specificity. *J Biol Chem* 1994, 269:20613–20621.
25. Khalid MF, Damha MJ, Shuman S, Schwer B. Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. *Nucleic Acids Res* 2005, 33:6349–6360.
26. Nam K, Lee G, Trambley J, Devine SE, Boeke JD. Severe growth defect in a *Schizosaccharomyces pombe* mutant defective in intron lariat degradation. *Mol Cell Biol* 1997, 17:809–818.
27. Wang H, Hill K, Perry SE. An Arabidopsis RNA lariat debranching enzyme is essential for embryogenesis. *J Biol Chem* 2004, 279:1468–1473.
28. Kim JW, Kim HC, Kim GM, Yang JM, Boeke JD, Nam K. Human RNA lariat debranching enzyme cDNA complements the phenotypes of *Saccharomyces cerevisiae* dbr1 and *Schizosaccharomyces pombe* dbr1 mutants. *Nucleic Acids Res* 2000, 28: 3666–3673.
29. Cheng Z, Menees TM. RNA branching and debranching in the yeast retrovirus-like element Ty1. *Science* 2004, 303:240–243.
30. Karst SM, Rütz ML, Menees TM. The yeast retrotransposons Ty1 and Ty3 require the RNA Lariat debranching enzyme, Dbr1p, for efficient accumulation of reverse transcripts. *Biochem Biophys Res Commun* 2000, 268:112–117.
31. Ye Y, De Leon J, Yokoyama N, Naidu Y, Camerini D. DBR1 siRNA inhibition of HIV-1 replication. *Retrovirology* 2005, 2:63.
32. Coombes CE, Boeke JD. An evaluation of detection methods for large lariat RNAs. *RNA* 2005, 11:323–331.
33. Pratico ED, Silverman SK. Ty1 reverse transcriptase does not read through the proposed 2″,5-″branched retrotransposition intermediate in vitro. *RNA* 2007, 13:1528–1536.
34. Lauermann V, Nam K, Trambley J, Boeke JD. Plus-strand strong-stop DNA synthesis in retrotransposon Ty1. *J Virol* 1995, 69:7845–7850.
35. Armakola M, Higgins MJ, Figley MD, Barmada SJ, Scarborough EA, Diaz Z, Fang X, Shorter J, Krogan NJ, Finkbeiner S, et al. Inhibition of RNA lariat debranching enzyme suppresses TDP-43 toxicity in ALS disease models. *Nat Genet* 2012, 44:1302–1309.
36. Burgess SM, Guthrie C. A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. *Cell* 1993, 73:1377–1391.
37. Hilleren PJ, Parker R. Cytoplasmic degradation of splice-defective pre-mRNAs and intermediates. *Mol Cell* 2003, 12:1453–1465.
38. Harigaya Y, Parker R. Global analysis of mRNA decay intermediates in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 2012, 109:11764–11769.
39. Danin-Kreiselman M, Lee CY, Chanfreau G. RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. *Mol Cell* 2003, 11:1279–1289.
40. del Campo EM, Casano LM. Degradation of plastid unspliced transcripts and lariat group II introns. *Biochimie* 2008, 90:474–483.
41. LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* 2005, 121:713–724.
42. San Paolo S, Vanacova S, Schenk L, Scherrer T, Blank D, Keller W, Gerber AP. Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. *PLoS Genet* 2009, 5:e1000555.
43. Ruskin B, Green MR. An RNA processing activity that debranches RNA lariats. *Science* 1985, 229:135–140.
44. Tseng C-K, Cheng S-C. Both catalytic steps of nuclear pre-mRNA splicing are reversible. *Science* 2008, 320:1782–1784.
45. Mayas RM, Maita H, Semlow DR, Staley JP. Spliceosome discards intermediates via the DEAH box ATPase Prp43p. *Proc Natl Acad Sci USA* 2010, 107:10020–10025.
46. Qian L, Vu MN, Carter M, Wilkinson MF. A spliced intron accumulates as a lariat in the nucleus of T cells. *Nucleic Acids Res* 1992, 20:5345–5350.
47. Guil S, Soler M, Portela A, Carrère J, Fonalleras E, Gómez A, Villanueva A, Esteller M. Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nat Struct Mol Biol* 2012, 19:664–670.
48. Gardner EJ, Nizami ZF, Talbot CC, Gall JG. Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of *Xenopus tropicalis*. *Genes Dev* 2012, 26:2550–2559.

49. Perng GC, Jones C, Ciacci-Zanella J, Stone M, Henderson G, Yukht A, Slanina SM, Hofman FM, Ghiasi H, Nesburn AB, et al. Virus-induced neuronal apoptosis blocked by the herpes simplex virus latency-associated transcript. *Science* 2000, 287:1500–1503.
50. Nicosia M, Zabolotny JM, Lirette RP, Fraser NW. The HSV-1 2-kb latency-associated transcript is found in the cytoplasm comigrating with ribosomal subunits during productive infection. *Virology* 1994, 204:717–728.
51. Clement JQ, Qian L, Kaplinsky N, Wilkinson MF. The stability and fate of a spliced intron from vertebrate cells. *RNA* 1999, 5:206–220.
52. Clement JQ, Maiti S, Wilkinson MF. Localization and stability of introns spliced from the Pem homeobox gene. *J Biol Chem* 2001, 276:16919–16930.
53. Coleclough C, Wood D. Introns excised from immunoglobulin pre-mRNAs exist as discrete species. *Mol Cell Biol* 1984, 4:2017–2022.
54. Yoshimoto R, Kataoka N, Okawa K, Ohno M. Isolation and characterization of post-splicing lariat-intron complexes. *Nucleic Acids Res* 2009, 37:891–902.
55. Fourmann J-B, Schmitzová J, Christian H, Urlaub H, Ficner R, Boon K-L, Fabrizio P, Lührmann R. Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev* 2013, 27:413–428.
56. Martin A, Schneider S, Schwer B. Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem* 2002, 277:17743–17750.
57. Small EC, Leggett SR, Winans AA, Staley JP. The EF-G-like GTPase Snu114p Regulates Spliceosome Dynamics Mediated by Brr2p, a DExD/H Box ATPase. *Mol Cell* 2006, 23:389–399.
58. Kannan R, Hartnett S, Voelker RB, Berglund JA, Staley JP, Baumann P. Intronic sequence elements impede exon ligation and trigger a discard pathway that yields functional telomerase RNA in fission yeast. *Genes Dev* 2013, 27:627–638.
59. Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C. Rapid, transcript-specific changes in splicing in response to environmental stress. *Mol Cell* 2007, 27:928–937.
60. Albulescu L-O, Sabet N, Gudipati M, Stepankiw N, Bergman ZJ, Huffaker TC, Pleiss JA. A quantitative, high-throughput reverse genetic screen reveals novel connections between Pre-mRNA splicing and 5′′ and 3′′ end transcript determinants. *PLoS Genet* 2012, 8:e1002530.
61. Tseng C-K, Cheng S-C. The spliceosome catalyzes debranching in competition with reverse of the first chemical reaction. *RNA* 2013, 320:1782–1784.
62. Ooi SL, Samarsky DA, Fournier MJ, Boeke JD. Intronic snoRNA biosynthesis in *Saccharomyces cerevisiae* depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. *RNA* 1998, 4:1096–1110.
63. Tycowski KT, Shu MD, Steitz JA. A mammalian gene with introns instead of exons generating stable RNA products. *Nature* 1996, 379:464–466.
64. Yin Q-F, Yang L, Zhang Y, Xiang J-F, Wu Y-W, Carmichael GG, Chen L-L. Long noncoding RNAs with snoRNA ends. *Mol Cell* 2012, 48:219–230.
65. Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. *Nature* 2007, 448:83–86.
66. Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell* 2007, 130: 89–100.
67. Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* 2009, 11:228–234.
68. Flynt AS, Greimann JC, Chung WJ, Lima CD, Lai EC. MicroRNA biogenesis via splicing and exosome-mediated trimming in Drosophila. *Mol Cell* 2010, 38:900–907.
69. Ladewig E, Okamura K, Flynt AS, Westholm JO, Lai EC. Discovery of hundreds of mirtrons in mouse and human small RNA data. *Genome Res* 2012, 22:1634–1645.
70. Curtis HJ, Sibley CR, Wood MJA. Mirtrons, an emerging class of atypical miRNA. *WIREs RNA* 2012, 3:617–632.
71. Gu W, Lee H-C, Chaves D, Youngman EM, Pazour GJ, Conte D, Mello CC. CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. *Cell* 2012, 151:1488–1500.
72. Saito TL, Hashimoto S-I, Gu SG, Morton JJ, Stadler M, Blumenthal T, Fire A, Morishita S. The transcription start site landscape of *C. elegans*. *Genome Res* 2013.
73. Morton JJ, Blumenthal T. Identification of transcription start sites of trans-spliced genes: uncovering unusual operon arrangements. *RNA* 2011, 17:327–337.
74. Kent JR, Kang W, Miller CG, Fraser NW. Herpes simplex virus latency-associated transcript gene function. *J Neurovirol* 2003, 9:285–290.
75. Wu TT, Su YH, Block TM, Taylor JM. Evidence that two latency-associated transcripts of herpes simplex virus type 1 are nonlinear. *J Virol* 1996, 70:5962–5967.
76. Rødahl E, Haarr L. Analysis of the 2-kilobase latency-associated transcript expressed in PC12 cells productively infected with herpes simplex virus type 1:

evidence for a stable, nonlinear structure. *J Virol* 1997, 71:1703–1707.

77. Kulesza CA, Shenk T. Murine cytomegalovirus encodes a stable intron that facilitates persistent replication in the mouse. *Proc Natl Acad Sci USA* 2006, 103:18302–18307.
78. Inman M, Perng GC, Henderson G, Ghiasi H, Nesburn AB, Wechsler SL, Jones C. Region of herpes simplex virus type 1 latency-associated transcript sufficient for wild-type spontaneous reactivation promotes cell survival in tissue culture. *J Virol* 2001, 75:3636–3646.
79. Goldenberg D, Mador N, Ball MJ, Panet A, Steiner I. The abundant latency-associated transcripts of herpes simplex virus type 1 are bound to polyribosomes in cultured neuronal cells and during latent infection in mouse trigeminal ganglia. *J Virol* 1997, 71:2897–2904.
80. Knipe DM, Lieberman PM, Jung JU, McBride AA, Morris KV, Ott M, Margolis D, Nieto A, Nevels M, Parks RJ, et al. Snapshots: chromatin control of viral infection. *Virology* 2013, 435:141–156.
81. Cliffe AR, Garber DA, Knipe DM. Transcription of the herpes simplex virus latency-associated transcript promotes the formation of facultative heterochromatin on lytic promoters. *J Virol* 2009, 83:8182–8190.
82. Kwiatkowski DL, Thompson HW, Bloom DC. The polycomb group protein Bmi1 binds to the herpes simplex virus 1 latent genome and maintains repressive histone marks during latency. *J Virol* 2009, 83:8173–8181.
83. Tsai M-C, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 2010, 329:689–693.
84. Krummenacher C, Zabolotny JM, Fraser NW. Selection of a nonconsensus branch point is influenced by an RNA stem-loop structure and is important to confer stability to the herpes simplex virus 2-kilobase latency-associated transcript. *J Virol* 1997, 71:5849–5860.
85. Jacquier A, Rosbash M. RNA splicing and intron turnover are greatly diminished by a mutant yeast branch point. *Proc Natl Acad Sci USA* 1986, 83:5835–5839.
86. Reineke LC, Lloyd RE. Animal virus schemes for translation dominance. *Curr Opin Virol* 2011, 1:363–372.
87. Roy SW, Gilbert W. The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat Rev Genet* 2006, 7:211–221.
88. Sharp PA. On the origin of RNA splicing and introns. *Cell* 1985, 42:397–400.
89. Tani T, Ohshima Y. mRNA-type introns in U6 small nuclear RNA genes: implications for the catalysis in pre-mRNA splicing. *Genes Dev* 1991, 5: 1022–1031.
90. Coghlan A, Wolfe KH. Origins of recently gained introns in Caenorhabditis. *Proc Natl Acad Sci USA* 2004, 101:11362–11367.
91. Schwer B. A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol Cell* 2008, 30:743–754.
92. Schwer B, Meszaros T. RNA helicase dynamics in pre-mRNA splicing. *EMBO J* 2000, 19:6582–6591.
93. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell* 2009, 136:701–718.
94. Huelga SC, Vu AQ, Arnold JD, Liang TY, Liu PP, Yan BY, Donohue JP, Shiue L, Hoon S, Brenner S, et al. Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep* 2012, 1:167–178.
95. Wang Y, Ma M, Xiao X, Wang Z. Intronic splicing enhancers, cognate splicing factors and context-dependent regulation rules. *Nat Struct Mol Biol* 2012, 19:1044–1052.
96. Brent MR. Steady progress and recent breakthroughs in the accuracy of automated genome annotation. *Nat Rev Genet* 2008, 9:62–73.
97. Zhang Z, Hesselberth JR, Fields S. Genome-wide identification of spliced introns using a tiling microarray. *Genome Res* 2007, 17:503–509.
98. Juneau K, Palm C, Miranda M, Davis RW. High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing. *Proc Natl Acad Sci USA* 2007, 104:1522–1527.
99. Clement JQ, Wilkinson MF. Rapid induction of nuclear transcripts and inhibition of intron decay in response to the polymerase II inhibitor DRB. *J Mol Biol* 2000, 299:1179–1191.
100. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 2009, 460:479–486.
101. Tsai R-T, Fu R-H, Yeh F-L, Tseng C-K, Lin Y-C, Huang Y-H, Cheng S-C. Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. *Genes Dev* 2005, 19:2991–3003.
102. Freeberg MA, Han T, Moresco JJ, Kong A, Yang Y-C, Lu ZJ, Yates JR, Kim JK. Pervasive and dynamic protein binding sites of the mRNA transcriptome in *Saccharomyces cerevisiae*. *Genome Biol* 2013, 14:R13.
103. Windhager L, Bonfert T, Burger K, Ruzsics Z, Krebs S, Kaufmann S, Malterer G, L'Hernault A, Schilhabel M, Schreiber S, et al. Ultrashort and progressive 4sU-tagging reveals key characteristics of RNA processing at nucleotide resolution. *Genome Res* 2012, 22:2031–2042.
104. Taggart AJ, DeSimone AM, Shih JS, Filloux ME, Fairbrother WG. Large-scale mapping of branchpoints

in human pre-mRNA transcripts *in vivo*. *Nat Struct Mol Biol* 2012, 19:719–721.

105. Suzuki H, Zuo Y, Wang J, Zhang MQ, Malhotra A, Mayeda A. Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res* 2006, 34:e63.
106. Wang Y, Silverman SK. Efficient one-step synthesis of biologically related lariat RNAs by a deoxyribozyme. *Angew Chem Int Ed* 2005, 44:5863–5866.
107. Mui TP, Silverman SK. Convergent and general one-step DNA-catalyzed synthesis of multiply branched DNA. *Org Lett* 2008, 10:4417–4420.
108. Lee CS, Mui TP, Silverman SK. Improved deoxyribozymes for synthesis of covalently branched DNA and RNA. *Nucleic Acids Res* 2010, 39: 269–279.
109. Carriero S, Damha MJ. Solid-phase synthesis of branched oligonucleotides. In: Beaucage S, ed. *Current protocols in nucleic acid chemistry*. Hoboken, NJ: John Wiley & Sons; 2002, Unit 4.14: 1–32.
110. Mourani R, Damha MJ. Synthesis, characterization, and biological properties of small branched RNA fragments containing chiral (Rp and Sp) 2″,5-″ phosphorothioate linkages. *Nucleosides Nucleotides Nucleic Acids* 2006, 25:203–229.

