

Dispersion of the Rmlnt1 group II intron in the *Sinorhizobium meliloti* genome upon acquisition by conjugative transfer

Rafael Nisa-Martínez, José I. Jiménez-Zurdo, Francisco Martínez-Abarca, Estefanía Muñoz-Adelantado and Nicolás Toro*

Grupo de Ecología Genética de la Rizosfera, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, Profesor Albareda 1, 18008 Granada, Spain

Received September 8, 2006; Revised October 23, 2006; Accepted November 9, 2006

ABSTRACT

Rmlnt1 is a self-splicing and mobile group II intron initially identified in the bacterium *Sinorhizobium meliloti*, which encodes a reverse transcriptase–maturase (Intron Encoded Protein, IEP) lacking the C-terminal DNA binding (D) and DNA endonuclease domains (En). Rmlnt1 invades cognate intronless homing sites (*IS_{Rm2011-2}*) by a mechanism known as retrohoming. This work describes how the Rmlnt1 intron spreads in the *S.meliloti* genome upon acquisition by conjugation. This process was revealed by using the wild-type intron Rmlnt1 and engineered intron-donor constructs based on ribozyme coding sequence (Δ ORF)-derivatives with higher homing efficiency than the wild-type intron. The data demonstrate that Rmlnt1 propagates into the *S.meliloti* genome primarily by retrohoming with a strand bias related to replication of the chromosome and symbiotic megaplasmids. Moreover, we show that when expressed *in trans* from a separate plasmid, the IEP is able to mobilize genomic Δ ORF ribozymes that afterward displayed wild-type levels of retrohoming. Our results contribute to get further understanding of how group II introns spread into bacterial genomes in nature.

INTRODUCTION

Group II introns are both catalytic RNAs and mobile retroelements initially found in organelles of plants and lower eukaryotes, but later identified in bacteria and in some archaeal genomes (1–3). A typical group II intron consists of a highly structured RNA in six distinct double-helical domains (DI to DVI) and an internally encoded (ORF within

DIV) reverse transcriptase–maturase (RT) that is required *in vivo* to fold the intron RNA into a catalytically active structure (4–8). Intron mobility may occur either to intronless alleles (retrohoming) or to ectopic sites (retrotransposition) and is mediated by RNP (ribonucleoprotein) complexes consisting of the Intron Encoded Protein (IEP) and the excised lariat RNA.

Three main phylogenetic classes (IIA, IIB and IIC) of group II introns have been described based on the IEP and conserved intron RNA structures (2,5,9–12). The well-characterized *Lactococcus lactis* LI.LtrB and yeast aI1 and aI2 introns belongs to the IIA class, the IEPs of which contain an N-terminal RT domain homologous to retroviral RTs, followed by domain X, associated with RNA splicing or maturase activity, and C-terminal DNA binding (D) and DNA endonuclease domains (En) (13–16). Retrohoming of these group IIA introns occurs by a target-DNA-primed reverse transcription mechanism in which the intron RNA reverse splices directly into one strand of a double-stranded DNA target site, whereas the IEP cleaves the opposite strand through its endonuclease domain and uses the 3' end as a primer for reverse transcription of the inserted intron RNA (3,17,18).

Group II intron retrotransposition occurs at low frequency and it is considered one of the mechanisms for intron dispersal in nature. Similar to retrohoming, retrotransposition also occurs via RNA intermediate that reverse splices into DNA targets (19–21). Retrotransposition of the LI.LtrB intron in *L.lactis* is biased toward the template for lagging-strand DNA synthesis and occurs in an endonuclease-independent manner, which suggests that the intron reverse splices into transiently single-stranded DNA at the replication fork (21,22). In *Escherichia coli* retrotransposition of LI.LtrB frequently occurs into double-stranded DNA in an endonuclease-dependent manner similar to the retrohoming pathway. Thus, LI.LtrB retrotransposes through two distinct mechanisms and the host environment seems to influence the choice of the integration pathway (23). Lateral inter- and intra-species transfer mediated by conjugation has been

*To whom correspondence should be addressed. Tel: +3 495 818 1600; Fax: +3 495 812 9600; Email: ntoro@eez.csic.es
Present address:
Estefanía Muñoz-Adelantado, Centre de Regulació Genòmica, 08003 Barcelona, Spain

recently assessed for the LI.LtrB intron (24,25) showing that upon acquisition by bacteria this intron is able to spread into the chromosome by retrotransposition, invading non-homologous sites. Nevertheless, the low level of *ltrB* transcript and relatively inefficient splicing of the intron may limit LI.LtrB mobility and dissemination in nature (26).

Strikingly, over half of bacterial group II IEPs annotated to date lack the En domain and many other like those belonging to the IIB3 subclass (2) also lack the D domain. It has been proposed that such group II introns are ancestral and they target DNA with minimal sequence constraints to be recognized by its IEP (3,18,19,21). The best studied intron of this subclass is the *Sinorhizobium meliloti* RmInt1, a very efficient mobile element (27). Recently, it was determined that RmInt1 has two retrohoming pathways for mobility with a preferred pathway consistent with reverse splicing of the intron RNA into single-stranded DNA at a replication fork, using a nascent lagging DNA strand as the primer for reverse transcription (28). Thus, the mechanism of retrohoming of RmInt1 resembles that of ectopic transposition of LI.LtrB in *L.lactis*.

Most bacterial group II introns are located within other mobile elements and outside the genes [(2) and references therein]. The natural homing site of RmInt1 is the insertion sequence (IS) *ISRm2011-2* (29), which is usually present in high copy number in the genome of virtually all *S.meliloti* isolates. RmInt1 is present in 90% of the *S.meliloti* isolates (30) and also in other *Sinorhizobium* and *Rhizobium* species, where the intron homing sites correspond to IS elements of the *ISRm2011-2* group (31). Thus, possibly this intron has been acquired by *Sinorhizobium* and *Rhizobium* species by vertical and horizontal inheritance (31). The absence of RmInt1 in ~10% of the *S.meliloti* isolates is intriguing, since in these isolates there is no restriction to intron mobility. RmInt1 is therefore a model system to study group II intron survival, evolution and dissemination in nature.

In this work we investigate the retrohoming ability and dispersion pathway of RmInt1 in the *S.meliloti* genome

upon intron acquisition by conjugative transfer. We used wild-type and engineered intron-donor constructs based on ribozyme coding sequence (Δ ORF)-derivatives with increased retrohoming efficiency. Our data show the extraordinary ability of RmInt1 for its dispersal in nature by retrohoming and reveal the pathway followed by the intron to colonize the natural homing sites in this bacterial genome.

MATERIALS AND METHODS

Bacterial strains and growth conditions

S.meliloti strains used in this work were 1021 (genome-sequenced strain) (32) and RMO17 (RmInt1 intronless strain) (33). *E.coli* DH5 α was used for cloning and maintenance of plasmid constructs. Rhizobial strains were routinely grown at 28°C on complete TY or defined minimal medium (MM) (34) and *E.coli* DH5 α at 37°C on Luria–Bertani (LB) medium. Media were supplemented when required with antibiotics at the following concentrations: kanamycin 200 μ g/ml for rhizobia and 50 μ g/ml for *E.coli*; tetracycline 10 μ g/ml and ampicillin 200 μ g/ml.

RmInt1-donor constructs for *in vivo* assays

Schematics of all the constructs described below for delivery of RmInt1 and derivatives are depicted in Figures 1A and 2B. Plasmids expressing the wild-type RmInt1 intron were all pBBR1MCS-2 derivatives (K_m^r). pKG2 was generated from pKG2.5 (27) by trimming the exon sequences to -50/+146 as described (35) whereas pBB2.5, pBB20/5H and pBB15/10H were generated *in vivo* by invasion of the corresponding target-recipient plasmids (27,36). To construct RmInt1- Δ ORF-derivatives a DNA fragment deleted of intron 611–1759 nt was generated as described previously from pKG2 (35) and subsequently inserted in the BamHI site of pKGIEP (37) just upstream the IEP coding sequence to yield pKG Δ ORFIEP.

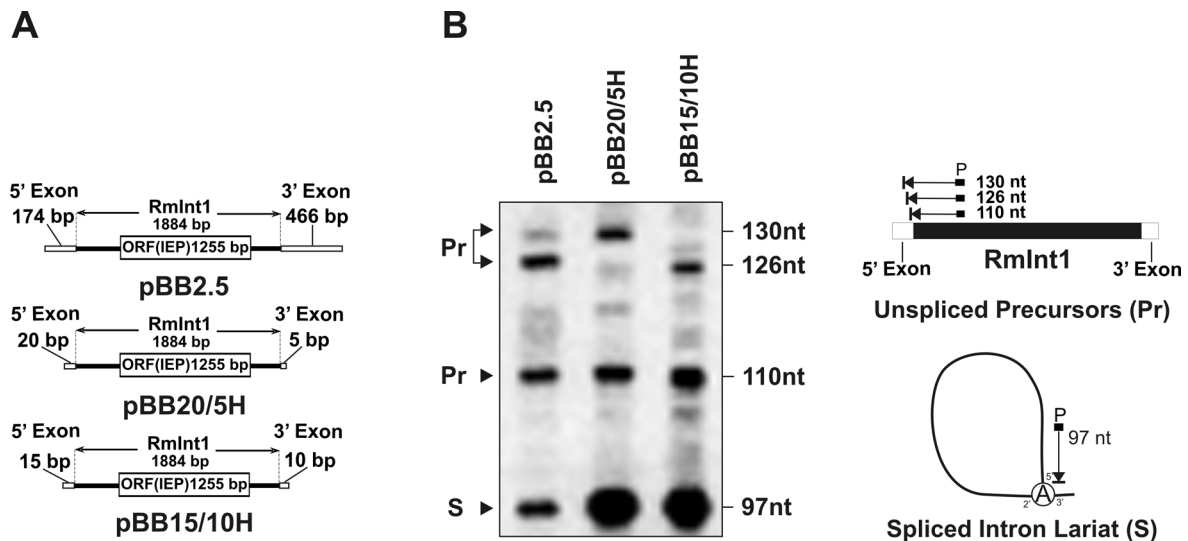


Figure 1. Effect of exon lengths on RmInt1 splicing *in vivo*. (A) Schematics of constructs used in the primer extension analysis. (B) Primer extension. Analysis was performed on total RNA (20 μ g) from RMO17 cells harboring intron-donor plasmids pBB2.5, pBB20/5H or pBB15/10H. The 97 nt cDNA product corresponds to the excised intron RNA (S) whereas the larger products of 110, 126 and 130 nt are derived from unsplined precursor RNA molecules (Pr). Schematics of primer extension products are shown to the right of the panel.

Δ ORF intron variants flanked by exon sequences $-20/+5$ were then generated by PCR amplification of the BamHI Δ ORF fragment using pKG Δ ORFIEP as template and the appropriate primers to yield DNA products that were inserted as BamHI or SacI fragments in pKGIEP upstream (pKG Δ ORFIEP20/5) or downstream (pKGIEP Δ ORF20/5) the IEP coding sequence, respectively. pKGEMA4 was used as a donor plasmid for delivery of RmInt1 into the *S.meliloti* RMO17 and 1021 genomes were generated by engineering PmlI and BlnI sites internal to SacI in pKGIEP Δ ORF20/5.

Further deletions of intron DIV extending positions 585–1781 (Δ ORFcd) and 576–1790 (Δ ORFab) were also generated as described by inverse PCR using primer pairs C (5'-GCCGCTCGAGATTTCGAAACGGCTTGTCTGT-3') and D (5'-GCCGCTCGAGTTC-GGAACGAACACGTTTA-3'); and A (5'-GCCGCTCGAGCT-TGTCTGCTGACTTTCC-3') and B (5'-GCCGCTCGAGAACACGTTTACCTGA-TGGGA-3'), respectively and assayed in a IEP+ Δ ORF ($-20/+5$) configuration for intron homing.

***In vivo* splicing and homing assays**

RmInt1 splicing *in vivo* was assessed by primer extension analysis on total RNA from RMO17 cells expressing the wild-type intron from plasmids pBB2.5, pBB20/5H and pBB15/10H as described previously (37). cDNA-bands corresponding to the resolved extension products were quantified with the Quantity One software package (Bio-Rad Laboratories) and intron splicing was measured as $100[S/(S+Pr)]$, where S is the spliced RNA and Pr the precursor RNA species. *In vivo* homing efficiency of the wild-type and Δ ORF-derivatives was determined by the double-plasmid assay in *S.meliloti* strain RMO17 using pJB0.6LEAD or pJB0.6LAG as recipient plasmids. The homing efficiency was calculated as the percentage of the ratio of homing product to the addition of homing product and non-invaded recipient plasmid. Data are the average of at least four independent determinations with their corresponding standard deviation (28,36,37). Recipient plasmid pJB Δ 129 that lacks the RmInt1 target was used as negative control in the assays (27).

Analysis of RmInt1 dispersal in the *S.meliloti* genome and mobility of the inserted introns

pKG2.5 and pKGEMA4 were independently mobilized to RMO17 and 1021 strains by triparental matings using pRK2013 as helper plasmid. Retrohoming events were then assessed in individual transconjugants from each cross (6–7 colonies) by Southern hybridization of BamHI- or SallI-digested genomic DNA (AquaPure Genomic DNA Isolation Kit, Bio-Rad Laboratories) with DIG-labeled probes specific to either the *ISRm2011-2* or the RmInt1 Δ ORF as described previously (38). pKG2.5x (27) and pKGEMA4x were used as controls in the assays carry a frameshift within the IEP coding sequence generated by eliminating the internal EcoRI in the intron sequence.

To assess homing efficiency *in vivo* of the newly acquired introns, two individual RMO17 derivatives containing genomic insertions of the wild-type (four insertions) or Δ ORF (five insertions) intron, were cured of the donor plasmids pKG2.5 or pKGEMA4 by subjecting bacteria to successive cultures in MM broth without the antibiotic K_m thus

generating strains RMO17-2.5c2 and RMO17 Δ Ic50. Plasmid pJB0.6LAG containing the RmInt1 target (28) was then mobilized to the cured strains and homing efficiency was assessed as described previously on plasmid DNA of individual transconjugants (36).

pKGIEP (37) expressing a functional IEP and pKG4dV that expresses a non-functional ribozyme were also independently mobilized by conjugation to RMO17 Δ Ic50 harboring plasmid pJB0.6LAG. pKG4dV is a pKGEMA4 derivative in which nucleotides GTT (RmInt1 nucleotide positions 1840–1842) within the Δ ORF wild-type fragment were replaced by CGA using the Altered Sites II *in vitro* Mutagenesis System (Promega). Homing efficiency of complemented Δ ORF variants was determined on plasmid DNA of the transconjugants as described. Plasmid pKGIEPYAHH, a pKGIEP derivative that expresses a non-functional IEP containing a mutation in the RT domain (37) was used as negative complementation control in these experiments.

RESULTS

Engineering RmInt1-donor constructs for intron delivery

Recently, we reported that RmInt1 self-splices *in vitro*, a reaction that was favored by removing most of the protein coding sequence from the large terminal loop of DIV and by trimming the 5' exon to its last 20 and 15 nt (35,39). Furthermore, in a previous work, we reported that a target site extending from positions -20 to -1 of the 5' exon and $+1$ to $+5$ of the 3' exon was sufficient to ensure efficient homing of intron RmInt1 (36).

In order to engineer a highly efficient RmInt1-donor construct for intron delivery, we first analyzed the influence of the length of the exons on intron splicing *in vivo* (Figure 1). The splicing reaction was characterized by primer extension using a primer P complementary to a sequence located 80–97 nt from the 5' end of the intron. In these assays, the excised intron RNA (S) is detected as an extension product of 97 nt, along with larger bands derived from unspliced precursor RNA molecules (Pr) (40). As shown in Figure 1B, the extension product corresponding to the excised intron was detected with all the RNA samples from cells harboring constructs (Figure 1A) that transcribe the intron with either long ($-174/+466$) or short stretches ($-20/+5$, $-15/+10$) of the wild-type flanking 5' and 3' exons sequences along with additional sequences from the plasmid. The estimated relative amount of excised intron was $\sim 32\%$ when transcribed with long exon sequences (pBB2.5), but it did increase 2-fold when only short stretches of the original exons were retained ($-20/+5$ and $-15/+10$ in pBB20/5H and pBB15/10H, respectively). Our results therefore indicate that exons trimming favors RmInt1 splicing *in vivo* as also occurs *in vitro*.

It has been previously reported that mobility frequencies of *L.lactis* group II intron Ll.LtrB can be significantly increased by using donor constructs expressing the LtrA protein from a position either downstream or upstream of the Δ ORF, in which a large portion of the LtrA ORF within DIV had been deleted (41,42). To assess the effect of ORF deletions in RmInt1, homing efficiency (fraction of recipient targets invaded by the intron) was determined by double-plasmid

mobility assays using a set of Δ ORF- and wild-type donor plasmids all expressing the intron from the kanamycin resistance gene constitutive promoter. As shown in Figure 2A a RmInt1 Δ ORF derivative deleted of intron 611–1759 nt (35) flanked by $-20/+5$ exons sequences and expressing the IEP upstream of the ribozyme (IEP+ Δ ORF) yielded invasion rates on pJB0.6LEAD of 40–42%. Larger deletions affecting subdomain DIVb reduced homing efficiency to $\sim 30\%$ (Δ ORFcd) and to $<2\%$ (Δ ORFab). Donor constructs pKG2, in which the wild-type intron is flanked by exon sequences $-50/+146$, and three Δ ORF-derivatives deleted of intron 611–1759 nt (35) and flanked by the same ($-50/+146$) or shorter ($-20/+5$) exon sequences, that express the IEP either downstream or upstream of the ribozyme (Δ ORF+IEP and IEP+ Δ ORF, respectively) were further tested for homing efficiency (Figure 2B). The target-recipient plasmids used in the homing assays were pJB0.6LEAD and pJB0.6LAG, in which the RmInt1 target site is cloned in both orientations with respect to the direction of DNA replication (28). As shown in Figure 2B the wild-type intron (pKG2) exhibited $\sim 10\%$ homing efficiency in pJB0.6LEAD whereas the use of a Δ ORF+IEP variant flanked by the same exon sequences ($-50/+146$) resulted in a higher homing efficiency ($\sim 23\%$). Interestingly, constructs Δ ORF+IEP and IEP+ Δ ORF flanked by exon sequences $-20/+5$ displayed even higher homing efficiency (33 and 37%, respectively). Consistent with RmInt1 retrohoming bias toward the strand that serves as template for the nascent lagging-strand in DNA replication forks (28) the invasion rates in pJB0.6LAG recipient plasmid were markedly increased for all the donor constructs, but again the highest efficiency corresponded to Δ ORF intron constructs flanked by short exons ($\sim 90\%$, Figure 2B). In addition to an increased homing efficiency, the combination of Δ ORF intron and short exons resulted even in rates of intron splicing close to 100% (data not shown). Taken together these data predict the Δ ORF constructs flanked by short exon sequences ($-20/+5$) and expressing the IEP in *cis* within the same precursor transcript as the most efficient shuttle configuration for RmInt1 delivery *in vivo*.

Dispersion of RmInt1 in the *S. meliloti* genome

RMO17 is an RmInt1 intronless *S. meliloti* strain that nevertheless harbors 13 copies of the RmInt1 homing site (*ISRm2011-2*). To determine how RmInt1 colonizes the *S. meliloti* genome, the wild-type intron (pKG2.5) and an IEP+ Δ ORF derivative flanked by short ($-20/+5$) exons (pKGEMA4) were transferred to RMO17 by conjugation. Upon acquisition of either pKG2.5 or pKGEMA4, 100% of the transconjugants exhibited genomic retrohoming events revealing the extraordinary spreading ability of RmInt1. Retrohoming did not occur when the intron-donor plasmids encode an IEP truncated in the RT domain (pKG2.5x or pKGEMA4x; Figure 3), which is consistent with the requirement of reverse transcriptase activity and a functional IEP for intron mobility. We analyzed six independent RMO17 transconjugants harboring pKG2.5 and interestingly, in all these six clones the insertion of the wild-type intron occurred into four particular *ISRm2011-2* genomic copies (IS3, IS4, IS6 and IS9) (the hybridization pattern of one of the clones

is shown in Figure 3, left panel). Consistent with higher retrohoming efficiency, the Δ ORF intron pKGEMA4 integrated into a larger number of homing sites. Seven independent transconjugants were analyzed, six displayed identical hybridization pattern to the RmInt1 probe sharing four retrohoming events with the RMO17 clones harboring pKG2.5 (IS3, IS4, IS6 and IS9) whereas insertion into IS2 was an additional retrohoming event. Retrohoming into IS13 was only detected in one clone of the seven analyzed (the hybridization pattern of this transconjugant is shown in Figure 3, right panel). These results suggest that certain homing sites favor RmInt1 retrohoming in the *S. meliloti* genome.

To get further insight into the dispersion pathway of RmInt1 in the *S. meliloti* genome, pKG2.5 and pKGEMA4 were transferred by conjugation into strain 1021, of which the genome has been sequenced (32). The tripartite genome of this *S. meliloti* strain is 6.7 Mb and contains a chromosome of 3.65 Mb, as well as two symbiotic megaplasmids known as pSymA (1.35 Mb) and pSymB (1.68 Mb). Strain 1021 harbors 13 homing sites for RmInt1 that include twelve copies of *ISRm2011-2* and one copy of the closely related insertion sequence *ISRm10-1*. Two copies of *ISRm2011-2* (one in pSymA and one in pSymB) are invaded by RmInt1, as occurs with the unique copy of *ISRm10-1* located in pSymB (Figure 4B). The three homing sites invaded by RmInt1 reside on the template for the lagging-strand synthesis, which is consistent with the preferred retrohoming pathway for RmInt1 and other bacterial group II introns (21,24,28). It should be noted that the *ISRm2011-2* intronless copies in strain 1021 all contain the characteristic nucleotide positions of RmInt1 wild-type target site. Analysis of independent 1021 transconjugants harboring pKG2.5 showed that the clones were not homogenous and the process of invasion of the homing sites by RmInt1 seems to occur in a gradual manner. This is revealed by the marked reduction in the intensity of hybridizing bands corresponding to intronless homing sites and the concomitant appearance of new ones containing the intron, with increased intensity with larger generations (compare hybridization patterns after 30 and 250 generations in Figure 4A, left panel). Our results revealed that upon acquisition of pKG2.5 five particular homing sites were gradually occupied by RmInt1. Likewise the Δ ORF intron variant integrated into a higher number of homing sites as expected (Figure 4A, right panel). The wild-type intron expressed from pKG2.5 integrated into four chromosomal copies of *ISRm2011-2* (ISc3, ISc4, ISc5 and ISc6) and one copy in pSymA (ISa1), all target-sites located on the template for the lagging-strand synthesis. The Δ ORF variant inserted in the same homing sites and two more, one in pSymA (ISa3) and other in the chromosome (ISc2) (Figure 4B). Interestingly, these two additional retrohoming events occurred in target-sites located on the template for the leading-strand synthesis. Furthermore, some other clones carried one additional insertion into ISa4 also in the leading-strand orientation (Figure 4B). Taken together, these results suggest that the retrohoming pathway followed by RmInt1 for its dispersion in the *S. meliloti* genome is related to replication of the chromosome and symbiotic megaplasmids and prefers the template for lagging-strand synthesis as the target site.

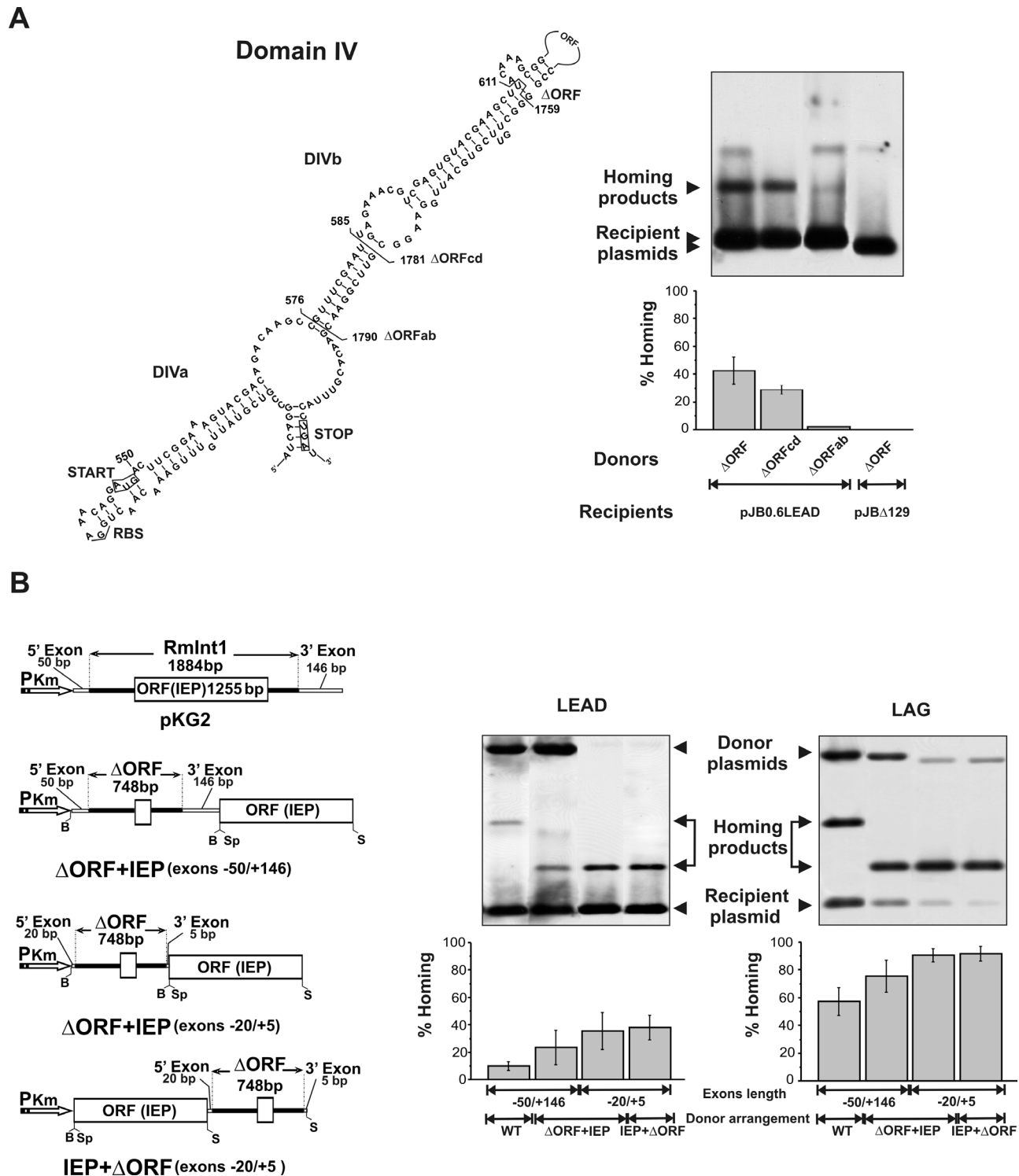


Figure 2. Homing efficiency of wild-type RmInt1 and Δ ORF-derivatives. (A) Deletion analysis of RmInt1 DIV. The predicted secondary structure of intron DIV is shown with indication of subdomains DIVa and DIVb. Start and stop codons of the IEP ORF are boxed and brackets demarcate the internal deletions (Δ ORF, Δ ORFcd and Δ ORFab) cloned downstream of the IEP coding sequence in the donor constructs. For homing assays plasmid pools from RMO17 cells harbouring donor and recipient (pJB0.6LEAD) plasmids were analyzed by Southern hybridization with a DNA probe specific to the insertion sequence *IS_{Rm2011-2}*. Recipient plasmid pJB Δ 129 was used as negative control in the assays. Target invasion rates in each homing assay were calculated as described in Materials and Methods and plotted in the graph-bar shown below the blot. (B) Effect of different donor genetic arrangements on intron homing. The configuration of the donor constructs is diagrammed to the left. PK_m, promoter of K_m resistance gene; B, *Bam*HI; Sp, *Spe*I; S, *Sac*I. Homing efficiency on recipient plasmids pJB0.6LEAD and pJB0.6LAG was quantified as described in A. The Southern blots are shown to the right along with the corresponding histograms. Features of the donor constructs are indicated below the plots. The hybridization signal corresponding to the donor plasmid with short flanking exons is variable and depends on the specific activity of the probe.

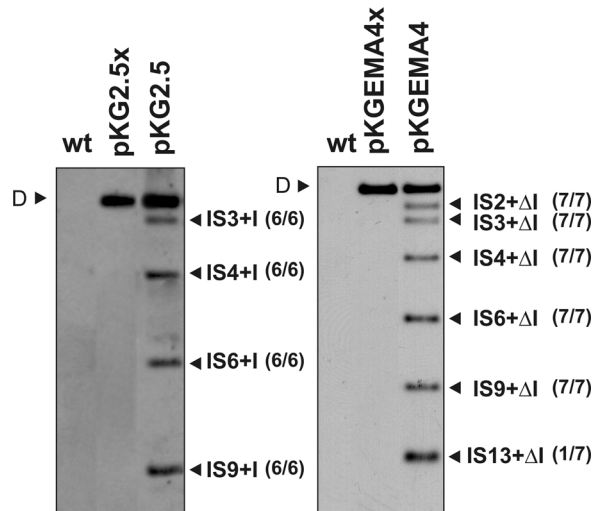


Figure 3. Retrohoming of RmInt1 into the intronless RMO17 genome. Panels show Southern hybridizations of *SalI*-digested DNA from the RMO17 strain (wt) and derivatives harboring plasmids pKG2.5 (donor of the wild-type intron, left panel), pKGEMA4 (expressing a Δ ORF variant, right panel), pKG2.5x or pKGEMA4x (both expressing non-mobile introns). The different lanes correspond to individual clones showing all the homing events, using as donor plasmids either pKG2.5 or pKGEMA4. Probe was specific to the 5' end of the intron ribozyme sequence. D indicates the hybridization signal corresponding to the donor plasmid DNA. Genomic insertions of the wild-type intron (IS+I) or its Δ ORF variant (IS+ Δ I) are marked. IS were denoted with numbers (1–13) according to the hybridization pattern provided by *SalI*-digested total DNA using an IS probe numbering from the larger to the smaller hybridizing IS-containing fragment. Number of transconjugants with the observed events is indicated in brackets alongside the bands.

Retrohoming ability of the genomic inserted intron copies and *trans* complementation of the Δ ORF intron with the IEP coding sequence

The results shown above suggest that the homing sites in the genome of *S. meliloti* will be gradually and sequentially colonized by RmInt1, as long as the intron is expressed in the bacterial cell. To test whether the newly acquired intron genomic copies retain retrohoming ability the RMO17 derivative strains were cured of donor plasmids (pKG2.5 or pKGEMA4) to generate RMO17-2.5c2 and RMO17- Δ Ic50, respectively (Figure 5A) and a target-recipient plasmid pJB0.6LAG was then transferred by conjugation into these two strains. Retrohoming efficiency of the acquired wild-type introns in RMO17-2.5c2 was estimated to be 10–13% (Figure 5B, lane 4), whereas no insertion events into the target-recipient plasmid were detected in RMO17- Δ Ic50, colonized by the Δ ORF intron variant (data not shown). These results indicate that the acquired wild-type intron genomic copies retain retrohoming ability and therefore they could contribute to further intron dispersal. Furthermore, retrohoming of the intron genomic copies requires the IEP, which cannot be replaced by the host genetic background.

Intron LI.LtrB maturase (LtrA) functions most efficiently when expressed in *cis* either from its normal location within the intron or from a position either upstream or downstream of the Δ ORF. The maturase also functions when expressed in *trans* from a separate plasmid, but splicing is <10% of that for a *cis* configuration (43). To investigate the *trans* complementation of the RmInt1 IEP functions, plasmids pKGIEP

and pKG4dV were individually conjugated into strain RMO17- Δ Ic50 harboring the compatible target-recipient plasmid pJB0.6LAG. Both plasmids express the wild-type IEP from a kanamycin resistance gene constitutive promoter either alone (pKGIEP) or within the context of a splicing defective ribozyme that contains a mutation within the catalytic DV (GUU \rightarrow CGA; pKG4dV). pKGIEPYAHH expressing a non-functional IEP with a mutation in the RT domain (YADD \rightarrow YAHH) was used as negative complementation control. Interestingly, when complemented in *trans* with a functional IEP either alone or within the context of a defective ribozyme, the genomic copies of the Δ ORF intron exhibited \sim 10% homing efficiency in pJB0.6LAG, which is comparable to that displayed by the wild-type intron (Figure 5B, lanes 2 and 3). As expected, the RT mutant IEP was not able to mobilize the Δ ORF intron (lane 1). These results presumably reflect the stability of the free protein and unspliced intron precursor RNA in the *S. meliloti* host and let us to hypothesize that inactive intron copies could still contribute to intron dispersal by supplying or being complemented in *trans* with an active IEP.

DISCUSSION

The dispersal of group II introns into bacterial genomes has been addressed mainly by analyzing retrotransposition of the LI.LtrB intron (21–24,44). Dissemination by retrohoming has been also reported, but these studies were limited to the analysis of a reduced number of chromosomal events and the invasion of target-recipient plasmids (24,25) or to the use of engineered intron-donors (targetron) that could be influenced by many factors inherent to such experimental systems (45,46). In this study, we determine how a bacterial genome is colonized by a group II intron upon acquisition by conjugative transfer. This analysis was possible due to the existence of numerous unoccupied RmInt1 homing sites in the genome of some *S. meliloti* strains, thus providing further insights on intron dispersal in natural bacterial populations.

We first showed that RmInt1 splicing *in vivo* and its homing efficiency were greatly improved by removing most of the protein coding sequence from the large terminal loop of DIV and by using short flanking exon sequences extending 20 nt in the 5' exon and 5 nt in the 3' exon. This improvement on splicing was previously observed for the RmInt1 *in vitro* splicing reaction (35,39), although the influence of a shorter 3' exon was not investigated. Furthermore, homing frequencies of a LI.LtrB- Δ ORF derivative expressing the IEP downstream or upstream of the Δ ORF was greatly stimulated, presumably because a greater resistance of Δ ORF intron to nucleolytic cleavage in DIV (41,42). For RmInt1, we determined that the IEP also functions efficiently either upstream or downstream the Δ ORF intron. The Δ ORF variant used in our assays still contains one DIVb stem of \sim 30 nt pairings (Figure 2A). Deletion of LI.LtrB DIVb inhibited splicing by \sim 60% (43) and similar deletions in the yeast aI2 intron coding sequence strongly-inhibited maturase-promoted reverse splicing in an *in vitro* assay, but its effect was not observed *in vivo*. We found that a larger deletion of RmInt1 DIVb reduced homing efficiency and a full deletion almost inhibited homing (Figure 2A). Thus, DIVb is required for efficient

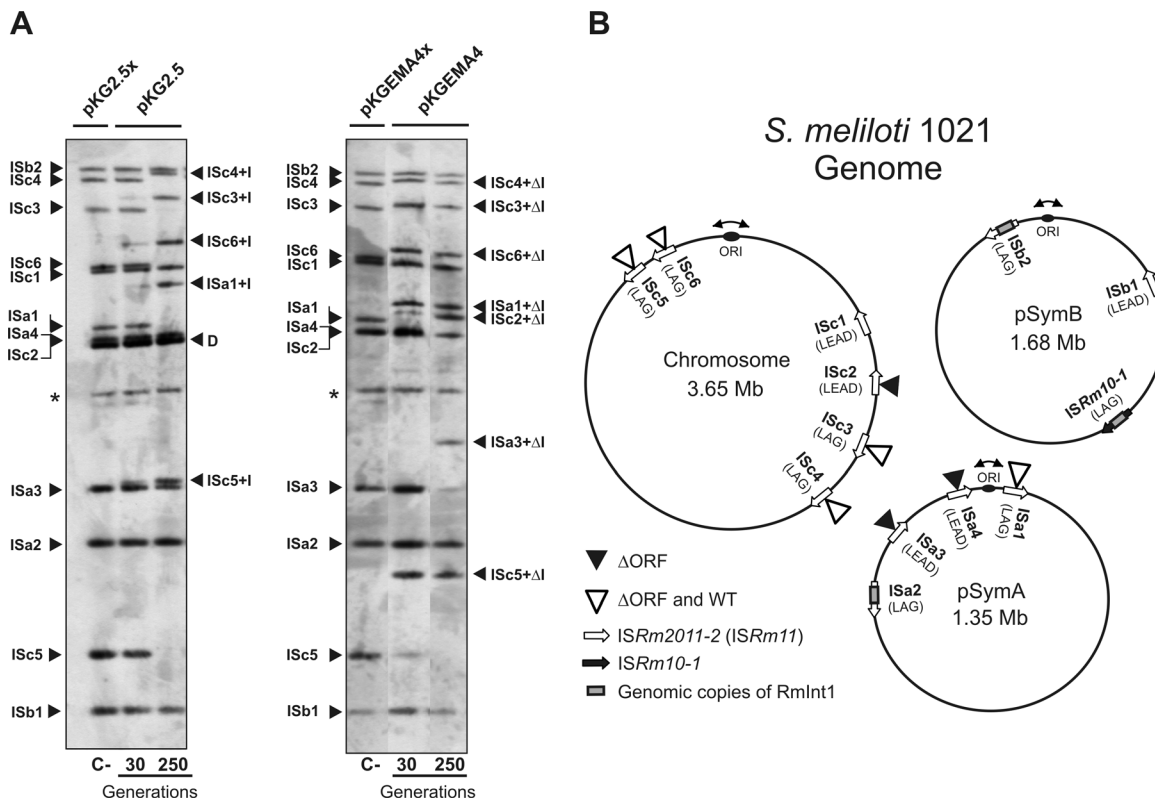


Figure 4. Retrohoming of RmInt1 into the genome of the *S. meliloti* reference strain 1021. (A) Southern hybridizations of *Bam*HI-digested DNA from 1021 transconjugants harboring plasmids pKG2.5 or pKGEMA4 grown for 30 and 250 generations. Membranes were probed with a DNA fragment specific to the 3' end of *ISRm2011-2*. First lane of each panel (C-) shows an *ISRm2011-2* (IS) fingerprint of 1021 derivatives expressing non-functional introns from plasmids pKG2.5x or pKGEMA4x. Insertions of the wild-type intron (IS+I) or the Δ ORF derivative (IS+ Δ I) are indicated. D indicates hybridization signals from pKG2.5 and pKG2.5x DNA and the asterisk (*) a DNA fragment partially digested presumably containing ISb1. (B) Distribution of intron RmInt1 and the newly acquired intron copies (wt or Δ ORF) within the 1021 genome. Schematic shows chromosome, pSymA and pSymB with indication of their replication origin, localization of RmInt1 targets, resident intron copies and new insertions upon acquisition of plasmids pKG2.5 or pKGEMA4 as deduced from the Southern analyses of all the transconjugants according to the *S. meliloti* 1021 genome annotation in (<http://bioinfo.genopole-toulouse.prd.fr/annotation/iANT/bacteria/rhime/>). IS targets correspond to the following annotated 1021 genes: SMC02205 (ISc1), SMC00848 (ISc2), SMC02411 (ISc3), SMC01382 (ISc4), SMC03084 (ISc5), SMC03305 (ISc6), SMA3000 (ISa1), SMA3014 (ISa2), SMA3015 (ISa3), SMA3018 (ISa4), SMB21616 (ISb1), SMB21703-5 (ISb2) and *ISRm10-1* (SMB21709-10).

homing of RmInt1 and presumably for IIB group II intron mobility.

Upon acquisition of the wild-type intron or an IEP+ Δ ORF variant with a marked increase of homing efficiency, the *S. meliloti* genome homing sites begin to be occupied by retrohoming events. This colonization occurs with an extraordinary frequency since these events were detected in 100% of the transconjugants analyzed. We previously reported that the preferred RmInt1 retrohoming pathway is consistent with reverse splicing of the intron RNA into single-stranded DNA at a replication fork, using a nascent DNA lagging-strand as the primer for reverse transcription (28) similar to the ectopic transposition of group II introns (21). The colonization of the homing sites in the *S. meliloti* genome corresponds to this preferred retrohoming pathway and our results predict that the homing sites located on the template for the lagging-strand synthesis are the first to be invaded by the intron, later the available homing sites on the leading strand template will be occupied as well. Interestingly, even though there is at least a proven ectopic site in *S. meliloti* named *oxil* (30) retrotransposition was not detected in our assays. Since this is a rather rare event, retrohoming seems to be the preferred mobility mechanism for intron

dissemination in the *S. meliloti* genome. Once the homing sites are saturated, retrotransposition could contribute to further intron dispersion.

It has been reported that LI.LtrB also retrotransposes into plasmid targets with a bias toward the template for the lagging-strand DNA synthesis even stronger than that for chromosomal events (22). It has been speculated that this bias may be related to the different location of chromosome and plasmids in the cell during replication. We have not observed such a bias for retrohoming of RmInt1 in the *S. meliloti* genome and in fact all the target sites located in the preferred strand for retrohoming into the chromosome were occupied. In *S. meliloti*, the two symbiotic plasmids pSymA and pSymB are considered chromosome-like replicons (32) and therefore they could share the same partitioning properties as the main chromosome being located at similar sites during cell division and therefore equally accessible to intron RNPs. Furthermore, our results do not reveal a preference for integration of the intron into homing sites near the Ori domain as it has been observed for retrotransposition of LI.LtrB (22,23) and retrohoming into *E. coli* genome of gene targeting vectors (targetron) based on LI.LtrB intron (45).

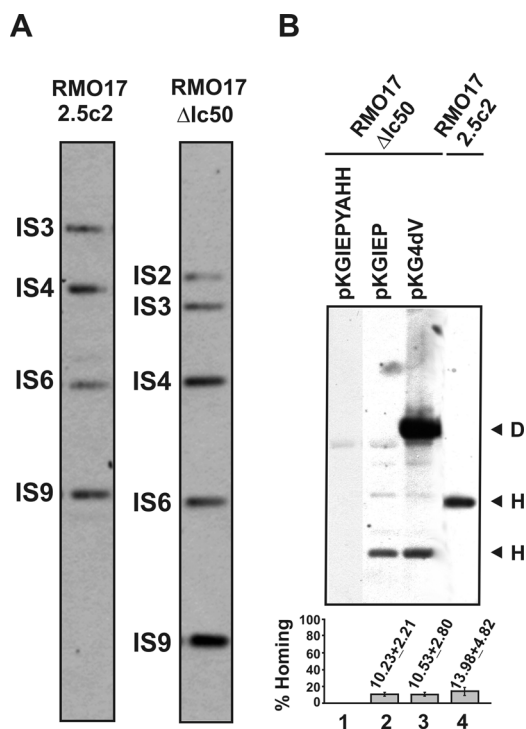


Figure 5. Mobility of the genomic intron copies and *trans* complementation of the Δ ORF derivative with the IEP. (A) Wild-type RmInt1 and Δ ORF profiling in RMO17 derivatives cured from donor plasmids pKG2.5 (RMO17-2.5c2) or pKGEMA4 (RMO17- Δ lc50). Genomic DNA from both strains was SalI-digested and probed with a DNA fragment specific to the 5' end sequence of the RmInt1 ribozyme. (B) *In vivo* homing of the acquired genomic introns. Target plasmid pJB0.6LAG was mobilized into RMO17-2.5c2 and RMO17- Δ lc50 strains. Compatible plasmids pKGI EP or pKGI 4dV expressing a functional IEP were then conjugated into RMO17- Δ lc50 containing pJB0.6LAG to complement the Δ ORF insertions. Homing of wild-type RmInt1 and Δ ORF derivative was assessed on plasmid pools from RMO17-2.5c2 and RMO17- Δ lc50, respectively, as described. pKGI EYPYAHH was used as negative complementation control in the assays. Invasion rates were calculated as described in materials and methods and plotted in the histogram shown below. D, donor plasmid; H, homing product.

We also showed that the retrohomed wild-type introns in the *S.meliloti* genome retain mobility, but not the Δ ORF variant, which imply that the functions of the IEP on intron mobility cannot be substituted by any other factor in this host genetic background. Complementation in *trans* with LtrA has been achieved for a Ll.LtrB Δ ORF derivative, but splicing is <10% of that in *cis* configuration (43). Interestingly, the RmInt1 Δ ORF variant inserted in the genome of *S.meliloti* could be also mobilized by supplying a functional IEP in *trans*, either alone or within the context of a defective ribozyme, displaying an invasion rate on a plasmid-borne target site similar to that exhibited by the wild-type intron. These findings suggest that some mutant introns can still contribute efficiently to intron dispersal by supplying a functional IEP.

From an evolutionary perspective, it has been proposed that RmInt1 has been inherited vertically from a common ancestor in *S.meliloti* and related bacteria and also by independent horizontal-transfer events (31). Presently it is estimated that only 10% of the *S.meliloti* isolates lack the intron and that presumably, in this population there is no restriction to intron homing, i.e. strain RMO17. The extraordinary

spreading ability of RmInt1 and the high number of homing sites available in these isolates predict that the intron might be disseminated into *S.meliloti* population by lateral transfer nowadays, with retrohoming being the main mechanism for the colonization of the recipient genomes.

ACKNOWLEDGEMENTS

This work was funded by the Spanish Ministerio de Educación y Ciencia (project no. BIO2005-02312). The authors are grateful to Vicenta Millán and Ascensión Martos for technical assistance. Funding to pay the Open Access publication charges for this article was provided by Spanish Ministerio de Educación y Ciencia.

Conflict of interest statement. None declared.

REFERENCES

- Michel,F. and Ferat,J.L. (1995) Structure and activities of group II introns. *Annu. Rev. Biochem.*, **64**, 435–461.
- Toro,N. (2003) Bacteria and Archaea group II introns; additional mobile genetic elements in the environment. *Environ. Microbiol.*, **5**, 143–151.
- Lambowitz,A.M. and Zimmerly,S. (2004) Mobile Group II Introns. *Annu. Rev. Genet.*, **38**, 1–35.
- Carignani,G., Groudinsky,O., Frezza,D., Schiavon,E., Bergantino,E. and Slonimski,P.P. (1983) An mRNA maturase is encoded by the first intron of the mitochondrial gene for the subunit I of cytochrome oxidase in *S.cerevisiae*. *Cell*, **35**, 733–742.
- Michel,F., Umesono,K. and Ozeki,H. (1989) Comparative and functional anatomy of group II catalytic introns—a review. *Gene*, **82**, 5–30.
- Moran,J.V., Mecklenburg,K.L., Sass,P., Belcher,S.M., Mahnke,D., Lewin,A. and Perlman,P. (1994) Splicing defective mutants of the *COX1* gene of yeast mitochondrial DNA: initial definition of the maturase domain of the group II intron α 2. *Nucleic Acids Res.*, **22**, 2057–2064.
- Saldanha,R., Chen,B., Wank,H., Matsuura,M., Edwards,J. and Lambowitz,A.M. (1999) RNA and protein catalysis in group II intron splicing and mobility reactions using purified components. *Biochemistry*, **38**, 9069–9083.
- Matsuura,M., Noah,J.W. and Lambowitz,A.M. (2001) Mechanism of maturase-promoted group II intron splicing. *EMBO J.*, **20**, 7259–7270.
- Toor,N., Hausner,G. and Zimmerly,S. (2001) Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA*, **7**, 1142–1152.
- Zimmerly,S., Hausner,G. and Wu,X.-C. (2001) Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res.*, **29**, 1238–1250.
- Toro,N., Molina-Sánchez,M.D. and Fernández-López,M. (2002) Identification and characterization of bacterial class E group II introns. *Gene*, **299**, 245–250.
- Ferat,J.L., Le Gouar,M. and Michel,F. (2003) A group II intron has invaded the genus *Azotobacter* and is inserted within the termination codon of the essential groEL gene. *Mol. Microbiol.*, **49**, 1407–1423.
- Mohr,G., Perlman,P.S. and Lambowitz,A.M. (1993) Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function. *Nucleic Acids Res.*, **21**, 4991–4997.
- Gorbalenya,A.E. (1994) Self-splicing group I and group II introns encode homologous (putative) DNA endonucleases of a new family. *Protein Sci.*, **3**, 1117–1120.
- Shub,D.A., Goodrich-Blair,H. and Eddy,S.R. (1994) Amino acid sequence motif of group I endonucleases is conserved in open reading frames of group II introns. *Trends Biochem. Sci.*, **19**, 402–404.
- San Filippo,J. and Lambowitz,A.M. (2002) Characterization of the C-Terminal DNA-Binding/DNA Endonuclease region of a group II Intron-Encoded Protein. *J. Mol. Biol.*, **324**, 933–951.

17. Lambowitz, A.M., Caprara, M.G., Zimmerly, S. and Perlman, P.S. (1999) Group I and group II ribozymes as RNPs: clues to the past and guides to the future. In Gesteland, R.F., Cech, T.R. and Atkins, J.F. (eds), *The RNA World, 2nd edn*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 451–485.
18. Belfort, M., Derbyshire, V., Parker, M.M., Cousineau, B. and Lambowitz, A.M. (2002) Mobile introns: pathways and proteins. In Craig, N.L., Craigie, R., Gellert, M. and Lambowitz, A.M. (eds), *Mobile DNA, 2nd edn*. ASM Press, Washington DC, pp. 761–783.
19. Martínez-Abarca, F. and Toro, N. (2000) Group II introns in the bacterial world. *Mol. Microbiol.*, **38**, 917–926.
20. Dickson, L., Huang, H.-R., Liu, L., Matsuura, M., Lambowitz, A.M. and Perlman, P.S. (2001) Retrotransposition of a yeast group II intron occurs by reverse splicing directly into ectopic DNA sites. *Proc. Natl Acad. Sci. USA*, **98**, 13207–13212.
21. Ichiyangi, K., Beauregard, A., Lawrence, S., Smith, D., Cousineau, B. and Belfort, M. (2002) Retrotransposition of the Ll.LtrB group II intron proceeds predominantly via reverse splicing into DNA targets. *Mol. Microbiol.*, **46**, 1259–1272.
22. Ichiyangi, K., Beauregard, A. and Belfort, M. (2003) A bacterial group II intron favors retrotransposition into plasmid targets. *Proc. Natl Acad. Sci. USA*, **100**, 15742–15747.
23. Coros, C.J., Landthaler, M., Piazza, C.L., Beauregard, A., Esposito, D., Perutka, J., Lambowitz, A.M. and Belfort, M. (2005) Retrotransposition strategies of the *Lactococcus lactis* Ll.LtrB group II intron are dictated by host identity and cellular environment. *Mol. Microbiol.*, **56**, 509–524.
24. 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.
25. Belhocine, K., Yam, K.K. and Cousineau, B. (2005) Conjugative transfer of the *Lactococcus lactis* chromosomal sex factor promotes dissemination of the Ll.LtrB group II intron. *J. Bacteriol.*, **187**, 930–939.
26. Chen, Y., Klein, J.R., McKay, L.L. and Dunny, G.M. (2005) Quantitative analysis of group II intron expression and splicing in *Lactococcus lactis*. *Appl. Environ. Microbiol.*, **71**, 2576–2586.
27. Martínez-Abarca, F., García-Rodríguez, F.M. and Toro, N. (2000) Homing of a bacterial group II intron with an intron-encoded protein lacking a recognizable endonuclease domain. *Mol. Microbiol.*, **35**, 1405–1412.
28. Martínez-Abarca, F., Barrientos-Durán, A., Fernández-López, M. and Toro, N. (2004) The RmInt1 group II intron has two different retrohoming pathways for mobility using predominantly the nascent lagging strand at DNA replication forks for priming. *Nucleic Acids Res.*, **32**, 2880–2888.
29. Martínez-Abarca, F., Zekrí, S. and Toro, N. (1998) Characterization and splicing *in vivo* of a *Sinorhizobium meliloti* group II intron associated with particular insertion sequences of the IS630-Tc1/IS3 retroposon superfamily. *Mol. Microbiol.*, **28**, 1295–1306.
30. Muñoz, E., Villadas, P.J. and Toro, N. (2001) Ectopic transposition of a group II intron in natural bacterial populations. *Mol. Microbiol.*, **41**, 645–652.
31. Fernández-López, M., Muñoz-Adelantado, E., Gillis, M., Willems, A. and Toro, N. (2005) Dispersal and evolution of the *Sinorhizobium meliloti* group II RmInt1 intron in bacteria that interact with plants. *Mol. Biol. Evol.*, **22**, 1518–1528.
32. Galibert, F., Finan, T.M., Long, S.R., Pühler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M.J., Becker, A., Boistard, P. *et al.* (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science*, **293**, 668–672.
33. Villadas, P.J., Velázquez, E., Martínez-Molina, E. and Toro, N. (1995) Identification of nodule-dominant *Rhizobium meliloti* strains carrying pRmeGR4b-type plasmid within indigenous soil populations by PCR using primers derived from specific DNA sequences. *FEMS Microbiol. Ecol.*, **17**, 161–168.
34. Robertsen, B.K., Aiman, P., Darvill, A.G., McNeil, M. and Albersheim, P. (1981) The structure of acidic extracellular polysaccharides secreted by *Rhizobium leguminosarum* and *Rhizobium trifolii*. *Plant Physiol.*, **67**, 389–400.
35. Costa, M., Michel, F., Molina-Sánchez, M.D., Martínez-Abarca, F. and Toro, N. (2006a) An alternative intron-exon pairing scheme implied by unexpected *in vitro* activities of group II intron RmInt1 from *Sinorhizobium meliloti*. *Biochimie*, **88**, 711–717.
36. Jiménez-Zurdo, J.I., García-Rodríguez, F.M., Barrientos-Durán, A. and Toro, N. (2003) DNA-target requirement for homing *in vivo* of a bacterial group II intron encoding a protein lacking the DNA endonuclease domain. *J. Mol. Biol.*, **326**, 413–423.
37. Muñoz-Adelantado, E., San Filippo, J., Martínez-Abarca, F., García-Rodríguez, F.M., Lambowitz, A.M. and Toro, N. (2003) Mobility of the *Sinorhizobium meliloti* group II intron RmInt1 occurs by reverse splicing into DNA, but requires an unknown reverse transcriptase priming mechanism. *J. Mol. Biol.*, **327**, 931–943.
38. Martínez-Abarca, F. and Toro, N. (2000) RecA-independent ectopic transposition *in vivo* of a bacterial group II intron. *Nucleic Acids Res.*, **28**, 4397–4402.
39. Costa, M., Michel, F. and Toro, N. (2006) Potential for alternative intron-exon pairings in group II intron RmInt1 from *Sinorhizobium meliloti* and its relatives. *RNA*, **12**, 338–341.
40. Molina-Sánchez, M.D., Martínez-Abarca, F. and Toro, N. (2006) Excision of the *Sinorhizobium meliloti* group II intron RmInt1 as circles *in vivo*. *J. Biol. Chem.*, **281**, 28737–28744.
41. Guo, H., Karberg, M., Long, M., Jones, J.P., III, Sullenger, B. and Lambowitz, A.M. (2000) Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. *Science*, **289**, 452–457.
42. Plante, I. and Cousineau, B. (2006) Restriction for gene insertion within the *Lactococcus lactis* Ll.LtrB group II intron. *RNA*, **12**, 1–13.
43. Cui, X., Matsuura, M., Wang, Q., Ma, H. and Lambowitz, A.M. (2004) A group II intron-encoded maturase functions preferentially *in cis* and requires both the reverse transcriptase and X domains to promote RNA splicing. *J. Mol. Biol.*, **340**, 211–231.
44. Conlan, L.H., Stanger, M.J., Ichiyangi, K. and Belfort, M. (2005) Localization, mobility and fidelity of retrotransposed group II introns in rRNA genes. *Nucleic Acids Res.*, **33**, 5262–5270.
45. Zhong, J., Karberg, M. and Lambowitz, A.M. (2003) Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Res.*, **31**, 1656–1664.
46. Yao, J., Zhong, J. and Lambowitz, A.M. (2005) Gene targeting using randomly inserted group II introns (targetrons) recovered from an *Escherichia coli* gene disruption library. *Nucleic Acids Res.*, **33**, 3351–3362.