Trans-splicing versatility of the Ll.LtrB group II intron

KAMILA BELHOCINE, ANTHONY B. MAK, and BENOIT COUSINEAU
Department of Microbiology and Immunology, McGill University, Montréal, Québec H3A 2B4, Canada

ABSTRACT

Group II introns are found in organelles, bacteria, and archaea. Some harbor an open reading frame (ORF) with reverse transcriptase, maturase, and occasionally endonuclease activities. Group II introns require the assistance of either intron-encoded or free-standing maturases to excise from primary RNA transcripts in vivo. Some ORF-containing group II introns were shown to be mobile retroelements that invade new DNA sites by retrohoming or retrotransposition. Group II introns are also hypothesized to be the ancestors of the spliceosome-dependent nuclear introns and the small nuclear RNAs (snRNAs—U1, U2, U4, U5, and U6) that are part of the spliceosome. The ability of some fragmented group II introns to undergo splicing in *trans* supports the theory that the snRNAs evolved from portions of group II introns. Here, we developed a Tn5-based genetic screen to explore the *trans*-splicing potential of the Ll.LtrB group II intron from the Gram-positive bacterium *Lactococcus lactis*. Proficient *trans*-splicing variants of Ll.LtrB were selected using a highly sensitive *trans*-splicing/conjugation screen. We report that numerous fragmentation sites located throughout Ll.LtrB support splicing in *trans*, showing that this intron is remarkably more tolerant to fragmentation than expected from the fragmentation sites uncovered within natural *trans*-splicing group II introns. This work unveils the great versatility of group II intron fragments to assemble and accurately *trans*-splice their flanking exons in vivo.

Keywords: *Lactococcus lactis*; conjugation; spliceosome; snRNAs; Tn5

INTRODUCTION

Group II introns were originally discovered interrupting mitochondrial genes of lower eukaryotes, as well as mitochondrial and chloroplastic genes of higher plants. In contrast, no group II introns were found interrupting nuclear genes of eukaryotes. Bacterial group II introns were uncovered more recently and are present in one to a few copies per genome in both Gram-negative and Gram-positive bacteria, while in archaea, group II introns are infrequent (Michel and Ferat 1995; Lambowitz and Zimmerly 2004). Interestingly, some group II introns that contain an open reading frame (ORF) are also mobile retroelements. Following their excision, they invade new DNA sites using an RNA intermediate and target either identical or similar sequences by retrohoming or retrotransposition, respectively (Lambowitz and Zimmerly 2004). From an evolutionary point of view, group II introns are considered to be the ancestors of the spliceosome-dependent nuclear introns and the small nuclear RNAs (snRNAs—U1, U2, U4, U5, U6) that are part of the eukaryotic spliceosome machinery (Sharp 1985, 1991; Cech 1986; Jarrell et al. 1988; Jacquier 1990; Cavalier-Smith 1991; Michel and Ferat 1995).

Group II introns are large RNA enzymes that need the assistance of maturases to remove themselves from pre-mRNAs in vivo. These ribozymes self-splice from RNA transcripts through two consecutive transesterification reactions (Fig. 1A; Michel and Ferat 1995; Lambowitz and Zimmerly 2004). Following transcription of the interrupted gene (Fig. 1A, step 1), the first nucleophilic attack at the exon 1–intron junction is initiated by the 2′-OH residue of a bulged adenosine (A) located near the 3′ end of the intron (Fig. 1A, step 2). The liberated 3′-OH of exon 1 then initiates the second transesterification reaction at the intron–exon 2 junction (Fig. 1A, step 3), which ligates the two flanking exons and releases the intron RNA in the form of a lariat. Eukaryotic nuclear introns, which require the spliceosome machinery to be excised, follow the same splicing pathway and are also released as lariats. This strongly suggests that group II and nuclear introns share a common ancestor (Sharp 1985; Cech 1986). The similarity of exon–intron boundaries between group II and nuclear introns also supports this hypothesis (Jacquier 1990). Moreover, substantial

Reprint requests to: Benoit Cousineau, Department of Microbiology and Immunology, McGill University, Lyman Duff Medical Building, 3775 University Street, Montréal, Québec H3A 2B4, Canada; e-mail: benoit.cousineau@mcgill.ca; fax: (514) 398-7052.

Article published online ahead of print. Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.1083508.

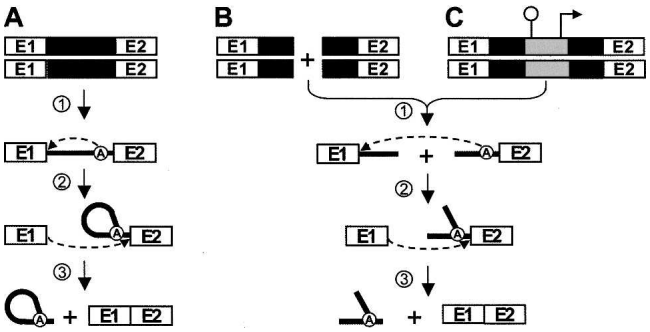

FIGURE 1. Group II intron splicing pathways. (*A*) *Cis*-splicing. Following transcription of the interrupted gene (step *1*), the 2′-OH residue of the branch point nucleotide (A) performs a nucleophilic attack at the exon 1–intron junction (step *2*). Then, the liberated 3′-OH at the end of exon 1 performs the second nucleophilic attack at the intron–exon 2 junction (step *3*), ligating the two exons and releasing the intron in the form of a lariat. (*B*) *Trans*-splicing. Some fragmented group II introns that are expressed from separate loci (step *1*) can splice in *trans* to ligate their flanking exons and release a Y-branched intron molecule (steps *2* and *3*). (*C*) *Trans*-splicing of Tn5-interrupted group II introns. Insertion of a Tn5 transposon, harboring a transcriptional terminator and a constitutive promoter, within the intron results in the expression of the gene as two separate transcripts (step *1*). These two transcripts may undergo *trans*-splicing and ligate their flanking exons (steps *2* and *3*). Group II intron, black box (DNA) and black line (RNA); branch point, circled A; exon 1 and 2, E1 and E2; Tn5 transposon, gray box; *pepN* transcriptional terminator, schematic stem–loop; P_{23} constitutive promoter, right-angle arrow.

evidence indicates that the splicing reaction achieved by the spliceosome is fundamentally RNA-mediated, strengthening the argument that this large splicing machinery evolved from a primordial ribozyme (for review, see Valadkhan 2007). Finally, the discovery that some organellar group II introns fragmented into two or three pieces can splice in *trans* (Fig. 1B; for review, see Bonen 1993) further suggests that the snRNAs of the spliceosome were generated from group II intron fragments (Sharp 1991).

Despite a lack of similitude at the primary sequence level, the secondary structure of group II introns is well conserved (Fedorova and Zingler 2007). These RNA elements fold into six domains (DI–DVI) that radiate from a central wheel (Fig. 2A). Domain I (DI), the largest functional domain, serves as a scaffold to dock the remaining domains (DII–DVI). The lower stem of DII is believed to recruit DIII to the catalytic core, with DIII acting as a catalytic effector (Fedorova and Zingler 2007). DIV is dispensable for splicing (Jarrell et al. 1988) and highly variable in size since it may contain an ORF. Accordingly, upon folding of the intron, this domain is thought to protrude from the catalytic core. DV is the catalytic domain and exhibits the most sequence conservation among group II introns; its interaction with DI forms the minimal catalytic core (Fedorova and Zingler 2007). DVI carries the bulged adenosine (A) residue, also termed the branch point, responsible for the first nucleophilic attack of the splicing pathway (Fig. 1; Fedorova and Zingler 2007). In addition to local interactions, a series of long-range tertiary contacts between the various domains contribute to the three-dimensional folding and compaction of group II introns, and must take place to enable splicing (Fedorova and Zingler 2007).

Functional and/or structural similarities exist between portions of group II introns and snRNAs (Fig. 2). For instance, DI is responsible for selection of the splice sites during group II intron splicing. The 5′ end of the intron is recognized through Watson–Crick base pairs (Fig. 2A, ε–ε′; Jacquier 1990; Fedorova and Zingler 2007). This is similar to the recognition of the 5′ end of spliceosomal introns by U1 (Fig. 2B; Jacquier 1990; Nilsen 1994; Patel and Steitz 2003). The 5′ splice site is also identified through a series of base pair interactions between intron binding sites (IBS) 1 and 2, which are at the 3′ end of exon 1, and exon binding sites (EBS) 1 and 2 in DI (Fig. 2A). The 3′ splice site is recognized by either the δ–δ′ (Fig. 2A) or EBS3–IBS3 base pair interactions between DI and exon 2 (Michel and Ferat 1995; Fedorova and Zingler 2007). The DI subdomain that harbors the EBS and δ sequences is thus functionally analogous to U5, which also recognizes the extremities of both exons at the intron–exon junctions in spliceosome-mediated splicing (Fig. 2B; Newman and Norman 1992; Nilsen 1994; Patel and Steitz 2003). In accordance, it was shown that a slightly modified U5 can complement the ID3 domain in group II intron catalysis (Hetzer et al. 1997). On the other hand, NMR studies showed that the stem–loop region of the catalytic snRNA U6, responsible for coordinating the metal ion in the spliceosome catalytic core, is structurally similar to DV, which plays the same role in group II intron splicing (Seetharaman et al. 2006). A slightly modified DV was shown to functionally complement

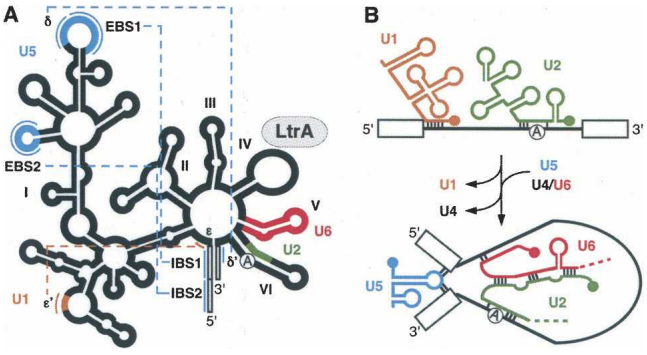

FIGURE 2. Functional and/or structural equivalences between portions of group II introns and the snRNAs of the spliceosome. (*A*) Schematic of group II intron conserved secondary structure. The six domains (I–VI) and pertinent long-range tertiary interactions are represented (dotted lines). (*B*) Interactions between snRNAs and spliceosomal introns. The functional and/or structural correspondences between snRNAs and portions of group II introns are color-coded. Group II intron, thick black line; spliceosomal intron, thin black line; branch point, circled A; exons, open boxes; exon binding sites 1 and 2; EBS1 and 2; intron binding sites 1 and 2, IBS1 and 2.

U6 in the U12-dependant spliceosome, which further strengthens the putative evolutionary link between these two RNA elements (Shukla and Padgett 2002). Finally, binding of the U2 snRNA to the distal end of nuclear introns causes the branch point nucleotide to bulge out in a structure similar to DVI of group II introns (Fig. 2; Jacquier 1990). These functional and/or structural similarities between portions of group II intron and the snRNAs further support the evolutionary hypothesis that snRNAs were derived from group II intron fragments.

The Ll.LtrB intron from the Gram-positive bacterium *Lactococcus lactis* is one of the most characterized bacterial group II introns (Lambowitz and Zimmerly 2004). This 2.5-kb intron carries an ORF coding for LtrA, a 599 amino acid protein with reverse transcriptase, maturase, and endonuclease functions (Matsuura et al. 1997). LtrA is an essential splicing co-factor for Ll.LtrB. The maturase activity of LtrA promotes intron splicing by stabilizing the catalytically active tertiary conformation of the ribozyme (Matsuura et al. 1997). Ll.LtrB was found on two highly similar conjugative elements of *L. lactis*: a conjugative plasmid, pRS01 (Mills et al. 1996), and a chromosomal sex factor (Shearman et al. 1996). In both elements, Ll.LtrB interrupts the *ltrB* gene coding for relaxase (LtrB) at the same position. Relaxase initiates conjugation by nicking the conjugative element at the origin of transfer (*oriT*) and delivers the DNA to recipient cells (Byrd and Matson 1997). This enzyme is thus essential for conjugation, and splicing of Ll.LtrB from its pre-mRNA transcript is necessary for LtrB production and subsequent DNA transfer. Therefore, Ll.LtrB splicing controls conjugation efficiency of its host elements (Mills et al. 1996; Shearman et al. 1996). We previously developed a sensitive splicing/conjugation assay in *L. lactis* to show that Ll.LtrB *trans*-splices when fragmented at natural group II intron fragmentation sites. We also demonstrated that, similarly to *cis*-splicing, LtrA is essential for Ll.LtrB *trans*-splicing in vivo (Belhocine et al. 2007).

Here, we engineered a Tn5-based genetic screen to explore what fragmentation sites throughout Ll.LtrB allow *trans*-splicing in vivo. *Trans*-splicing Ll.LtrB variants were selected using a *trans*-splicing/conjugation assay. Our data reveal that Ll.LtrB can *trans*-splice efficiently when fragmented at various locations throughout its structure. Therefore, this study demonstrates that Ll.LtrB is more versatile in *trans*-splicing than expected from natural *trans*-splicing group II introns.

RESULTS

A pLE12-based conjugation assay to monitor Ll.LtrB splicing

We previously developed a sensitive splicing/conjugation assay in *L. lactis* (10^7-fold detection range) where conjugation of a relaxase-deficient chromosomal sex factor was used as a read-out for Ll.LtrB splicing. In that assay, the relaxase gene, interrupted by different *cis*- and *trans*-splicing Ll.LtrB variants, was provided from a non-mobilizable plasmid (Belhocine et al. 2007).

A similar splicing/conjugation assay was used here; however, Ll.LtrB splicing is monitored by the conjugation efficiency of its carrying mobilizable plasmid (pLE12; Belhocine et al. 2004). pLE12 harbors a portion of the pRS01 conjugative plasmid including the origin of transfer (*oriT*) and the *ltrB* relaxase gene, interrupted by Ll.LtrB (Fig. 3A, PstI fragment; Belhocine et al. 2004). Recognition of the *oriT* by the relaxase initiates pLE12 mobilization, which then allows its transfer by the chromosomally encoded sex factor conjugation machinery (Belhocine et al. 2004). The *L. lactis* strain NZ9800Δ*ltrB*::*tet* bears a defective relaxase gene and is used as the donor strain for pLE12 conjugation. In this strain, the Ll.LtrB intron and parts of its flanking exons were deleted from the chromosomal sex factor (Belhocine et al. 2004). The absence of relaxase in this strain prevents sex factor conjugation, as shown by the significant reduction in transfer efficiency from 10^{-3} (NZ9800) to 10^{-9} (NZ9800Δ*ltrB*) (Belhocine et al. 2007). However, since pLE12 encodes its own relaxase, it is efficiently transferred from NZ9800Δ*ltrB* to the LM0231 recipient strain (Table 1, pLE12, 3.17×10^{-5}). To assess the dependence of pLE12 transfer on Ll.LtrB splicing, two plasmid variants carrying splicing defective introns were tested for conjugation. Either when pLE12 bears a catalytic mutant of Ll.LtrB (pLE12ΔDV) or when most of the LtrA coding region is deleted (pLE12ΔORF), the transfer efficiency drops significantly ($\sim 10^4$-fold) to background level (10^{-9}) (Table 1). These data show that conjugation of pLE12 from NZ9800Δ*ltrB* to LM0231 is initiated by the plasmid-borne relaxase, and that its conjugative transfer depends on splicing of Ll.LtrB from the *ltrB* transcript.

Tn5-based genetic screen to study Ll.LtrB *trans*-splicing versatility

We previously demonstrated that Ll.LtrB splices in *trans* when fragmented independently within DI, DIII, or DIV, either upstream or downstream from *ltrA* (Fig. 4, sites labeled S1–S5; Belhocine et al. 2007). Interestingly, all fragmented Ll.LtrB variants, which were engineered to mimic natural *trans*-splicing introns, were able to *trans*-splice (Belhocine et al. 2007). We therefore decided to develop a more comprehensive approach to investigate the full *trans*-splicing potential of group II introns.

We engineered a Tn5 transposon carrying the transcriptional terminator from *pepN* followed by the P_{23} *L. lactis* constitutive promoter (Fig. 3B). Insertion of this transposon into a target gene severs the RNA transcript into two independent fragments: one fragment extending from the natural promoter of the gene up to the Tn5 insertion site (transcriptional terminator) and the second fragment

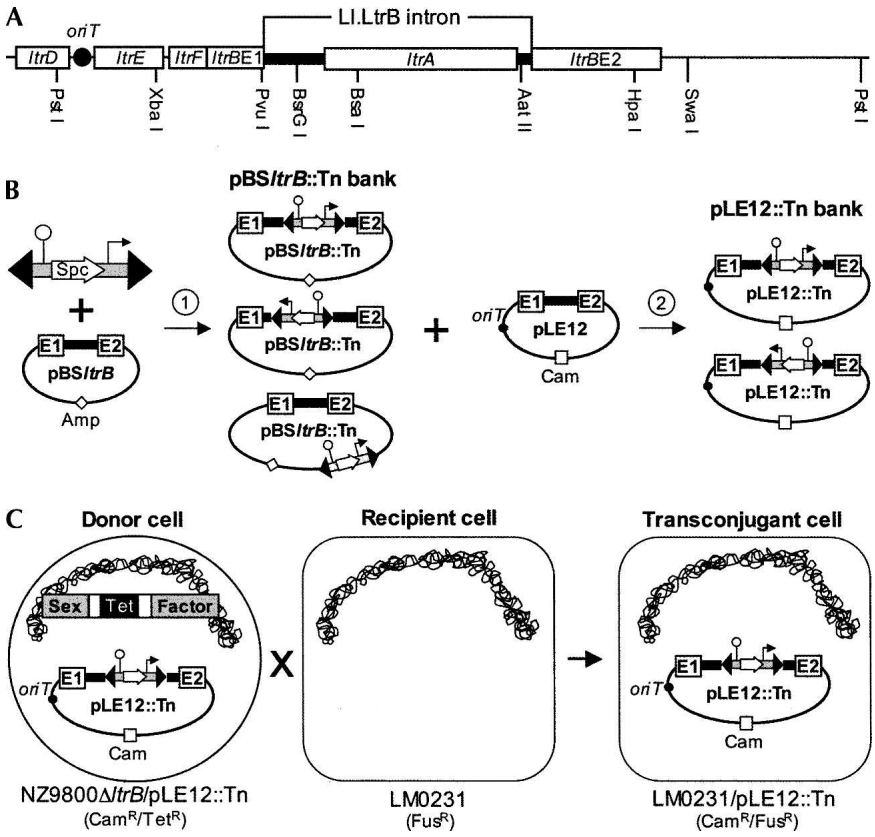

FIGURE 3. *Trans*-splicing/conjugation assay. (*A*) Schematic of the pRS01 fragment subcloned within pLE12 (PstI). (*B*) Generation of saturated Tn5 insertion banks. The Tn5 transposon was mixed with the target plasmid (pBS*ltrB*) in presence of the Tn5 transposase (step *1*). This in vitro reaction created a bank of Tn5 insertions within the target plasmid (pBS*ltrB*::Tn) consisting of unique insertions between every nucleotide. To generate various pLE12::Tn banks bearing Tn5 insertions in specific sections of Ll.LtrB, the corresponding sections were excised from the pBS*ltrB*::Tn bank and transferred to pLE12 (step *2*). (*C*) *Trans*-splicing/conjugation assay. The pLE12::Tn bank is introduced in the NZ9800Δ*ltrB* donor strain and mated with the LM0231 recipient strain. This assay selects for plasmids harboring proficient Ll.LtrB *trans*-splicing variants. Every donor cell contains a different plasmid, only the pLE12::Tn variants that harbor a *trans*-splicing proficient Ll.LtrB intron produce relaxase and are transferred by conjugation. Origin of conjugative transfer (*oriT*), black circle.

extending from the Tn5 insertion site (P_{23} promoter) to the end of the transcript (Fig. 1C, step 1). We first used this randomly inserting transposon (Reznikoff et al. 2004) to generate a saturated bank of transposon-containing pBS*ltrB* plasmids (pBS*ltrB*::Tn) (Fig. 3B, step 1). Next, the section of pBS*ltrB*::Tn containing Ll.LtrB (Fig. 3A, PvuI/HpaI fragment) was transferred by subcloning to pLE12 (pLE12::Tn) (Fig. 3B, step 2) so that only the intron and portions of exon 1 and 2 were saturated with Tn5 insertions. Statistically, the pLE12::Tn bank contains independent plasmids, each carrying a single Tn5 insertion between each nucleotide of Ll.LtrB.

The pLE12::Tn bank was transformed into *L. lactis* NZ9800Δ*ltrB*, the transformants were pooled and used as donor cells for conjugation (Fig. 3C). Since relaxase production is needed to initiate pLE12 transfer, a pLE12::Tn variant can only be transferred if the *ltrB* transcript tolerates fragmentation at the Tn5 insertion site. More specifically, the plasmid will only be transferred if it carries a Tn5-fragmented Ll.LtrB that can *trans*-splice and ligate its flanking exons, allowing for relaxase production. Plasmids were recovered from 66 independent transconjugant cells, i.e., recipient cells that acquired a pLE12::Tn variant, and the transposon insertion sites were localized by sequence analysis. As seen in Figure 4A (black arrowheads), the Tn5 insertions are not randomly distributed within Ll.LtrB following the screen. Instead, the great majority of insertions (56/66) are found at the 5′ end of DI (42/66) and within DIV (14/66) either upstream of the LtrA start codon or within the last 300 nucleotides (nt) of *ltrA* downstream from the maturase domain. Interestingly, these three locations correspond to fragmentation sites observed in *trans*-splicing group II introns found in organelles (Bonen 1993). The rest of the Tn5 insertion sites (10/66) are novel fragmentation sites within DI and DII, which were never previously observed in characterized *trans*-splicing group II introns. No Tn5 insertions were isolated within DIII, DV, or DVI.

The presence of multiple Tn5 insertions within the endonuclease domain of *ltrA*, downstream from the maturase domain, suggests that these *trans*-splicing introns express truncated versions of active maturases. We repeated the conjugation screen in the presence of LtrA expressed from a second plasmid (pDE*ltrA*). Interestingly, when the pressure to produce an active maturase from the intron was relieved, numerous Tn5 insertions were found randomly distributed within *ltrA*, including insertions upstream and within the maturase domain (Fig. 4A, open arrowheads).

These genetic screens demonstrate that the 5′ end of DI and DIV correspond to the best fragmentation regions to create proficient Ll.LtrB *trans*-splicers. In addition, these experiments confirm the importance of the maturase function of LtrA for Ll.LtrB *trans*-splicing in *L. lactis*.

Trans-splicing efficiency of some Tn5 fragmented Ll.LtrB variants

The transfer rate of conjugative elements between *L. lactis* strains was shown to be directly proportional to the Ll.LtrB

TABLE 1. Conjugation efficiencies of pLE12 variants

Construct/ isolate	Conjugation efficiency	Conjugation efficiency with LtrA
pLE12	$(3.17 \pm 0.79) \times 10^{-5}$	$(2.06 \pm 0.90) \times 10^{-4}$
pLE12ΔD5	$(8.37 \pm 2.73) \times 10^{-9}$	
pLE12ΔORF	$(5.40 \pm 3.98) \times 10^{-9}$	$(4.98 \pm 1.98) \times 10^{-4}$
F1	$(2.62 \pm 1.19) \times 10^{-5}$	$(2.89 \pm 0.74) \times 10^{-5}$
F2	$(1.94 \pm 0.21) \times 10^{-5}$	$(4.78 \pm 6.42) \times 10^{-5}$
F3	$(3.16 \pm 0.54) \times 10^{-5}$	$(9.47 \pm 7.54) \times 10^{-5}$
F4	$(4.29 \pm 1.22) \times 10^{-6}$	$(2.14 \pm 0.99) \times 10^{-5}$
F5	$(1.80 \pm 0.57) \times 10^{-6}$	$(2.57 \pm 1.48) \times 10^{-5}$
F6	$(2.94 \pm 0.70) \times 10^{-6}$	$(1.13 \pm 0.42) \times 10^{-5}$
F7	$(1.86 \pm 0.29) \times 10^{-5}$	$(1.37 \pm 0.54) \times 10^{-4}$
F8	$(1.16 \pm 0.27) \times 10^{-5}$	$(2.07 \pm 0.34) \times 10^{-5}$
F9	$(3.32 \pm 0.84) \times 10^{-6}$	$(7.49 \pm 2.29) \times 10^{-6}$
F10	$(3.18 \pm 0.74) \times 10^{-6}$	$(9.51 \pm 2.50) \times 10^{-6}$
F11	$(1.89 \pm 0.58) \times 10^{-6}$	$(5.88 \pm 1.95) \times 10^{-6}$
$F1P^-$	$(1.37 \pm 0.98) \times 10^{-7}$	$(4.50 \pm 2.83) \times 10^{-7}$
$F3P^-$	$(3.17 \pm 0.36) \times 10^{-7}$	$(5.75 \pm 4.20) \times 10^{-7}$

splicing efficiency from the *ltrB* transcript (Belhocine et al. 2007). We thus assessed the *trans*-splicing efficiency of some selected Tn5-fragmented Ll.LtrB introns by conjugation of their pLE12 carriers (Fig. 4A, F1–F6). The six isolates tested showed significantly higher conjugation efficiencies (10^{-5}–10^{-6}) compared to the splicing-deficient introns (10^{-9}); some of these isolates (Fig. 4A, F1–F3) supported conjugation at a similar rate as pLE12 harboring the *cis*-splicing intron (Table 1). Moreover, detection of ligated exons by RT-PCR confirmed that both Ll.LtrB fragments are able to fold correctly, assemble, and *trans*-splice in order to ligate their flanking exons (Fig. 5, F1–F6). As expected from our functional analyses, sequencing of the RT-PCR products (Fig. 5, F1–F6) confirmed that removal of Ll.LtrB by *trans*-splicing is accurate, precisely joining the flanking exons.

Ll.LtrB *trans*-splicing variants fragmented both upstream of (Fig. 4A, F1–F4) and downstream (Fig. 4A, F5,F6) from *ltrA* were selected. We assumed that insertion of a constitutive promoter upstream of the ORF leads to higher expression levels of LtrA than when expressed from its native promoter. Indeed, Western blot analyses confirmed that significantly more LtrA is present in cells that contain an Ll.LtrB intron bearing a Tn5 insertion upstream of *ltrA* (data not shown). Nevertheless, when LtrA was provided in *trans*, all fragmented introns *trans*-spliced more efficiently (Table 1, 1.8- to 14.2-fold). The increase in *trans*-splicing efficiency parallels the increase of Ll.LtrB *cis*-splicing from pLE12 in the same conditions (Table 1, 6.5-fold). These results suggest that, when expressed from its natural promoter, the level of LtrA is suboptimal for both *cis*- and *trans*-splicing of Ll.LtrB.

As a control, we deleted the P_{23} promoter from the transposon of two selected Ll.LtrB *trans*-splicers (Fig. 4A, F1,F3). This resulted in a significant reduction of their respective conjugation efficiency (Table 1, ~100-fold decrease) and a drastic reduction of the RT-PCR signal corresponding to ligated exons (Fig. 5, $F1P^-$,$F3P^-$). These results demonstrate that expression of the second intron fragment is required to produce active relaxase. However, since the conjugation efficiencies of both controls are slightly over background (Table 1, cf. $F1P^-$ and $F3P^-$ and pLE12ΔD5 and pLE12ΔORF) and that a faint RT-PCR signal can be detected for $F1P^-$ (Fig. 5), we cannot completely rule out the possibility that some Tn5-interrupted Ll.LtrB variants splice in *cis* following a transcriptional readthrough of the terminator present within the transposon.

Screening Ll.LtrB subdomains for functional fragmentation sites

The full-length intron screen yielded a majority of fragmentation sites in a section of DI and within DIV (Fig. 4A). To facilitate the identification of novel fragmentation sites outside these regions, subdomains of the intron were subjected to the Tn5 screen (Fig. 4B; see brackets, I–III,II–III,V–VI). We first subjected domains I to III of Ll.LtrB to our screen, excluding most of *ltrA* (Fig. 3A, PvuI/BsaI fragment; Fig. 4B, I–III brackets). Following conjugation, the transposon insertion sites from 68 transferred plasmids were identified (Fig. 4B, open arrowheads). As we found in the original full-length intron screen (Fig. 4A), almost two thirds of the Tn5 insertions (41/68) were located in DI within a 120-nt section near the 5′ end of the intron. Moreover, numerous fragmentation sites that were selected within DI in the original full-length screen were isolated again (Fig. 4, cf. A and B). Nevertheless, the strategy of excluding most of DIV from the screen allowed the identification of more fragmentation sites within DII and two novel sites within DIII. These results confirm that the region located between the basal stem of DI and the ζ region constitutes a region of predilection for Ll.LtrB fragmentation.

We also screened the DII–III region of Ll.LtrB, excluding most of *ltrA* as well as the beginning of DI (Fig. 3A, fragment BsrGI/BsaI; Fig. 4B, II–III brackets). The transposon insertion sites from 41 transconjugants were sequenced (Fig. 4B, black arrowheads). Twelve transposon insertions were found within DI, concentrated in a 30-nt stretch located upstream of the ζ acceptor site. The majority of fragmentations (15/41) occurred within DII while seven insertions were found in DIII. Finally, four isolates harbored the Tn5 transposon within DIV, upstream of *ltrA*. These results suggest that DII is more flexible and tolerant to fragmentation than the 3′ portion of DI and DIII.

Domains V and VI are the most conserved domains among group II introns and both play key roles in splicing (Michel and Ferat 1995). DV is the catalytic domain while DVI contains the branch point nucleotide (A) that initiates

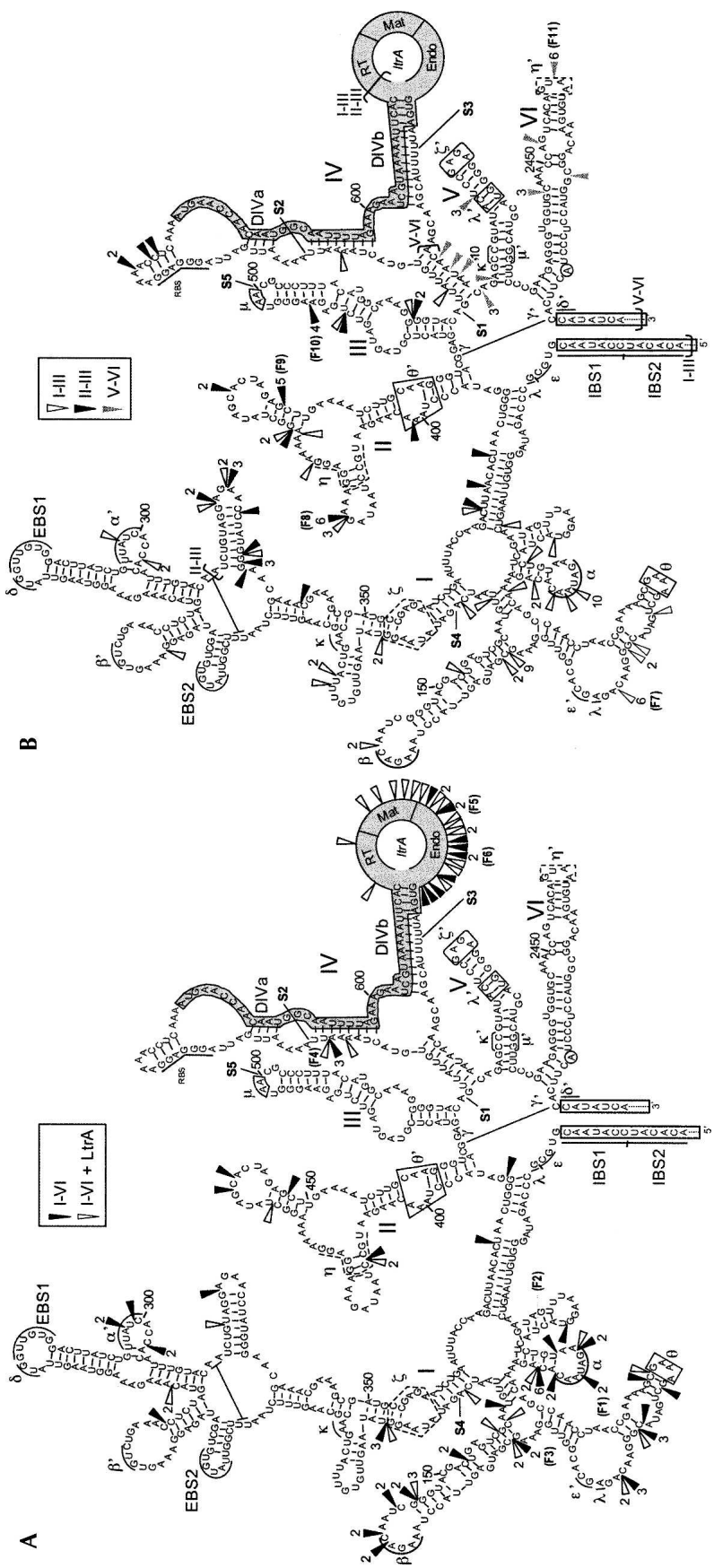

FIGURE 4. Selected Tn5 insertion sites supporting Ll.LtrB *trans*-splicing. (*A*) Ll.LtrB full-length screen (Fig. 3A, PvuI/HpaI). The screen was performed with (open arrowheads) or without (black arrowheads) the expression of LtrA in *trans* (pDE*ltrA*). (*B*) Screens of Ll.LtrB subdomains. The various sections of Ll.LtrB that were independently screened are delimited by brackets (I–III, open arrowheads; II–III, black arrowheads; V–VI, gray arrowheads). Isolates found multiple times (numbers) and the ones further analyzed are indicated (F1 to F11). The fragmentation points that were previously engineered to mimic natural *trans*-splicing group II introns are denoted (S1–S5, Belhocine et al. 2007). Exon binding sites 1 and 2, EBS1 and 2; intron binding sites 1 and 2, IBS1 and 2; exons, open boxes; *ltrA*, gray box; LtrA reverse transcriptase domain, RT; LtrA maturase domain, Mat; LtrA endonuclease domain, Endo; tertiary interactions, pairs of Greek letters.

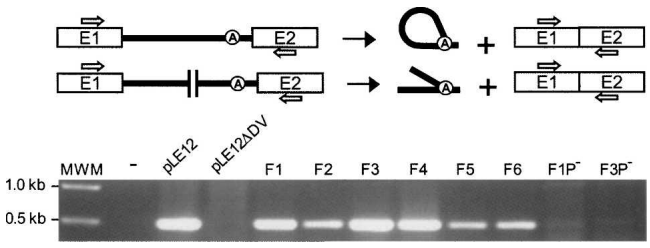

FIGURE 5. RT-PCR analysis of *ltrB* ligated exons. The primers used to amplify ligated exons are represented as arrows on the Ll.LtrB splicing schematics. The agarose gel shows the RT-PCR amplification products across the junction of *ltrB* ligated exons. Amplifications were performed on total RNA extracts of NZ9800Δ*ltrB*-containing plasmids harboring either continuous (pLE12, pLE12ΔD5) or fragmented introns (F1–F6). Variants of the F1 and F6 constructs missing the P_{23} promoter were also analyzed ($F1P^-$ and $F3P^-$).

splicing. We subjected the DV–VI region of Ll.LtrB to the transposon screen (Fig. 3A, fragment AatII/HpaI; Fig. 4B, V–VI brackets). The transposon insertion sites were obtained from 30 transconjugants (Fig. 4B, gray arrowheads). Half of the insertions were located within the short 10-nt stretch of DIV that was included in the screen. Four insertions occurred in DV, with three isolates mapping to the same position. However, these four insertion sites produced largely intact DV and are thus equivalent to the insertions found at the end of DIV. Insertion of the Tn5 transposon creates a 9-nt direct repeat of the target sequence (Goryshin and Reznikoff 1998). Therefore, the second *ltrB* fragment, initiated from the P_{23} promoter, starts 9 nt upstream of the indicated arrowheads in Figure 4. The remaining 11 insertions were found in DVI, with six mapping to the same position in the middle of the GUAA tetraloop.

We compared the conjugation efficiencies of selected samples harboring transposon insertions within different domains of Ll.LtrB (Fig. 4B, F7–F11). These conjugation efficiencies are similar to the ones previously observed (Fig. 4B, F1–F6). When LtrA is expressed from a second plasmid, slightly higher conjugation efficiencies were observed, which were significantly higher than the background (Table 1), therefore indicating that Ll.LtrB *trans*-splices efficiently in all cases.

DISCUSSION

In this study, we describe the use of a Tn5-based genetic screen to explore the versatility of group II introns to splice in *trans*. The Tn5-fragmented Ll.LtrB variants were selected for proficient *trans*-splicing introns using a highly sensitive splicing/conjugation assay (Fig. 3C). This experimental approach is both powerful and reliable. First, we identified the Tn5 insertion sites in 43 independent plasmids from the initial pLE12::Tn bank. Insertions were randomly distributed in both the intron and the *ltrB* exons, and occurred in both orientations. Conversely, the 247 plasmids selected by conjugation did not display a random distribution of Tn5 insertions (Fig. 4). Moreover, all Tn5 insertions were found in the same orientation as the relaxase gene such that the terminator and promoter are functional and result in the creation of a bipartite intron. In addition, fragmented introns that were selected with our screen harbored either a full-length or truncated version of *ltrA* that contained the maturase domain, which is essential for Ll.LtrB *trans*-splicing in vivo (Belhocine et al. 2007). However, when LtrA was provided in *trans* during selection, the pressure to express a maturase from the intron was relieved, and introns containing Tn5 insertions throughout the whole *ltrA* gene were selected. As well, removal of the P_{23} promoter from two Ll.LtrB *trans*-splicing variants severely reduced the transfer efficiency of their corresponding plasmid carrier. Taken together, these results confirm the reliability of our screen to identify functional fragmentation sites within Ll.LtrB.

The first two Tn5-based genetic screens that we performed, covering the entire intron, revealed that the majority of the fragmentation sites selected were found in DIB-C and DIV of Ll.LtrB (Fig. 4A). These *trans*-splicing variants of Ll.LtrB are very proficient, some of them splicing almost as efficiently as the non-fragmented Ll.LtrB splicing in *cis* (Table 1, cf. F1–F6 and pLE12). Interestingly, these sites correspond to natural break points identified in organellar *trans*-splicing group II introns (Bonen 1993). Then, by targeting subdomains of Ll.LtrB for fragmentation, we demonstrated that group II introns are considerably more versatile than expected from previously described natural *trans*-splicing group II introns. Indeed, fragmentation sites supporting efficient *trans*-splicing were found not only in DI and DIV but rather distributed throughout Ll.LtrB. As expected, fragmentations are not tolerated in functionally important regions of Ll.LtrB such as EBS1, EBS2, DV, the 5′ and 3′ splice sites, and the area surrounding the branch point in DVI (Fig. 4). This is also generally true for the structurally important long-range tertiary interactions (Fig. 4, pairs of Greek letters). One exception is the α–α′ long-range interaction where both counterparts of this motif were independently interrupted numerous times. We assume that the α–α′ interaction may still be maintained in these Ll.LtrB *trans*-splicing variants since insertion of the Tn5 transposon creates a direct repeat of the target sequence which causes the second *ltrB* fragment to start 9 nt upstream of the fragmentation site, therefore allowing the transcription of an intact α region. However, it is unclear why this small sequence motif is more susceptible to fragmentation than the other small structural motifs present throughout the Ll.LtrB structure.

Our data also confirmed the importance of the maturase function of LtrA for Ll.LtrB *trans*-splicing in vivo (Belhocine et al. 2007). All *trans*-splicing variants selected encoded either a full-length or truncated version of LtrA harboring

the RT and maturase domains. Also, taking into consideration the Tn5 target site duplication of 9 nt, the high-affinity binding site for LtrA, DIVa, is intact in almost all cases. It is interesting to note that the high-affinity binding site for LtrA, and *ltrA*, can be present within either the 3′ or the 5′ intron fragment. LtrA may thus be involved in the assembly of the two intron fragments by binding first to its high-affinity binding site in one fragment, and subsequently making secondary contacts with either DI, DII, or DVI (Wank et al. 1999; Matsuura et al. 2001) in the second fragment, therefore contributing to functional *trans*-splicing of the intron. Potential base pairs and long-range interactions (Fig. 4, pairs of Greek letters) between intron fragments could also be involved in the assembly and *trans*-splicing of Ll.LtrB.

The discovery of group II introns fragmented into two or three pieces strongly supports the theories proposing that self-splicing group II introns are the progenitors of nuclear introns and that the five spliceosomal snRNAs were derived from fragments of group II introns (Sharp 1991). While numerous functional and/or structural parallels can be drawn between regions of group II introns and some snRNAs (Fig. 2), the specific group II intron segments that correspond to the RNA components of the spliceosome machinery are not well defined and were never precisely delineated. Nevertheless, the discovery that a group II intron can be functionally fragmented at multiple locations throughout its structure and between regions that are functionally and/or structurally analogous to the five snRNAs supports the theory suggesting that the snRNAs, which constitute the heart of the spliceosome machinery, originated from the fragmentation of an ancestral group II intron.

MATERIALS AND METHODS

Strains

L. lactis strains NZ9800Δ*ltrB*::*tet* (Belhocine et al. 2004) and LM0231 (Belhocine et al. 2007) were used as donor and recipient for conjugation, respectively. They were grown in M17 broth supplemented with 0.5% glucose (GM17) at 30°C without shaking. *Escherichia coli* strains DH10β and DH5α were used for cloning and plasmid amplification, respectively, and were grown in LB broth at 37°C with shaking. The following concentrations of antibiotics were used: chloramphenicol (Cam), 10 μg/mL; fusidic acid (Fus), 25 μg/mL; spectinomycin (Spc), 300 μg/mL; ampicillin (Amp), 100 μg/mL; erythromycin (Erm), 300 μg/mL.

Plasmids

The pBS*ltrB* plasmid was constructed by cloning the XbaI/SwaI fragment from pLE12 (Fig. 3A; Belhocine et al. 2004) into the unique SmaI site of pBlueScript SK-. This plasmid was used as the target for the Tn5 in vitro transposition assay. The insertion bank generated (pBS*ltrB*::Tn) statistically contained 22 copies of each plasmid carrying a single unique Tn5 insertion between each nucleotide (22 insertions/position); every possible Tn5 insertion was represented 22 times. The statistical representation of the bank was calculated by dividing the number of Spc colonies recuperated following transformation of the transposition assay (Tn5-containing plasmids) by twice the size of the plasmid. This method of calculation takes into consideration that Tn5 can insert randomly on either strand of the target plasmid (Goryshin and Reznikoff 1998; Reznikoff et al. 2004). The pBS*ltrB*::Tn bank was used to construct the various pLE12::Tn banks except the V–VI bank. For the latter, a variant of the target plasmid, pBS*ltrB*AII, carrying an engineered AatII restriction site within DIV was used to create the pBS*ltrB*AII::Tn bank (1.9 insertions/position). The AatII site was engineered by site-directed mutagenesis changing intron nucleotides 2384–2387 from 5′-ACGA-3′ to 5′-CGTC-3′. We confirmed by conjugation that splicing of the resulting intron was not impaired (data not shown). The various pLE12::Tn banks were built by subcloning the corresponding fragments from either the pBS*ltrB*::Tn (full-length, PvuI/HpaI; I–III, PvuI/BsaI; II–III, BsrGI/BsaI) or pBS*ltrB*AII::Tn bank (V–VI, AatII/HpaI) into the pLE12 plasmid, where the unique BamHI site present in the MCS was removed. The pMOD−2<MCS> Transposon Construction Vector from Epicentre (pMOD) was used to construct the Tn5 variant used in this study (Fig. 3B). The P_{23} promoter (two complementary oligos, Table 2) was cloned between the two Tn5 mosaic ends within pMOD (BamHI). The *pepN* terminator was PCR-amplified from pLEII*td*+KR″ (Table 2, primers; Cousineau et al. 1998 and cloned upstream of P_{23} (HindIII). Finally, the *aad9*

TABLE 2. Primers

Name	Sequence
P_{23} top	5′-GATCCTGATGACAAAAAGAGCAAATTTTGATAAAATAGTATTAGAATTGCGGCCGCG-3′
P_{23} bottom	5′-GATCCGCGGCCGCAATTCTAATACTATTTTATCAAAATTTGCTCTTTTTGTCATCAG-3′
5′ *pepN* term	5′-ATTAAGCTTCCAAAAAGCAGCAGTTGATAAAG-3′
3′ *pepN* term	5′-ATTAAGCTTCTTCCTTATTCTCGCTTTGATTG-3′
5′ *aad9*	5′-TAACTGCAGAAGCTTGATTTTCGTTCGTGAATACATGTT-3′
3′ *aad9*	5′-TAACTGCAGCGCTTACCAATTAGAATGAATATTTCCCAA-3′
5′ *ermAM*	5′-AAAGAATTCGACTTAGAAGCAAACTTAAGAGT-3′
3′ *ermAM*	5′-AAAGAATTCCCTTTAGTAACGTGTAACTT-3′
5′ *ltrA*	5′-CGGGGTACCTTGATTAGGGAGGAAAACCTCAA-3′
3′ *ltrA*	5′-CGGGGTACCTCACTTGTGTTTATGAATCACGT-3′

gene conferring spectinomycin resistance (Spc), flanked by its promoter and terminator, was PCR-amplified from pDL278 (Table 2, primers; Belhocine et al. 2004) and cloned between the *pepN* terminator and P_{23} (Fig. 2B, PstI). pDE*ltrA* was constructed by first cloning the PCR-amplified *ltrA* gene (Table 2, primers) in the pDL-P_{23}[2] plasmid (Belhocine et al. 2007) under the control of P_{23} (NotI). The *ermAM* gene conferring erythromycin resistance (Erm) was PCR-amplified from pMSP3545 (Bryan et al. 2000) and cloned in the resulting plasmid (SmaI). Finally, the *aad9* gene conferring Spc resistance was deleted (PacI/SwaI).

Conjugation, RT-PCR, and Tn5 transposition assays

Conjugations were performed on milk plates as previously described (Belhocine et al. 2007). Conjugation efficiency was calculated as the ratio of transconjugant cells (Cam/Fus) to donor cells (Cam) for three independent assays. RT-PCR assays were performed as previously described (Belhocine et al. 2007). For the in vitro transposition assay, the Tn5 transposon was excised from pMOD (PvuII) and purified on agarose gel. The transposon was incubated with the pBS*ltrB* or pBS*ltrB*AII target plasmids (14:1 ratio) at 37°C for 2 h, in the presence of the EZ::Tn5 transposase recombinant protein (Epicentre). A fraction of the transposition mix (10%) was then electroporated into DH10β ElectroMAX competent cells (Invitrogen) and plated on LB/Spc plates to select for transposon-containing plasmids. Notably, insertion of the Tn5 transposon creates a 9-nt direct repeat (Goryshin and Reznikoff 1998). Therefore, the second *ltrB* fragment that initiated from the P_{23} promoter within the transposon starts 9 nt upstream of the indicated arrowheads in Figure 4.

ACKNOWLEDGMENTS

We thank L. Bonen, L. Kronstad, F. Lang, R. Lease, C. Ritlop, A. Stern, and K.K. Yam for critical reading of the manuscript, and L. Kronstad for technical assistance. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to B.C. B.C. is a FRSQ Chercheur-Boursier Junior 2 and a McGill University William Dawson Scholar. K.B. was supported by a Postgraduate Studies Fellowship from NSERC.

Received March 17, 2008; accepted May 12, 2008.

REFERENCES

Belhocine, K., Plante, I., and Cousineau, B. 2004. Conjugation mediates transfer of the Ll.LtrB group II intron between different bacterial species. *Mol. Microbiol.* **51:** 1459–1469.

Belhocine, K., Mak, A.B., and Cousineau, B. 2007. *Trans*-splicing of the Ll.LtrB group II intron in *Lactococcus lactis*. *Nucleic Acids Res.* **35:** 2257–2268.

Bonen, L. 1993. *Trans*-splicing of pre-mRNA in plants, animals, and protists. *FASEB J.* **7:** 40–46.

Bryan, E.M., Bae, T., Kleerebezem, M., and Dunny, G.M. 2000. Improved vectors for nisin-controlled expression in gram-positive bacteria. *Plasmid* **44:** 183–190.

Byrd, D.R. and Matson, S.W. 1997. Nicking by transesterification: The reaction catalysed by a relaxase. *Mol. Microbiol.* **25:** 1011–1022.

Cavalier-Smith, T. 1991. Intron phylogeny: A new hypothesis. *Trends Genet.* **7:** 145–148.

Cech, T.R. 1986. The generality of self-splicing RNA: Relationship to nuclear mRNA splicing. *Cell* **44:** 207–210.

Cousineau, B., Smith, D., Lawrence-Cavanagh, S., Mueller, J.E., Yang, J., Mills, D., Manias, D., Dunny, G., Lambowitz, A.M., and Belfort, M. 1998. Retrohoming of a bacterial group II intron: Mobility via complete reverse splicing, independent of homologous DNA recombination. *Cell* **94:** 451–462.

Fedorova, O. and Zingler, M. 2007. Group II introns: Structure, folding and splicing mechanism. *Biol. Chem.* **388:** 665–678.

Goryshin, I.Y. and Reznikoff, W.S. 1998. Tn5 in vitro transposition. *J. Biol. Chem.* **273:** 7367–7374.

Hetzer, M., Wurzer, G., Schweyen, R.J., and Mueller, M.W. 1997. *Trans*-activation of group II intron splicing by nuclear U5 snRNA. *Nature* **386:** 417–420.

Jacquier, A. 1990. Self-splicing group II and nuclear pre-mRNA introns: How similar are they? *Trends Biochem. Sci.* **15:** 351–354.

Jarrell, K.A., Dietrich, R.C., and Perlman, P.S. 1988. Group II intron domain 5 facilitates a *trans*-splicing reaction. *Mol. Cell. Biol.* **8:** 2361–2366.

Lambowitz, A.M. and Zimmerly, S. 2004. Mobile group II introns. *Annu. Rev. Genet.* **38:** 1–35.

Matsuura, M., Saldanha, R., Ma, H., Wank, H., Yang, J., Mohr, G., Cavanagh, S., Dunny, G.M., Belfort, M., and Lambowitz, A.M. 1997. A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: Biochemical demonstration of maturase activity and insertion of new genetic information within the intron. *Genes & Dev.* **11:** 2910–2924.

Matsuura, M., Noah, J.W., and Lambowitz, A.M. 2001. Mechanism of maturase-promoted group II intron splicing. *EMBO J.* **20:** 7259–7270.

Michel, F. and Ferat, J.L. 1995. Structure and activities of group II introns. *Annu. Rev. Biochem.* **64:** 435–461.

Mills, D.A., McKay, L.L., and Dunny, G.M. 1996. Splicing of a group II intron involved in the conjugative transfer of pRS01 in lactococci. *J. Bacteriol.* **178:** 3531–3538.

Newman, A.J. and Norman, C. 1992. U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. *Cell* **68:** 743–754.

Nilsen, T.W. 1994. RNA–RNA interactions in the spliceosome: Unraveling the ties that bind. *Cell* **78:** 1–4.

Patel, A.A. and Steitz, J.A. 2003. Splicing double: Insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* **12:** 960–970.

Reznikoff, W.S., Goryshin, I.Y., and Jendrisak, J.J. 2004. Tn5 as a molecular genetics tool: In vitro transposition and the coupling of in vitro technologies with in vivo transposition. *Methods Mol. Biol.* **260:** 83–96.

Seetharaman, M., Eldho, N.V., Padgett, R.A., and Dayie, K.T. 2006. Structure of a self-splicing group II intron catalytic effector domain 5: Parallels with spliceosomal U6 RNA. *RNA* **12:** 235–247.

Sharp, P.A. 1985. On the origin of RNA splicing and introns. *Cell* **42:** 397–400.

Sharp, P.A. 1991. Five easy pieces. *Science* **254:** 663.

Shearman, C., Godon, J.J., and Gasson, M. 1996. Splicing of a group II intron in a functional transfer gene of *Lactococcus lactis*. *Mol. Microbiol.* **21:** 45–53.

Shukla, G.C. and Padgett, R.A. 2002. A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. *Mol. Cell* **9:** 1145–1150.

Valadkhan, S. 2007. The spliceosome: A ribozyme at heart? *Biol. Chem.* **388:** 693–697.

Wank, H., SanFilippo, J., Singh, R.N., Matsuura, M., and Lambowitz, A.M. 1999. A reverse transcriptase/maturase promotes splicing by binding at its own coding segment in a group II intron RNA. *Mol. Cell* **4:** 239–250.