



Contribution of Mobile Group II Introns to *Sinorhizobium meliloti* Genome Evolution

Nicolás Toro*, Francisco Martínez-Abarca, María D. Molina-Sánchez, Fernando M. García-Rodríguez and Rafael Nisa-Martínez

Structure, Dynamics and Function of Rhizobacterial Genomes (Grupo de Ecología Genética de la Rizosfera), Department of Soil Microbiology and Symbiotic Systems, Consejo Superior de Investigaciones Científicas, Estación Experimental del Zaidín, Granada, Spain

OPEN ACCESS

Edited by:

Marco Bazzicalupo,
Università degli Studi di Firenze, Italy

Reviewed by:

Trevor Carlos Charles,
University of Waterloo, Canada

Julie Ardley,
Murdoch University, Australia
Antonio Lagares,
UNLP-CONICET, Argentina

*Correspondence:

Nicolás Toro
nicolas.toro@eez.csic.es

Specialty section:

This article was submitted to
Microbial Symbioses,
a section of the journal
Frontiers in Microbiology

Received: 16 January 2018

Accepted: 19 March 2018

Published: 04 April 2018

Citation:

Toro N, Martínez-Abarca F,
Molina-Sánchez MD,
García-Rodríguez FM and
Nisa-Martínez R (2018) Contribution
of Mobile Group II Introns
to *Sinorhizobium meliloti* Genome
Evolution. *Front. Microbiol.* 9:627.
doi: 10.3389/fmicb.2018.00627

Mobile group II introns are ribozymes and retroelements that probably originate from bacteria. *Sinorhizobium meliloti*, the nitrogen-fixing endosymbiont of legumes of genus *Medicago*, harbors a large number of these retroelements. One of these elements, Rmlnt1, has been particularly successful at colonizing this multipartite genome. Many studies have improved our understanding of Rmlnt1 and phylogenetically related group II introns, their mobility mechanisms, spread and dynamics within *S. meliloti* and closely related species. Although Rmlnt1 conserves the ancient retroelement behavior, its evolutionary history suggests that this group II intron has played a role in the short- and long-term evolution of the *S. meliloti* genome. We will discuss its proposed role in genome evolution by controlling the spread and coexistence of potentially harmful mobile genetic elements, by ectopic transposition to different genetic loci as a source of early genomic variation and by generating sequence variation after a very slow degradation process, through intron remnants that may have continued to evolve, contributing to bacterial speciation.

Keywords: catalytic RNAs, genome evolution, intron, insertion sequences, retroelements, rhizobia, reverse transcriptase, ribozymes

INTRODUCTION

Rhizobia are soil bacteria that elicit root nodules on leguminous plants (Oldroyd et al., 2011; Haag et al., 2013), in which they transform nitrogen (N₂) from the atmosphere into ammonia. *Sinorhizobium meliloti* and *Sinorhizobium medicae* are closely related species that form nitrogen-fixing symbioses on *Medicago* plants. These two bacterial species have a large composite multipartite genome with a chromosome (~3.65 Mb) and two large symbiotic (Sym) megaplasmids of ~1.3 and ~1.6 Mb in size, and some strains also have smaller accessory plasmids (Barloy-Hubler et al., 2000; Barnett et al., 2001; Capela et al., 2001; Finan et al., 2001; Galibert et al., 2001; Reeve et al., 2010). These species are found in the soil, the rhizosphere, and within root nodules. The different ecological niches colonized by these bacteria, the boundaries of gene flow between these microhabitats and plant selection have shaped the evolution of these bacterial species.

Whole-genome sequence analyses have been performed for *S. meliloti* and *S. medicae* (Bailly et al., 2011; Galardini et al., 2011, 2013; Epstein et al., 2012; Sugawara et al., 2013), and some studies have indicated that the pSymB megaplasmid plays a key role in intraspecies differentiation (Galardini et al., 2013, 2015). Whole-genome sequencing of isolates from a natural *S. meliloti* population (GR4-type) isolated from *M. sativa* (alfalfa) root nodules revealed the existence of genotypic variation underlying early genetic and ecological differentiation (Toro et al., 2016, 2017). In addition to the small numbers of single-nucleotide polymorphisms uniformly distributed over the multipartite genome, the isolates analyzed exhibited other types of genomic variation, indels of various sizes, some recombination events including the excision and acquisition of genomic islands, and the transposition of mobile elements, which seem to be the early microevolutionary forces influencing this multipartite genome. Mobile genetic elements have caused genetic variation in the three main replicons, and the variation due to one group of mobile elements, group II introns, is particularly interesting (Lambowitz and Zimmerly, 2011; Novikova and Belfort, 2017; Toro et al., 2017). These mobile introns are self-splicing RNAs and retroelements, which are highly abundant in *S. meliloti* and, to a lesser extent, in closely related species. A number of studies on the mobile group II intron RmInt1 (Martínez-Abarca et al., 1998) harbored by most *S. meliloti* strains (Muñoz et al., 2001; Fernández-López et al., 2005; Molina-Sánchez and Toro, 2015) have provided clues to the dynamics and behavior of these retroelements and their possible contribution to the short- and long-term evolutionary events shaping these bacterial genomes.

In this review, we provide a brief overview of what is presently known about group II introns, including RmInt1 in particular, in the context of their possible contribution to *S. meliloti* genome evolution.

MAIN FEATURES OF GROUP II INTRONS AND RmInt1

Group II introns were originally identified in the mitochondrial and chloroplast genomes of lower eukaryotes and plants (Michel et al., 1989), and subsequently identified in bacteria (Ferat and Michel, 1993), and in some archaeal species (Dai and Zimmerly, 2003; Toro, 2003).

Despite the fact that the nuclear genomes of eukaryotes do not contain group II introns, the nuclear spliceosome and spliceosomal introns are thought to have evolved from mobile group II introns (Koonin, 2006). Moreover, group II introns have also been identified as the probable evolutionary ancestors of non-LTR retrotransposons, telomerase, and retroviruses in eukaryotes (Lambowitz and Belfort, 2015). Group II introns have been maintained in bacteria because they home preferentially to sites outside of functional genes, within intergenic regions or in mobile genetic elements behaving predominantly as retroelements (Simon et al., 2008; Chillón et al., 2011; Nisa-Martínez et al., 2013). Typical group II introns comprises

six typical stem-loop domains (DI–DVI), with an internal open reading frame (ORF) encoding an RT-maturase within DIV (**Figure 1A**). Exon-binding recognition sequences (EBS) have been identified in DI, and the most important catalytic residues are present in DV, whereas the branch point adenosine, the nucleophile responsible for initiating the splicing reaction resulting in intron lariat formation, is located in DVI. The intron-encoded protein (IEP) is a multidomain protein consisting of a reverse transcriptase (RT) domain followed by a maturase (X) domain. The RT domain resembles to the finger and palm subdomains of a polymerase, and the X domain may be equivalent to a polymerase thumb domain. Some IEPs have an endonuclease domain (EN) involved in intron mobility at their C-terminus (Belfort et al., 2002; Lambowitz and Zimmerly, 2004; Toro et al., 2007). The intron lariat (Costa et al., 2016), the RT domain without the thumb domain (Zhao and Pyle, 2016), and a full-length (Stamos et al., 2017) group II intron RT have been crystallized. The cryo-EM structure of a group II intron and its reverse transcriptase has also been reported (Qu et al., 2016).

The non-coding portion of group II intron RNAs has coevolved with their IEPs, and on the basis of conserved RNA structure, group II introns have been classified in three main classes (IIA, IIB, and IIC) (for a review see Lambowitz and Zimmerly, 2011). On the basis of their IEPs they are classified in several classes and varieties: A, B, C, D, E, F, G (*g1*), CL1 (chloroplast-like 1), CL2 (chloroplast-like 2), and ML (mitochondrion-like) (Simon et al., 2008; Toro and Martínez-Abarca, 2013).

Group II intron splicing involves two sequential transesterification reactions (de Lencastre et al., 2005; Marcia et al., 2013; Kruschel et al., 2014; Zhao and Pyle, 2017). The first reaction involves a nucleophilic attack on the 5' splice site by the 2'-OH of a bulged adenosine residue (bulging A) located in DVI, releasing the 5' exon and generating an intron-3' exon branched intermediate. The second step involves a nucleophilic attack on the 3' splice site by the free 3'-OH of the 5'-exon, yielding the ligated exons and the intron lariat RNA with a 2'-5' phosphodiester bond. The intron binding sequences (IBS) at the 5' and 3' splice sites are recognized by base-pairing with the intron RNA exon-binding sequences (EBS). The EBS1–IBS1 interaction is essential for 5' splice site recognition during the splicing reaction whereas the EBS2–IBS2 pairing is dispensable for splicing (Barrientos-Durán et al., 2011).

The mechanism of group II intron mobility was initially established (Zimmerly et al., 1995) for yeast introns (aI1 and aI2) and for the *Lactococcus lactis* LLtrB intron. The mobility of these introns is mediated by a target DNA-primed reverse transcription mechanism (TPRT endonuclease-dependent) involving a ribonucleoprotein (RNP) complex containing both the intron RNA and the IEP (Lambowitz et al., 1999). Once the target sequence (homing site) has been recognized (20–35 bp) mainly by the RNA component of the RNP complex, through base pairing between the EBS and IBS sequences, retrohoming occurs by TPRT. The intron

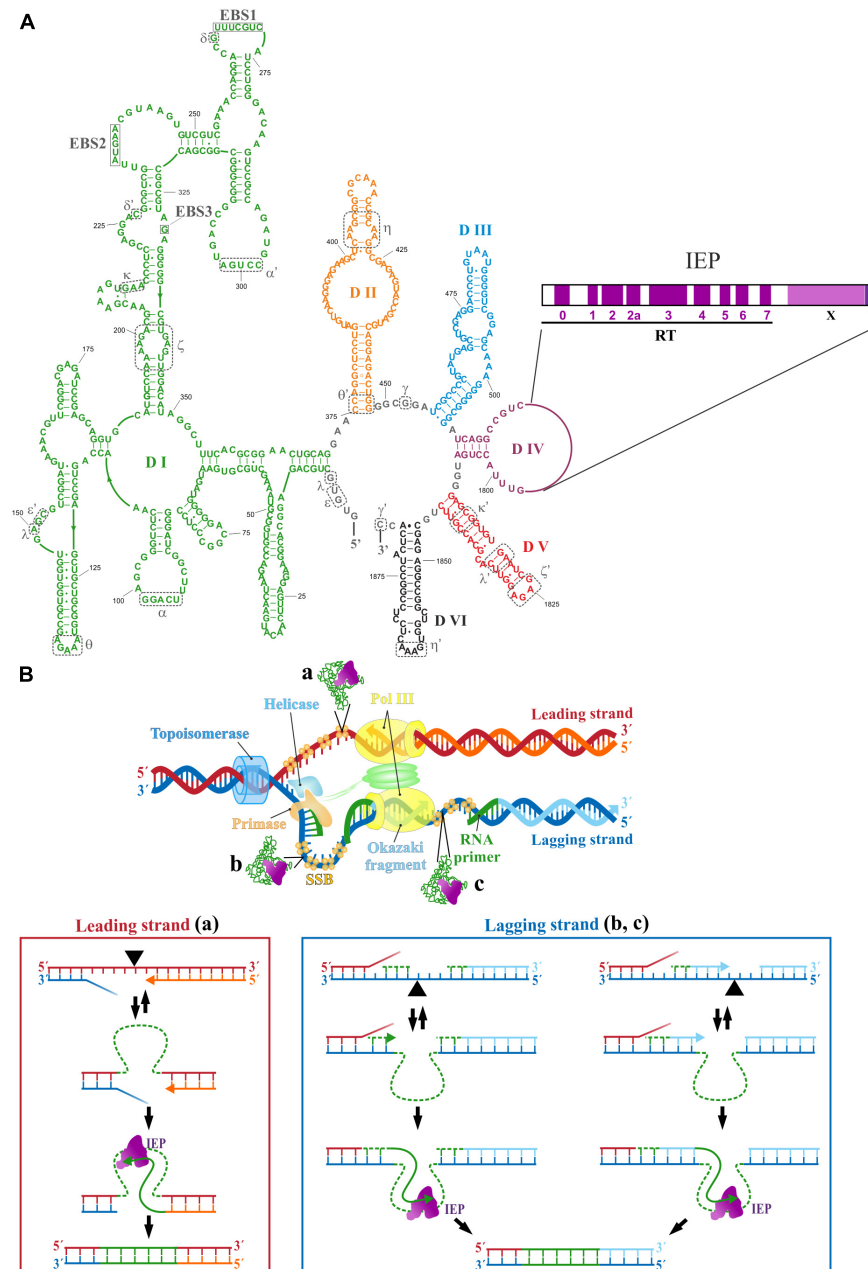


FIGURE 1 | (A) The secondary structure of the RmInt1 ribozyme is shown. Intron RNA domains (DI–DVI) are shown in different colors. Nucleotides identify as exon binding sites 1, 2, and 3 (EBSs) in DI are boxed. Nucleotides involved in tertiary contacts are indicated by dashed gray lines and Greek letters. The intron-encoded protein (IEP) domains encoded by DIV are also shown. RT: reverse transcriptase domain with conserved RT sequence blocks (0–7); X: maturase domain; *C-terminal tail. **(B)** A schematic diagram of the replication fork with intron insertion in the template for leading strand (a), and lagging strand (b,c) is shown. The process of intron integration is represented below. Once the intron RNA reverse splices into the target DNA, the IEP retrotranscribes the RNA into cDNA, using the nascent leading strand (a), the 3' end of the RNA primers synthesized by the primase (b), or the nascent Okazaki fragments (c) as the primer. Green dashed lines represent the RNA molecules. The black triangle indicates the intron insertion site.

RNA cleaves the sense strand at the exon junctions, and integrates into the target site. At the same time, the IEP cleaves the antisense strand through its EN activity (for the Ll.ltrB, this cleavage occurs at position +9). The 3' end of the antisense strand is used by the RT domain of the IEP as the primer site for reverse transcription of the inserted

intron RNA. The intron cDNA is then integrated into the host target site by homologous recombination-independent repair mechanisms. Some mobile group II introns, such as the *S. meliloti* RmInt1 intron (IIB3/class D), have an IEP with no En domain. The C-terminal maturase region (C-tail) of this intron is responsible for its maturase and

DNA-insertion functions (Molina-Sánchez et al., 2010). The RmInt1 IEP has RT activity, and its RNP can mediate reverse splicing into target sites at double- or single-stranded DNA substrates but cannot achieve site-specific second-strand cleavage (Muñoz-Adelantado et al., 2003). An alternative to the TPRT retrohoming pathway is therefore required. RmInt1 retrohomes very efficiently (Martínez-Abarca et al., 2000; Martínez-Abarca and Toro, 2000a,b; Nisa-Martínez et al., 2007), and its mobility is dependent on the intron RNA and the IEP, recognizing a target site extending 20 nt into the 5' exon and 5 nt into the 3' exon (Jiménez-Zurdo et al., 2003). The preferred retrohoming pathway of RmInt1 involves the reverse splicing of the intron RNA into single-stranded DNA at replication forks, with a bias toward the template for lagging strand synthesis (Martínez-Abarca et al., 2004), probably through the use of the RNA primers synthesized by the primase or the Okazaki fragments to prime reverse transcription (Figure 1B, down, right side).

RmInt1 AND CLOSELY RELATED INTRONS IN *S. meliloti* CONTROL THE SPREAD OF POTENTIALLY HARMFUL MOBILE GENETIC ELEMENTS

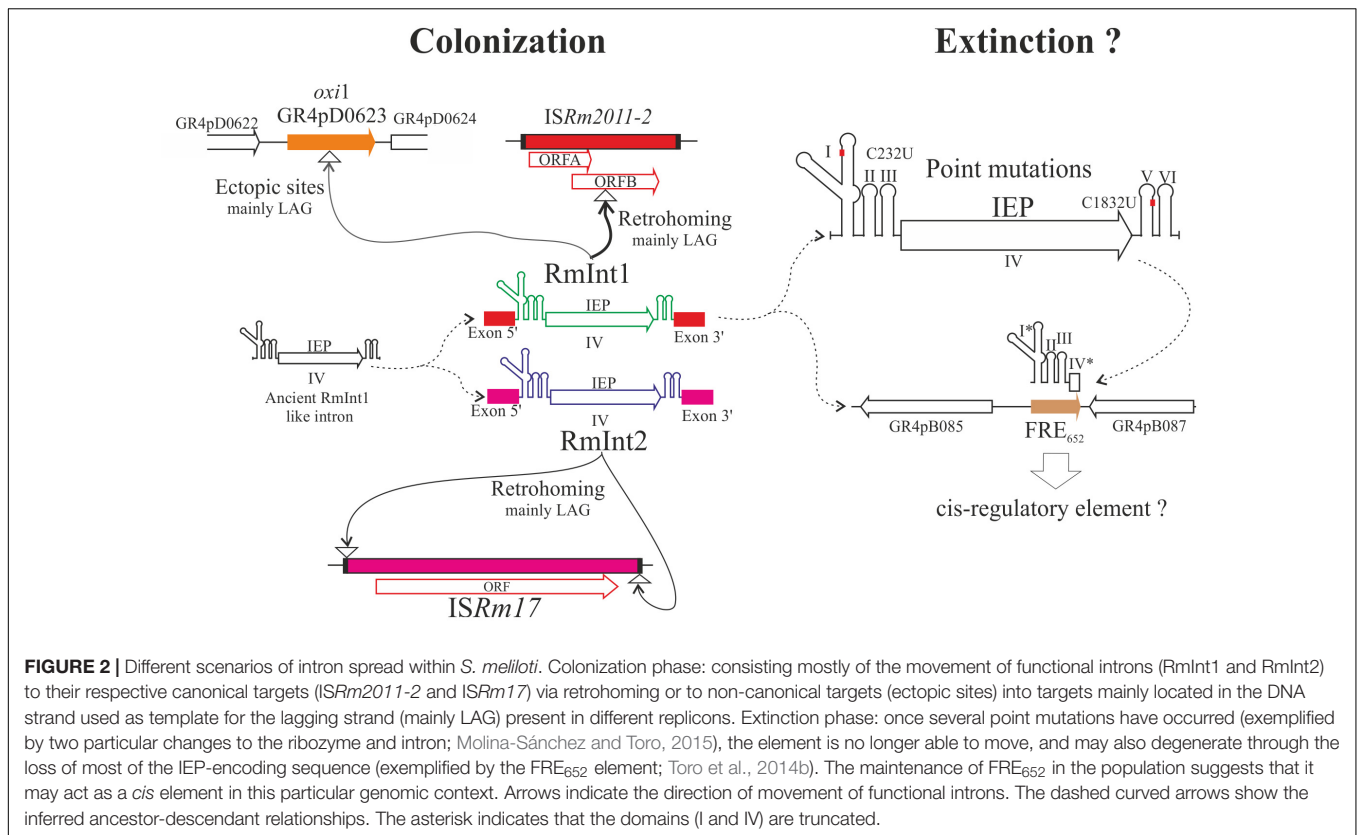
The *S. meliloti* genome harbors many different mobile genetic elements, including group II introns from classes C, D, E, and G (Toro et al., 2002, 2003; Candales et al., 2012; Toro and Martínez-Abarca, 2013). RmInt1 is widespread in *S. meliloti* (Muñoz et al., 2001) and was first described in the GR4 strain (Martínez-Abarca et al., 1998), which contains 10 copies of RmInt1 (Martínez-Abarca et al., 2013) distributed between the different replicons [four copies on the chromosome, five copies on pRmeGR4c (pSymA), and one copy on the accessory plasmid pRmeGR4b]. This bacterial strain also harbors a closely related mobile intron called RmInt2 (Martínez-Rodríguez et al., 2014). RmInt2 also belongs to class D, it has a nucleotide sequence 72% identical to that of RmInt1, and it is present in seven copies (four copies on pSymA and three on pSymB). Other group II introns are present as single copies in this bacterial genome: a class C intron (pSymB IEP: WP_015243078.1), a class G intron (chromosome IEP: WP_015241266.1), and a disrupted class E intron (*S.me* I4) (Toro et al., 2002) located on the accessory plasmid pRmeGR4b. It is thought that 90% of *S. meliloti* isolates harbor RmInt1 intron, with copy numbers differing between *S. meliloti* strains. Full-length close relatives of RmInt1 (85–99% nucleotide identity) are also present in *S. medicae*, *E. adhaerens* and *S. terangae* (closest relatives of *S. meliloti*), whereas closely related fragmented introns have been identified in these bacterial species and in other *Sinorhizobium* and *Rhizobium* species (Fernández-López et al., 2005). These RmInt1-like elements were likely acquired by vertical inheritance from a common ancestor and through independent lateral transfer events.

The natural target site of RmInt1 lies within *ISRm2011-2* (*ISRm11*) an insertion sequence from the IS630 family (Figure 2;

Selbitschka et al., 1995), and this intron is also found in other closely related insertion elements (Toro et al., 2003; Biondi et al., 2011). *ISRm2011-2* is present virtually in all *S. meliloti* strains, usually at high copy numbers (4–13; Biondi et al., 2011). In its natural host, RmInt1 has a very low splicing ability *in vivo*; insertion of the intron into *ISRm2011-2* therefore blocks its spread (Chillón et al., 2011). Group II introns are highly flexible in terms of their ability to colonize the *S. meliloti* genome (Nisa-Martínez et al., 2007). For example, RmInt2 has a lower retrohoming efficiency than RmInt1, and a greater probability of survival and spread in the genome through the relaxation of target-site specificity using the left and right inverted repeats of *ISRm17* as DNA targets (Figure 2), thereby ensuring that at least one of its targets is on the template of the lagging strand during DNA replication (Martínez-Rodríguez et al., 2014). The spread, gain, and losses of these IS elements not carrying adaptive traits likely impose costs on their hosts in the short term, particularly when their copy number increases, generating a source of genetic instability and a burden for the cell replication machinery (Rankin et al., 2011). The mobile introns that use these elements as a target not only help the host cell to reduce such fitness costs by limiting their ability to spread within the genome, but also contribute to their coexistence with the host genome, allowing time for selective forces to exert their effects on both the mobile elements and the host genome. Thus, group II introns in bacteria, and RmInt1 in particular, may be considered key elements in the short- and long-term evolution of the *S. meliloti* genome.

ECTOPIC TRANSPOSITION OF RmInt1 AS A SOURCE OF EARLY GENOMIC VARIATION IN *S. meliloti*

As indicated above, group II introns can insert into intronless alleles at intron-specific locations (homing), and at novel (ectopic) sites, albeit at low frequency, and this has been identified as an important mechanism of intron spread in bacterial populations (Muñoz et al., 2001). Initial studies on the GR4-type *S. meliloti* population showed that RmInt1 was able to invade the ectopic *oxi1* site. The intron is able to recognize this target site because this gene carries IBS sequences similar to those of *ISRm2011-2*. Following the sequencing of the genome of GR4 (Martínez-Abarca et al., 2013), *oxi1* was identified in pRmeGR4d (pSymB) as a gene encoding an oxidoreductase (WP_010975806.1), resulting from an ectopic transposition event that has occurred independently several times in the natural population (Toro et al., 2016). In the GR4-type population analyzed, 4.1% of the isolates displayed such intron insertions, but this variant is fixed in 95% of EM2-type isolates, a different population obtained from nodules also occupied by GR4-type strains, suggesting that it may be subject to selective pressure (Muñoz et al., 2001). Moreover, whole-genome sequencing in GR4-type isolates identified another ectopic site, an intron insertion



in a gene encoding a NAD(P)-dependent oxidoreductase (WP_010969452.1), corresponding to an event that probably occurred about 1700 years ago (Toro et al., 2016, 2017). The biological significance of the targeting of these genes by *RmInt1* in *S. meliloti* is unknown. The characterization of these loci, together with genome-wide sequence analysis of the EM2-type population, will therefore provide insight into the importance of the ectopic transposition of group II introns for early genome evolution in natural bacterial populations.

THE DYNAMICS OF *RmInt1* IN *S. meliloti*

The *S. meliloti* strains lacking *RmInt1*, such as RMO17 (Toro et al., 2014a), have been shown to remain suitable for intron colonization. The arrival of the intron is followed by its spread through high-frequency retrohoming (using an RNA intermediate) with a strand bias related to the DNA replication (Nisa-Martínez et al., 2007). Like other genetic mobile elements, group II introns appear to follow a gain-loss cycle that would account for the frequency and copy number of *RmInt1* in *S. meliloti* species (Molina-Sánchez and Toro, 2015). It has been suggested that bacteria with highly colonized genomes by a particular intron are removed from the population by purifying selection in genome-wide selective sweeps (Leclercq and Cordaux, 2012). This elimination probably occurs due to the genomic instability resulting from possible recombination

between intron copies, leading to the deletion of parts of the genome. For *RmInt1*, gradual eradication (Figure 2) begins with specific mutations of the intron ribozyme RNA and the IEP, inactivating intron splicing and retrohoming (Molina-Sánchez and Toro, 2015). This inactivation is followed by further fragmentation, with the loss of the intron 3'-end, including the IEP (Dai and Zimmerly, 2002; Fernández-López et al., 2005), which seems to be a general process underlying the loss of function of bacterial group II introns. This gradual process of inactivation may still result in a transitory active mobile element rescued by either the intron RNA or IEPs from other active copies present or arriving to the same cell. Nonetheless, this process is unlikely to remove intron sequences completely (Figure 2). Instead, some intron fragments are likely to remain buried within the genome. For example, a 652 nt *RmInt1*-like fragment known as *FRE₆₅₂*, which is 88.9% identical to *RmInt1* and 89.7% identical to the *S. medicae* intron Sr.md.II, has persisted over long periods of evolutionary time in the genomes of *S. meliloti* and *S. medicae* (Toro et al., 2014b). This intron fragment covers ribozyme DI–DIII and includes part of DIV, truncated at a position corresponding to position 653 of *RmInt1*. *FRE₆₅₂* is located on the accessory plasmid pRmeGR4b of strain GR4, whereas it is carried by pSymA and orthologous plasmids in other *S. meliloti* and *S. medicae* strains, respectively. *FRE₆₅₂* is located close to a predicted helix–turn–helix transcriptional regulator and a diguanylate cyclase/phosphodiesterase (DGC/PDEA) gene, and is followed by a pectate lyase and a carbonate dehydratase. All these genes are absent from

S. meliloti strains 1021 and Rm41. It therefore seems likely that FRE₆₅₂ and the surrounding loci are involved in recombination events within the various replicons of the multipartite genome of these bacterial species and that this element may even spread within populations, through recombination, without causing genome-wide selective sweeps. The significance and impact of fragmented introns within a particular genome remains uncertain, but we suggest that FRE₆₅₂ and other intron fragments have persisted and are continuing to evolve in the genome perhaps by contributing to the environmental adaptation and symbiotic capacity of these rhizobial species, perhaps as *cis*-regulatory elements (Figure 2).

CONCLUSION

The presence of active mobile introns, such as RmInt1 and close relatives, in the *S. meliloti* genome constitutes an interesting model for exploring the dynamics and possible influence of these retroelements in bacterial genome evolution. We suggest that the persistence of these self-splicing catalytic RNAs in

bacteria, such as RmInt1 in *S. meliloti*, results from their possible involvement in short- and long-term mechanisms underlying genome evolution. We therefore consider these apparently “parasitic” genetic elements as a source of direct and indirect genetic variation contributing to genome evolution and ecological differentiation in natural bacterial populations.

AUTHOR CONTRIBUTIONS

NT prepared the manuscript. MM-S and FG-R prepared the Figure 1, and FM-A prepared the Figure 2. All authors critically reviewed the manuscript.

FUNDING

This work was supported by research grants BIO2014-51953-P and BIO2017-82244-P of the *Plan Nacional de I+D+i*, Biotechnology Program from the Spanish *Ministerio de Economía, Industria y Competitividad* including ERDF (European Regional Development Funds).

REFERENCES

- Bailey, X., Giuntini, E., Sexton, M. C., Lower, R. P., Harrison, P. W., Kumar, N., et al. (2011). Population genomics of *Sinorhizobium medicae* based on low-coverage sequencing of sympatric isolates. *ISME J.* 5, 1722–1734. doi: 10.1038/ismej.2011.55
- Barloy-Hubler, F., Capela, D., Batut, J., and Galibert, F. (2000). High-resolution physical map of the pSymB megaplasmid and comparison of the three replicons of *Sinorhizobium meliloti* strain 1021. *Curr. Microbiol.* 41, 109–113. doi: 10.1007/s002840010103
- Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., et al. (2001). Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proc. Natl. Acad. Sci. U.S.A.* 98, 9883–9888. doi: 10.1073/pnas.161294798
- Barrientos-Durán, A., Chillón, I., Martínez-Abarca, F., and Toro, N. (2011). Exon sequence requirements for excision *in vivo* of the bacterial group II intron RmInt1. *BMC Mol. Biol.* 12:24. doi: 10.1186/1471-2199-12-24
- Belfort, M., Derbyshire, V., Parker, M. M., Cousineau, B., and Lambowitz, A. M. (2002). “Mobile introns: pathways and proteins,” in *Mobile DNA*, eds N. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (Washington DC: ASM Press), 761–783. doi: 10.1128/9781555817954.ch31
- Biondi, E. G., Toro, N., Bazzicalupo, M., and Martínez-Abarca, F. (2011). Spread of the group II intron RmInt1 and its insertion sequence target sites in the plant endosymbiont *Sinorhizobium meliloti*. *Mob. Genet. Elements* 1, 2–7. doi: 10.4161/mge.1.1.15316
- Candales, M. A., Duong, A., Hood, K. S., Li, T., Neufeld, R. A., Sun, R., et al. (2012). Database for bacterial group II introns. *Nucleic Acids Res.* 40, 187–190. doi: 10.1093/nar/gkr1043
- Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., et al. (2001). Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proc. Natl. Acad. Sci. U.S.A.* 98, 9877–9882. doi: 10.1073/pnas.161294398
- Chillón, I., Martínez-Abarca, F., and Toro, N. (2011). Splicing of the *Sinorhizobium meliloti* RmInt1 group II intron provides evidence of retroelement behavior. *Nucleic Acids Res.* 39, 1095–1104. doi: 10.1093/nar/gkq847
- Costa, M., Walbott, H., Monachello, D., Westhof, E., and Michel, F. (2016). Crystal structures of a group II intron lariat primed for reverse splicing. *Science* 354:aaf9258. doi: 10.1126/science.aaf9258
- Dai, L., and Zimmerly, S. (2002). Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res.* 30, 1091–1102. doi: 10.1093/nar/30.5.1091
- Dai, L., and Zimmerly, S. (2003). ORF-less and reverse-transcriptase-encoding group II introns in archaeobacteria, with a pattern of homing into related group II introns ORFs. *RNA* 8, 1294–1307. doi: 10.1261/rna.2126203
- de Lencastre, A., Hamill, S., and Pyle, A. M. (2005). A single active-site region for a group II intron. *Nat. Struct. Mol. Biol.* 12, 626–627. doi: 10.1038/nsmb957
- Epstein, B., Branca, A., Mudge, J., Bharti, A. K., Briskine, R., Farmer, A. D., et al. (2012). Population genomics of the facultatively mutualistic bacteria *Sinorhizobium meliloti* and *S. medicae*. *PLoS Genet.* 8:e1002868. doi: 10.1371/journal.pgen.1002868
- Ferat, J. L., and Michel, F. (1993). Group II self-splicing introns in bacteria. *Nature* 364, 358–361. doi: 10.1038/364358a0
- 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. doi: 10.1093/molbev/msi144
- Finan, T. M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorhölter, F. J., et al. (2001). The complete sequence of the 1,683-kb pSymB megaplasmid from the N₂-fixing endosymbiont *Sinorhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* 98, 9889–9894. doi: 10.1073/pnas.161294698
- Galardini, M., Brilli, M., Spini, G., Rossi, M., Roncaglia, B., Bani, A., et al. (2015). Evolution of intra-specific regulatory networks in a multipartite bacterial genome. *PLoS Comput. Biol.* 11:e1004478. doi: 10.1371/journal.pcbi.1004478
- Galardini, M., Mengoni, A., Brilli, M., Pini, F., Fioravanti, A., Lucas, S., et al. (2011). Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. *BMC Genomics* 12:235. doi: 10.1186/1471-2164-12-235
- Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G., and Mengoni, A. (2013). Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. *Genome Biol. Evol.* 5, 542–558. doi: 10.1093/gbe/evt027
- Galibert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., et al. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293, 668–672. doi: 10.1126/science.1060966
- Haag, F. A., Arnold, M. F., Myka, K. K., Kerscher, B., Dall’Angelo, S., Zanda, M., et al. (2013). Molecular insights into bacteroid development during *Rhizobium*-legume symbiosis. *FEMS Microbiol. Rev.* 37, 364–383. doi: 10.1111/1574-6976.12003
- Jiménez-Zurdo, J. I., García-Rodríguez, F. M., Barrientos-Durán, A., and Toro, N. (2003). DNA target site requirements for homing *in vivo* of a bacterial group II

- intron encoding a protein lacking the DNA endonuclease domain. *J. Mol. Biol.* 326, 413–423. doi: 10.1016/S0022-2836(02)01380-3
- Koonin, E. V. (2006). The origins of introns and their role in eukaryogenesis: a compromise solution to introns-early versus introns-late debate. *Biol. Direct* 1:22.
- Kruschel, D., Skilandat, M., and Sigel, R. K. (2014). NMR structure of the 5' splice site in the group IIB intron Sc.a15 γ -conformational requirements for exon-intron recognition. *RNA* 20, 295–307. doi: 10.1261/rna.041137.113
- Lambowitz, A. M., and Belfort, M. (2015). Mobile bacterial group II introns at the crux of eukaryotic evolution. *Microbiol. Spectr.* 3:MDNA3-0050-2014. doi: 10.1128/microbiolspec.MDNA3-0050-2014
- 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 *The RNA World*, eds R. F. Gesteland, T. R. Cech, and J. F. Atkins (New York, NY: Cold Spring Harbor), 451–485. doi: 10.1101/087969589.37.451
- Lambowitz, A. M., and Zimmerly, S. (2004). Mobile group II introns. *Ann. Rev. Genet.* 38, 1–35. doi: 10.1146/annurev.genet.38.072902.091600
- Lambowitz, A. M., and Zimmerly, S. (2011). Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb. Perspect. Biol.* 3:a003616. doi: 10.1101/cshperspect.a003616
- Leclercq, S., and Cordaux, R. (2012). Selection-driven extinction dynamics for group II introns in Enterobacteriales. *PLoS One* 7:e252268. doi: 10.1371/journal.pone.0052268
- Marcia, M., Somarowthu, S., and Pyle, A. M. (2013). Now on display: a gallery of group II intron structures at different stages of catalysis. *Mob. DNA* 4:14. doi: 10.1186/1759-8753-4-14
- 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. doi: 10.1093/nar/gkh616
- 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. doi: 10.1046/j.1365-2958.2000.01804
- Martínez-Abarca, F., Martínez-Rodríguez, L., López-Contreras, J. A., Jiménez-Zurdo, J. I., and Toro, N. (2013). Complete genome sequence of the alfalfa symbiont *Sinorhizobium/Ensifer meliloti* strain GR4. *Genome Announc.* 1:e00174-12. doi: 10.1128/genomeA.00174-12
- Martínez-Abarca, F., and Toro, N. (2000a). Group II introns in the bacterial world. *Mol. Microbiol.* 38, 917–926. doi: 10.1046/j.1365-2958.2000.02197
- Martínez-Abarca, F., and Toro, N. (2000b). RecA-independent ectopic transposition in vivo of a bacterial group II intron. *Nucleic Acids Res.* 28, 4397–4402. doi: 10.1093/nar/28.21.4397
- Martínez-Abarca, F., Zekri, 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. doi: 10.1046/j.1365-2958.1998.00894
- Martínez-Rodríguez, L., García-Rodríguez, F. M., Molina-Sánchez, M. D., Toro, N., and Martínez-Abarca, F. (2014). Insights into the strategies used by related group II introns to adapt successfully for the colonization of a bacterial genome. *RNA Biol.* 11, 1061–1071. doi: 10.4161/rna.32092
- Michel, F., Umesomo, K., and Ozeki, H. (1989). Comparative and functional anatomy of group II catalytic introns—a review. *Gene* 82, 5–30. doi: 10.1016/0378-1119(89)90026-7
- Molina-Sánchez, M. D., Martínez-Abarca, F., and Toro, N. (2010). Structural features in the C-terminal region of the *Sinorhizobium meliloti* RmInt1 group II intron-encoded protein contribute to its maturase and intron DNA-insertion function. *FEBS J.* 277, 244–254. doi: 10.1111/j.1742-4658.2009.07478
- Molina-Sánchez, M. D., and Toro, N. (2015). Inactivation of group II intron RmInt1 in the *Sinorhizobium meliloti* genome. *Sci. Rep.* 5:12036. doi: 10.1038/srep12036
- 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. doi: 10.1046/j.1365-2958.2001.02540
- 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. doi: 10.1016/S0022-2836(03)00208-0
- Nisa-Martínez, R., Jiménez-Zurdo, J. I., Martínez-Abarca, F., Muñoz-Adelantado, E., and Toro, N. (2007). Dispersion of the RmInt1 group II intron in the *Sinorhizobium meliloti* genome upon acquisition by conjugative transfer. *Nucleic Acids Res.* 35, 214–222. doi: 10.1093/nar/gkl1072
- Nisa-Martínez, R., Laporte, P., Jiménez-Zurdo, J. I., Frugier, F., Crespi, M., and Toro, N. (2013). Localization of a bacterial group II intron-encoded protein in eukaryotic nuclear splicing-related cell compartments. *PLoS One* 8:e84056. doi: 10.1371/journal.pone.0084056
- Novikova, O., and Belfort, M. (2017). Mobile group II introns as ancestral eukaryotic elements. *Trends Genet.* 33, 773–783. doi: 10.1016/j.tig.2017.07.009
- Oldroyd, G. E., Murray, J. D., Poole, P. S., and Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* 45, 119–144. doi: 10.1146/annurev-genet-110410-132549
- Qu, G., Kaushal, P. S., Wang, J., Shigematsu, H., Piazza, C. L., Agrawa, R. K., et al. (2016). Structure of a group II intron in complex with its reverse transcriptase. *Nat. Struct. Mol. Biol.* 23, 549–557. doi: 10.1038/nsmb.3220
- Rankin, D. J., Rocha, E. P. C., and Brown, S. P. (2011). What traits are carried on mobile genetic elements, and why? *Heredity* 106, 1–10. doi: 10.1038/hdy.2010.24
- Reeve, W., Chain, P., O'Hara, G., Ardley, J., Nandesana, K., Bräu, L., et al. (2010). Complete genome sequence of the Medicago microsymbiont Ensifer (*Sinorhizobium*) medicae strain WSM419. *Stand. Genomics Sci.* 2, 77–86. doi: 10.4056/signs.43526
- Selbitschka, W., Arnold, W., Jording, D., Kosier, B., Toro, N., and Pühler, A. (1995). The insertion sequence element ISRm2011-2 belongs to the IS630-Tc1 family of transposable elements and is abundant in *Rhizobium meliloti*. *Gene* 163, 59–64. doi: 10.1016/0378-1119(95)00371-C
- Simon, D. M., Clarke, N. A., McNeil, B. A., Johnson, I., Pantuso, D., Dai, L., et al. (2008). Group II introns in eubacteria and archaea: ORF-less introns and new varieties. *RNA* 14, 1704–1713. doi: 10.1261/rna.1056108
- Stamos, J. L., Lentzsch, A. M., and Lambowitz, A. M. (2017). Structure of a thermostable group II intron reverse transcriptase with template-primer and its functional and evolutionary implications. *Mol. Cell* 68, 926–939. doi: 10.1016/j.molcel.2017.10.024
- Sugawara, M., Epstein, B., Badgley, B. D., Unno, T., Xu, L., Reese, J., et al. (2013). Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. *Genome Biol.* 14:R17. doi: 10.1186/gb-2013-14-2-r17
- Toro, N. (2003). Bacteria and archaea group II introns: additional mobile elements in the environment. *Environ. Microbiol.* 5, 143–151. doi: 10.1046/j.1462-2920.2003.00398
- Toro, N., Jiménez-Zurdo, J. I., and García-Rodríguez, F. M. (2007). Bacterial group II introns: not just splicing. *FEMS Microbiol. Rev.* 31, 342–358. doi: 10.1111/j.1574-6976.2007.00068
- Toro, N., and Martínez-Abarca, F. (2013). Comprehensive phylogenetic analysis of bacterial group II intron-encoded ORFs lacking the DNA endonuclease domain reveals new varieties. *PLoS One* 8:e55102. doi: 10.1371/journal.pone.0055102
- Toro, N., Martínez-Abarca, F., and Fernández-López, M. (2016). The early events underlying genome evolution in a localized *Sinorhizobium meliloti* population. *BMC Genomics* 17:556. doi: 10.1186/s12864-016-2878-9
- Toro, N., Martínez-Abarca, F., Fernández-López, M., and Muñoz-Adelantado, E. (2003). Diversity of group II introns in the genome of *Sinorhizobium meliloti* strain 1021: splicing and mobility of RmInt1. *Mol. Gen. Genomics* 268, 628–636. doi: 10.1007/s00438-002-0778-y
- Toro, N., Martínez-Abarca, F., and Nisa-Martínez, R. (2014a). Complete genome sequence of the RmInt1 group II intronless *Sinorhizobium meliloti* strain RMO17. *Genome Announc.* 2:e1001-14. doi: 10.1128/genomeA.01001-14
- Toro, N., Martínez-Rodríguez, L., and Martínez-Abarca, F. (2014b). Insights into the history of a bacterial group II intron remnant from the genomes of the nitrogen-fixing symbionts *Sinorhizobium meliloti* and *Sinorhizobium medicae*. *Heredity* 113, 306–315. doi: 10.1038/hdy.2014.32
- 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. doi: 10.1016/S0378-1119(02)01079
- Toro, N., Villadas, P. J., Molina-Sánchez, M. D., Navarro-Gómez, P., Vinardell, J. M., Cuesta-Berrio, L., et al. (2017). The underlying process of early ecological

- and genetic differentiation in a facultative mutualistic *Sinorhizobium meliloti* population. *Sci. Rep.* 7:675. doi: 10.1038/s41598-017-00730-7
- Zhao, C., and Pyle, A. M. (2016). Crystal structures of a group II intron maturase reveal a missing link in spliceosome evolution. *Nat. Struct. Mol. Biol.* 23, 558–565. doi: 10.1038/nsmb.3224
- Zhao, C., and Pyle, A. M. (2017). Structural insights into the mechanism of group II intron splicing. *Trends Biochem. Sci.* 42, 470–482. doi: 10.1016/j.tibs.2017.03.007
- Zimmerly, S., Guo, H., Perlman, P. S., and Lambowitz, A. M. (1995). Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* 82, 545–554. doi: 10.1016/0092-8674(95)90027-6

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Toro, Martínez-Abarca, Molina-Sánchez, García-Rodríguez and Nisa-Martínez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.