# Exploring the Evolution of Group II Introns Using Ll.LtrB from *Lactococcus lactis* as a Model System

**Doctoral Thesis**

**Z. Kamila Belhocine**

**Department of Microbiology and Immunology**
**McGill University, Montreal**

**November 2007**

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of *Philosophiæ Doctorate*

*Your file Votre référence*
*ISBN: 978-0-494-51486-3*
*Our file Notre référence*
*ISBN: 978-0-494-51486-3*

# Table of contents

# Abstract

Group II introns are self-splicing introns found in bacterial, archaeal and organellar genomes. Some group II introns are also mobile retroelements that are capable of invading DNA using an RNA intermediate. Group II introns are the subject of several evolutionary hypotheses. It was suggested that these mobile elements are able to spread via lateral (horizontal) transfer. The association of many bacterial group II introns with other mobile elements further strengthened this theory. Moreover, group II introns are hypothesized to be the progenitors of nuclear spliceosome-dependant introns, as well as the five small nuclear RNAs (snRNAs) part of the spliceosome. The ability of some fragmented organellar group II introns to undergo splicing *in trans* supports this evolutionary theory.

We used the Ll.LtrB group II intron from the Gram-positive bacterium *Lactococcus lactis* as a model system to explore the evolutionary hypotheses described above. Ll.LtrB was the first bacterial group II intron shown to splice and invade new DNA sites *in vivo*. Ll.LtrB is found on three conjugative elements of *L. lactis*: two conjugative plasmids and the chromosomal sex factor. In these three elements, Ll.LtrB interrupts the gene coding for relaxase, an enzyme that initiates the intercellular transfer of its host conjugative element.

We demonstrated that Ll.LtrB can be disseminated between *L. lactis* strains and to another Gram-positive bacterium, *Enterococcus faecalis*, by the

conjugative transfer of its host elements. The Ll.LtrB intron can invade its recognition site in the recipient cells upon its conjugative transfer. Moreover, the intron can invade ectopic sites in the chromosome of the host cell. This work constituted the first experimental demonstration of the spread of a group II intron by the transfer of its host mobile element.

We developed an Ll.LtrB-based model system to study *trans*-splicing of group II introns, a phenomenon poorly studied and understood. We demonstrated that Ll.LtrB is capable of *trans*-splicing in *L. lactis* when independently fragmented at several locations that mimic natural fragmentation sites of group II introns. Subsequently, we used a Tn5 transposon-based genetic screen to identify other sites where Ll.LtrB could be fragmented and maintain its ability to *trans*-splice. The selected fragmentation sites clustered between segments that are structurally and/or functionally analogous to snRNAs. Therefore, our work supports the long-standing evolutionary theory that links spliceosomal RNAs and nuclear intervening sequences to group II introns.

We successfully addressed evolutionary questions surrounding group II introns using experimental approaches. Our work contributes to understanding the evolution of group II introns and provides support to long-standing theories involving these self-splicing retroelements.

# Résumé

Les introns de groupe II sont des introns autocatalytiques que l'on trouve dans les génomes d'organelles, de bactéries et d'archaea. Certains introns de groupe II sont aussi des éléments rétromobiles capables d'envahir des sites d'ADN à l'aide d'un intermédiaire ARN. Les introns de groupe II sont au centre de plusieurs théories évolutives. Ces éléments mobiles sont considérés extrêmement sujets au transfert latéral (horizontal). Cette théorie est aussi supportée par le fait que les introns de groupe II sont souvent associés à d'autres éléments mobiles dans les génomes bactériens. De plus, les introns de groupe II sont considérés comme les ancêtres des introns nucléaires eucaryotes, ainsi que des 5 petits ARN nucléaires (snRNAs) qui font partie du spliceosome. Cette hypothèse évolutive est supportée par la découverte d'introns de groupe II fragmentés qui s'épissent en *trans* dans certaines organelles.

Nous avons utilisé l'intron de groupe II Ll.LtrB de la bactérie Gram-positive *Lactococcus lactis* comme modèle pour explorer certaines de ces théories évolutives. Ll.LtrB fût le premier intron bactérien de groupe II pour lequel l'épissage et la mobilité ont été démontrés *in vivo*. Ll.LtrB réside dans trois éléments conjugatifs de *L. lactis* : deux plasmides conjugatifs et un facteur sexuel chromosomique. Dans ces trois éléments, Ll.LtrB interrompt le gène codant pour une relaxase, enzyme qui initie le transfert intercellulaire de son élément hôte.

Nous avons démontré que Ll.LtrB peut être disséminé entre diverses souches de *L. lactis* et vers une autre bactérie Gram-positive, *Enterococcus faecalis*, par le transfert de ses éléments conjugatifs hôtes. Ll.LtrB peut envahir son site de reconnaissance dans les cellules receveuses lors de son transfert conjugatif. Il peut aussi envahir des sites hétérologues sur le chromosome de son nouvel hôte. Notre travail constitue la première démonstration expérimentale de la dissémination d'un intron de groupe II par son élément mobile hôte.

Nous avons développé un système modèle basé sur Ll.LtrB afin d'étudier l'épissage en *trans* des introns de groupe II, un phénomène peu étudié et compris. Nous avons démontré que Ll.LtrB peut s'épisser en *trans* dans *L. lactis* s'il est indépendamment fragmenté à différents points correspondant aux fragmentations naturelles de certains introns d'organelles. Subséquemment, nous avons utilisé un essai génétique basé sur le transposon Tn5 afin de déterminer tous les sites potentiels auxquels un intron de groupe II peut être fragmenté tout en maintenant sa capacité à s'épisser en *trans*. La majorité des sites de fragmentation isolés se situait entre des régions structurellement et/ou fonctionnellement analogues aux petits ARN du spliceosome. Notre travail procure donc un nouveau support expérimental à la théorie évolutive qui lie les introns de groupe II aux introns nucléaires et leur machinerie d'épissage.

Nous avons investigué des théories évolutives concernant les introns de groupe II à l'aide d'approches expérimentales. Notre travail apporte une contribution significative à la compréhension de l'évolution des introns de groupe II et fournit un support expérimental aux théories évolutives qui entourent ces éléments autocatalytiques et rétromobiles.

# Acknowledgements

These last five years were a fantastic experience for me and I am extremely grateful to all persons who contributed to the wonderful adventure that it has been. My warmest thanks go to my supervisor, Professor Benoit Cousineau, who dared hiring a freshly graduated physicist in his lab who had no clue what the 5' end of DNA was, and barely knew that an intron was some weird RNA that ... how do you say again, splice?... Ben, I am profoundly indebted for your guidance, your patience and your constant availability, but above all, for your trust in me and your encouragements. It has been a fantastic collaboration and I am very proud to have contributed to the journey of the Cousineau lab.

I also extend my thanks to the four great honour's students that I had the extreme privilege to train and that did a wonderful job, investing their time, energy and enthusiasm into their projects. Karen, Bonnie, Vickie and Anthony, it has been a real pleasure working with you and I am extremely proud to have known and trained each one of you. We had wonderful collaborations and I am convinced that you will all do very well. My best wishes accompany each one of you...

I would also like to thank the members of my advisory committee, Dr. Greg Marczynski and Dr. Hervé LeMoual for those long committee meetings

where I would tend to get excited and talk for hours about my projects. Thank you for your precious time, advice and input into our work.

Moreover, I wish to thank all past and present members of the Cousineau lab, with special thoughts for Karen and Isabelle and their arachnophobia and those Friday evening conversations where every other sane person would have left the Duff building, and Karine and her wicked sense of humour. I also thank all people who transited by our lab and made the place a wonderful environment to work, discuss about very important topics such as sales at the Eaton center, laugh and fight about whose turn it is for lab meetings and who left the balance so dirty... I also thank Martin, a genuine Cousineau lab member, for always interrupting my conversations with Ben.

I want to thank all my friends both in and outside of the Duff building, with special thoughts for Lei, my "Chinese twin sister", for always being there and for your support at different moments of my thesis during all these years. I also thank all people that contribute to make Thomson House THE place to be on Friday nights... I won't list names otherwise this thesis will really be way too long...

I thank Benoit, Anthony, Will Spencer, and Yassine for their precious help in editing some parts of my thesis.

Furthermore, I would have never achieved any of all this without the tremendous support of my family. I thank my wonderful parents for always pushing us to surpass ourselves and accept no limits, and for providing us with the financial and moral support to achieve our goals. I also thank my brothers, Nazim and Lyès, for being two fantastic gentlemen always taking care of their

"favourite" (and unique) sister, and Laura for her kindness. Special thoughts also for my two outstanding roommates, my "guys": my brother and my husband, who made home a fantastic place to live in during these five years!! Special thanks for Lyès for always finding the right joke or the right tune for both joyful and difficult moments (especially the joyful ones...).

Finally and most importantly I thank my husband Yassine, my soul mate, for unconditional support and encouragements during all these years, for always believing in me, for learning that group II introns are both splicing and mobile elements and being able to hold a discussion about lateral transfer, not bad for a physicist!! Yassine, this thesis is dedicated to you...

# LIST OF FIGURES

# Contributions of Authors

**Chapter Two**

K.B. conducted the conjugation/retrotransposition assay and characterized the chromosomal intron insertion sites. K.B. also performed many repeats of conjugation to obtain consistent data. I. P. performed various repeats for conjugation efficiencies. B.C. designed all experiments and conducted all experiments involving intron retrohoming. B.C. wrote the article and K.B. participated in manuscript editing.

**Chapter Three**

K. B. and B.C. designed the experiments. K.B. performed most of the experiments described in the article. K.K.Y. performed various repeats for conjugation efficiencies. B.C. wrote the article and K.B. participated in manuscript editing.

**Chapter Four**

K.B. and B.C. designed the experiments. K.B. generated the marked sex factor, performed some of the described experiments and mentored B.Y. and V.M. B.Y. characterized intraspecies transfer of the marked sex factor and subsequent

intron retrohoming. V.M. generated the pool of sex factor co-integrates and performed the majority of interspecies conjugations. V.M., K.B. and B.C. wrote the article.

**Chapter Five**

K.B. and B.C. designed the experiments. K.B. built the whole system and performed most of the experiments. A.B.M. performed various repeats for conjugation. A.B.M. also conducted all experiments for the section of implication of the protein in trans-splicing. B.C. and K.B. wrote the article and A.B.M. participated in editing the manuscript.

**Chapter Six**

B.C. designed the screen. K.B. built the Tn5 transposon and performed the screens. K.B. and A.B.M. identified the Tn5 insertion sites and analyzed the conjugation efficiency of independent *trans*-splicers. K.B. performed the RT-PCR assays. K.B. and B.C. wrote the article and A.B.M. participated in editing the manuscript.

# CHAPTER ONE

## Literature review and objective of the thesis

In 1977, it was discovered that some DNA coding sequences were interrupted by non-coding segments that were processed prior to translation (Berget *et al.*, 1977). These intervening sequences termed introns were removed, or spliced, from the RNA precursor leaving the coding portions of the gene, or exons, ligated. By 1983, all known introns could be classified into four groups: nuclear introns, tRNA introns, organellar group I and group II introns. In the latter two classes, introns were sorted based on the observed conservation of their secondary structures (Michel *et al.*, 1982; Michel and Dujon, 1983). It became evident that the splicing mechanism of each intron class was fundamentally different, with only group II and nuclear introns sharing identical splicing pathways. The latter observation prompted the hypothesis of an evolutionary relationship between group II and nuclear introns (Sharp, 1985; Cech, 1986).

Group II introns are highly structured catalytic RNAs or ribozymes. They are capable to achieve their own splicing from RNA transcripts. In that respect, they are the largest ribozymes observed in nature (Michel and Ferat, 1995; Lehmann and Schmidt, 2003). The self-splicing feature of group II introns was

simultaneously discovered in 1986 by van der Veen and co-workers and Peebles and co-workers. *In vitro* splicing of the yeast mitochondrial aI5γ group II intron from the cytochrome *c* oxidase subunit I (*cox I*) was reported, independent of any cellular factors. Van der Veen and co-workers showed that splicing resulted in the formation of a lariat molecule branched at an A residue located eight nucleotides upstream of the 3' splice site (van der Veen *et al.*, 1986). Moreover, Peebles and co-workers showed that splicing produced an RNA with a circular portion branched by a linkage that differs from a 3' – 5' phosphodiester bond (Peebles *et al.*, 1986).

The idiosyncratic distribution of group II introns in nature prompted the hypothesis that these intervening sequences were potentially mobile genetic elements able to invade new sites (Lambowitz, 1989). The presence of an open reading frame (ORF) related to reverse transcriptase genes within the known group II introns further supported this hypothesis (Michel and Lang, 1985). *In vivo* mobility of group II introns was first reported in 1990 with the yeast mitochondrial intron *coxI*-I1 (Meunier *et al.*, 1990; Skelly *et al.*, 1991). The field of group II introns expanded with the advance of molecular biology and the development of new *in vitro* and *in vivo* experimental systems to study these unique genetic elements. The following literature review highlights the multiple facets of group II introns.

## 1.1 Distribution and classification of group II introns

Group II introns consist of a ~ 600 bp self-splicing RNA structure that may contain an ORF of approximately 2 kb (Michel and Ferat, 1995; Zimmerly *et al.*, 2001). Group II introns were initially reported only in organelles of lower eukaryotes and higher plants (Michel *et al.*, 1989). Based on structural differences, they were classified into two major subclasses, IIA and IIB, with further subgroups, A1 and A2, B1 and B2 (Michel *et al.*, 1989; Michel and Ferat, 1995). In organelles, group II introns often interrupt essential genes at conserved positions, underlying their ability to invade homologous sites. The majority of organellar group II introns do not harbour ORFs; these introns rely on cellular-encoded factors to act as maturases to stabilize their RNA structure and assist splicing (see section 1.2) (Michel and Ferat, 1995).

A systematic PCR-based search in the putative bacterial ancestors of organelles, proteobacteria (mitochondria) and cyanobacteria (chloroplasts), revealed the existence of group II introns in bacteria (Ferat and Michel, 1993). As a consequence of growing genome sequencing projects, more than a hundred introns have been identified to date in both Gram-negative and Gram-positive bacteria (Dai *et al.*, 2003). The vast majority of bacterial group II introns harbour an ORF; however, this observation may be biased by the fact that intron detection is primarily achieved through sequence identification of reverse-transcriptase motifs (Dai and Zimmerly, 2002a). Group II intron distribution in bacteria differs considerably from their distribution in organelles. Bacterial group II introns are often found in non-coding regions, within other mobile elements and are not associated with conserved genes (Dai and Zimmerly, 2002a). An additional

subclass of introns, subgroup IIC introns, was identified in bacteria. Introns from this class are typically located downstream of transcriptional terminator motifs (Toor *et al.*, 2001; Dai and Zimmerly, 2002a; Dai *et al.*, 2003). Finally, several ORF-containing and ORF-less group II introns were discovered in two related archaeal species, *Methanosarcina acetivorans* and *M. mazei*. These introns are believed to have been acquired from a bacterial host through horizontal gene transfer (see section 1.6.2.2) (Dai and Zimmerly, 2003; Toro, 2003).

## 1.2 Group II intron splicing pathway

Splicing proceeds via two consecutive transesterification reactions. The first reaction is initiated by a bulged nucleotide, most often an adenosine, located near the 3' end of the intron and known as the branch point. The 2'-OH group of the branch point nucleotide performs a nucleophilic attack on the 5' splice site. This results in the release of exon 1 and the branching of the 5' end of the intron to the branch point nucleotide via a 2'–5' linkage (Figure 1.1, step 1). Subsequently, the 3'-OH group of the released exon 1 performs a second nucleophilic attack on the 3' splice site, resulting in the ligation of the two exons and the release of the intron in the form of a lariat (Figure 1.1, step 2). This splicing pathway is identical to the removal of eukaryotic nuclear introns, suggesting common evolutionary origins between the two classes of introns (see section 1.5.3.1). The splicing reaction is fully RNA-catalyzed and only requires correct folding of the intron ribozyme. *In vitro*, folding can only be achieved at non-physiological conditions. *In vivo*, group II introns require the assistance of maturases to stabilize their correct tertiary conformation. These proteins function

as RNA chaperones only and do not participate in the catalysis. Maturases can be either intron-encoded (see section 1.4.1) or *trans*-acting host factors (Michel and Ferat, 1995). These host factors typically consist in RNA chaperones, such as tRNA synthetases and DEAD-box proteins, required in other cellular processes involving correctly folded RNA. Whether these proteins function by unwinding RNA to resolve kinetic traps or by stabilizing intramolecular arrangements is still under debate (Del Campo M. *et al.*, 2007).

Alternatively, the first transesterification reaction can be initiated by a water molecule, resulting in the release of exon 1 and a linear intron + exon 2 molecule. The second step is still initiated by the 3'-OH of exon 1. Splicing via the hydrolytic pathway releases ligated exons and a linear intron (Lehmann and Schmidt, 2003).

Notably, a group II intron-containing gene may become fragmented following genome rearrangements, resulting in its expression into two fragments: a first fragment harbouring exon 1 and the 5' portion of the intron, and a second fragment harbouring the 3' part of the intron and exon 2. In organelles, these intron fragments may re-associate and catalyze the splicing reaction *in trans* (Figure 1.1B) (Bonen, 1993). *Trans*-splicing thus ligates exons initially expressed on separate transcripts. This phenomenon has only been reported in organelles which demonstrate considerable genome plasticity (Fauron *et al.*, 1995; Knoop, 2004), and where the number of identified group II introns exceeds that found in bacteria and archaea (Dai *et al.*, 2003; Fedorova and Zingler, 2007; Lang *et al.*, 2007).

## 1.3 Folding and architectural organization of group II introns

Proper folding and compaction of group II introns are essential for the onset of splicing. The majority of group II introns fold into a characteristic secondary structure consisting of six helical domains radiating from a central hub (Figure 1.2) (Michel *et al.*, 1989). Furthermore, these domains interact through several long-range tertiary interactions (Figure 1.2, pairs of Greek letters). Most tertiary interactions were originally identified by phylogenetic analyses, and were later experimentally confirmed.

### 1.3.1 Domain I – folding as a scaffold and recognizing the exons

Domain I, the largest catalytic domain of group II introns, is divided into four subdomains (Ia to Id, Figure 1.2). It harbours critical elements for exon recognition and splice site selection, as well as essential sites of tertiary interactions with other domains. Domain I is an independently folding unit that provides a scaffold for the tertiary assembly of the remaining domains to form the catalytic core (Qin and Pyle, 1997; Su *et al.*, 2005).

Subdomain Id typically displays two sequences termed Exon Binding Sites (EBS) 1 and 2, complementary to Intron Binding Sites (IBS) 1 and 2 located at the 3' end of exon 1 (Figure 1.2) (Jacquier and Michel, 1987). The 10-13 base-pairing interaction between EBS and IBS takes place after assembly of the active site (Costa and Michel, 1999). These interactions ensure proper recognition of exon 1 and the 5' splice site; they also play a critical role in substrate recognition during intron mobility (see section 1.4.2). Subgroup IIC introns lack the EBS2 sequence and rely on the EBS1-IBS1 interaction only for exon 1 recognition. As

previously mentioned, these introns are often found downstream of transcriptional termi-nators. The terminator's stem-loop motif is believed to participate in defining the 5' splice site, thus compensating for the missing EBS2-IBS2 interaction (Toor *et al.*, 2006). Subdomain Id also contains elements responsible for exon 2 recognition. For subgroup IIA introns, a single nucleotide-sequence termed δ lies immediately upstream of EBS1 and interacts with the δ' sequence on the 5' end of exon 2 (Figure 1.2) (Jacquier and Jacquesson-Breuleux, 1991). For subgroups IIB and IIC introns, exon 2 recognition is achieved through an EBS3-IBS3 interaction, with EBS3 located in subdomain Id downstream of EBS1, and IBS3 at the 5' end of exon 2 (Figure 1.2) (Costa *et al.*, 2000).

Domain I harbours multiple sequences involved in intra- and inter-domain tertiary interactions (Figure 1.2, pairs of Greek letters) (Lehmann and Schmidt, 2003; Fedorova and Zingler, 2007). A base-pairing interaction links the ε sequence located at nucleotides 3 and 4 of the intron to the ε' sequence contained in a conserved internal loop of domain Ic (Jacquier and Michel, 1990). Along with the EBS-IBS interactions, this intron-intron base-pairing interaction contributes to 5' splice site recognition. The six base-pair α – α' interaction is highly conserved phylogenetically (Jacquier and Michel, 1987). In addition to its involvement in the first transesterification reaction of the splicing pathway (Harris-Kerr *et al.*, 1993), it participates in intron folding and compaction (Waldsich and Pyle, 2007). The β – β' interaction is less conserved phylogenetically and apart from participating in intron folding and compaction (Waldsich and Pyle, 2007), its role remains unclear (Michel and Ferat, 1995; Fedorova and

Zingler, 2007). Subgroup IIB introns exhibit an additional interaction, δ – δ', which likely facilitates base-pairing between the intron and exon 1 (Lehmann and Schmidt, 2003).

Furthermore, interaction of domain I with domain V is believed to form the minimal catalytic core (Koch *et al.*, 1992; Qin and Pyle, 1998). Domain coupling is achieved through three essential long range interactions: ζ – ζ' (Costa and Michel, 1995) and κ – κ' (Boudvillain and Pyle, 1998), which are essential for domain V anchoring and intron compaction, and λ – λ' (Boudvillain *et al.*, 2000), which creates the framework for catalysis by positioning domain V in close proximity to the 5' splice site (Figure 1.2).

### 1.3.2 Domain II

This domain displays substantial size and structure variations, notably between subclasses IIA and IIB. Its poor phylogenetic conservation suggests a limited if any participation in catalysis (Lehmann and Schmidt, 2003). Accordingly, replacement of domain II with a short linker region between domains I and III has minimal effect on splicing (Kwakman *et al.*, 1989; Bachl and Schmelzer, 1990; Qin and Pyle, 1998).

Domain II contacts domains I and VI via the θ – θ' and η – η' tertiary interactions, respectively (Figure 1.2). The θ – θ' interaction stabilizes the catalytic core (Costa *et al.*, 1997), while the η – η' interaction plays a key role in the conformational rearrangement that occurs between the two transesterification reactions of the splicing pathway (Chanfreau and Jacquier, 1996).

### 1.3.3 Domain III – enhancing the rate of splicing

Deletion of domain III significantly inhibits, but does not abolish, splicing *in vitro* (Koch *et al.*, 1992). Supplying this domain *in trans* to reconstituted intron ribozymes *in vitro* rescues the splicing reaction (Xiang *et al.*, 1998). Domain III is therefore considered a catalytic effector for splicing and has been hypothesized to contribute to formation of the catalytic core (Qin and Pyle, 1998). However, phylogenetic analyses failed to identify potential tertiary interactions between domain III and any other domain. The first tertiary interaction involving domain III was identified biochemically and links the terminal loop of domain III and domain V (Fedorova and Pyle, 2005). This newly defined μ – μ' interaction is believed to anchor domain III to the catalytic core (Figure 1.2). Additionally, a long-range interaction exists between the segment connecting domains II and III and the last nucleotide of the intron (Jacquier and Michel, 1990). This γ – γ' interaction stabilizes the structure of the intron and is important for the second transesterification reaction (Jacquier and Michel, 1990).

### 1.3.4 Domain IV – potentially containing an ORF

Domain IV is greatly variable in size and displays no phylogenetic conservation among group II introns (Michel and Ferat, 1995). It protrudes from the catalytic core (de Lencastre *et al.*, 2005) and is dispensable for splicing *in vitro* (Jarrell *et al.*, 1988; Koch *et al.*, 1992). In many cases, this domain contains

an ORF coding for an Intron-Encoded Protein (IEP) with maturase (Mat), reverse-transcriptase (RT) and endonuclease (Endo) activities (Figure 1.3A) (Lambowitz and Zimmerly, 2004). The IEP is essential for intron splicing and mobility *in vivo* (see section 1.4.1).

### 1.3.5 Domain V – catalyzing the splicing reaction

Along with domain I, domain V is the only required domain for catalysis as its deletion completely abolishes splicing (Koch *et al.*, 1992). This short stem-loop domain is the most conserved region among group II introns (Michel and Ferat, 1995). The multiple interactions linking domain V with domains I ($\zeta - \zeta'$, $\kappa - \kappa'$ and $\lambda - \lambda'$) and III ($\mu - \mu'$) were described above and contribute to the formation of the catalytic core (de Lencastre *et al.*, 2005). The X-ray and NMR structures of domain V were determined and confirmed its role in coordinating a magnesium ion essential for formation of the catalytic core (Zhang and Doudna, 2002; Sigel *et al.*, 2004; Seetharaman *et al.*, 2006).

### 1.3.6 Domain VI – presenting the branch site

Domain VI contains the branch-point nucleotide, which is most often an adenosine (Michel and Ferat, 1995). The 2'-OH group of this bulged nucleotide acts as the nucleophile in the first transesterification reaction of the splicing pathway (Qin and Pyle, 1998). Replacing this adenosine by another nucleotide dramatically decreases the splicing efficiency (Gaur *et al.*, 1997; Liu *et al.*, 1997). Moreover, natural introns presenting aberrant branch points, either mutated or non

bulged, splice via alternate pathways and do not produce intron lariats (Li-Pook-Than and Bonen, 2006).

The spatial arrangement of domains V and VI is critical for positioning the branch-point within the catalytic core (Dib-Hajj *et al.*, 1993; Qin and Pyle, 1998). It has been shown that modification of the 3-nt linker between these two domains decreases splicing efficiency and the fidelity of branch-point selection (Boulanger *et al.*, 1996).

### 1.3.7 A model for the 3D structure of the catalytic core

3D modeling of the yeast ai5γ intron catalytic core revealed that other elements in addition to domain V are essential components of the catalytic heart of the ribozyme. These include the linker region between domains II and III, nucleotides involved in the $\varepsilon - \varepsilon$' interaction and some parts of domain I. Importantly, the model shows that both splicing steps occur at a single active site that is fully pre-organized prior to catalysis (de Lencastre *et al.*, 2005). Moreover, this model is consistent with the hypothesis that a conformational change occurs between the two steps of splicing (Chanfreau and Jacquier, 1996).

## 1.4 Group II intron mobility pathways

Groups II introns are not only self-splicing elements. Some of them are also mobile retroelements that can invade new genomic locations using an RNA intermediate. Group II introns can relocate into new DNA sites by two main mechanisms. They can either invade their cognate homing site, which consists of

an intronless allele, by retrohoming or invade ectopic or non-homologous sites by retrotransposition. Mobility requires the assistance of the intron-encoded protein (IEP).

### 1.4.1 Intron-Encoded Proteins

Some group II introns harbour an open reading frame in their domain IV. In some mitochondrial introns, this ORF is fused to the upstream exon and translated in frame with the exon (Zimmerly *et al.*, 2001). The product is then cleaved to release an active IEP. IEPs typically have four functional domains: reverse-transcriptase (RT), maturase (Mat or X), DNA binding domain (D) and endonuclease (Endo or Zn) (Figure 1.3A) (Lambowitz and Zimmerly, 2004; Toro *et al.*, 2007).

The reverse-transcriptase domain, located in the N-terminal portion of the IEP, shows significant homology to retroviral RT portions 1 to 7. It also harbours additional regions, RT 0 and 2a, analogous to those of non-LTR retrotransposon RTs (Figure 1.3A) (Zimmerly *et al.*, 2001). The RT domain is the most conserved domain of IEPs and this conservation allows for the identification of new group II introns based on sequence analyses.

The maturase domain is essential for promoting intron splicing *in vivo*, by binding the intron and stabilizing its active tertiary conformation (Lambowitz and Zimmerly, 2004). Mutating or deleting the Mat domain abolishes splicing. This domain binds the intron RNA although it displays no detectable RNA-binding motifs (Mohr *et al.*, 1993). The Mat domain exhibits poor sequence conservation and maturases are generally specific for their host introns. ORF-less intron

splicing typically relies on external maturases, which can either be encoded by closely related introns or by the host cell (Lehmann and Schmidt, 2003; Lambowitz and Zimmerly, 2004; Meng *et al.*, 2005).

The C-terminus of many IEPs harbours a DNA-binding domain along with a DNA endonuclease domain (Figure 1.3A). The endonuclease acts during the intron mobility reaction and is not required for splicing. More than half of annotated bacterial IEPs do not contain an Endo domain and their host introns use alternate mobility pathways (see section 1.4.2) (Lambowitz and Zimmerly, 2004; Toro *et al.*, 2007).

### 1.4.2 Retrohoming

Upon transcription of the intron-harbouring gene and IEP translation, the protein binds the intron on the unspliced pre-mRNA and catalyzes splicing through its maturase activity (Matsuura *et al.*, 1997). The IEP remains associated to the liberated intron lariat in the form of a ribonucleoprotein particle (RNP) (Figure 1.3B). The RNP recognizes the cognate intron homing site (HS), consisting of an intronless allele (Belfort *et al.*, 2002). The RNP binds duplex DNA initially in a non-site specific fashion and then screens for the HS sequence (Aizawa *et al.*, 2003). HS reco-gnition is mostly achieved by the intron RNA: the EBS and δ sites located within intron domain I (see section 1.3.1) recognize the IBS and δ' sites present on the HS (which consists of ligated exons) by direct base-pairing (Mohr *et al.*, 2000). The protein component of the RNP also recognizes specific residues of the HS and contributes to the target site specificity

of retrohoming (Guo *et al.*, 1997; Mohr *et al.*, 2000). The IEP unwinds the DNA and the intron invades the top strand of the target DNA by partial or complete reverse-splicing (Figure 1.3B, step 1) (Zimmerly *et al.*, *1995a*; Cousineau *et al.*, 1998). This reaction consists in the reversal of the transesterification reactions that lead to splicing (Eickbush, 1999). It obeys the same rules as splicing: it is fully RNA-catalyzed and requires the IEP's maturase function only to maintain the correct folding of the intron ribozyme. The IEP's endonuclease domain cleaves the bottom strand of the target DNA 9 or 10 nucleotides downstream of the intron insertion site (Figure 1.3B, pathway a, step 2) (Zimmerly *et al.*, 1995a; Matsuura *et al.*, 1997; Lambowitz and Zimmerly, 2004). The IEP's RT uses the released 3'-OH end of the cleaved anti-sense strand as a primer to synthesize a copy of the intron RNA (Figure 1.3B, step 3). This process is known as Target-Primed Reverse-Transcription (TPRT) (Zimmerly *et al.*, 1995b; Eickbush, 1999). The final step of intron mobility consists of integration of the intron cDNA into the target (Figure 1.3B, step 4) and depends on the host DNA repair mechanisms (Lambowitz and Zimmerly, 2004; Smith *et al.*, 2005). In yeast, where DNA recombination is extremely proficient, three different pathways have been identified depending on the pattern of recombination between the cDNA and the exons (Eskes *et al.*, 2000; Lambowitz and Zimmerly, 2004). In bacteria, retrohoming was shown to be independent from the major RecA homologous recombination system (Mills *et al.*, 1997; Cousineau *et al.*, 1998; Martinez-Abarca *et al.*, 2000).

Only ~40% of the annotated bacterial group II IEPs harbour an endonuclease domain (Toro *et al.*, 2007). Endonuclease-deficient introns cannot

perform second-strand cleavage during mobility, raising the question of how reverse-transcription is primed (Munoz-Adelantado *et al.*, 2003). It was suggested that these introns reverse-splice either into transiently single-stranded or double-stranded DNA (Munoz-Adelantado *et al.*, 2003; Robart and Zimmerly, 2005). The favoured candidate for single-stranded DNA is the DNA strand that serves as a template for the lagging strand synthesis at a DNA replication fork (Martinez-Abarca *et al.*, 2004). The reverse-transcription process would then be primed by the nascent DNA strand at the replication fork (Figure 1.3B, pathway a') (Robart and Zimmerly, 2005; Toro *et al.*, 2007).

An alternate mobility pathway was also described for bacterial group IIC introns. These introns are often found downstream of transcriptional terminators (Toor *et al.*, 2001; Dai and Zimmerly, 2002a; Dai *et al.*, 2003). These introns have a shortened EBS1/IBS1 pairing (4 nt vs. 6) and lack EBS2/IBS2 sequences (Toor *et al.*, 2001). A thorough mutational analysis showed that the intron specifically recognizes single-stranded DNA stem-loop motifs (Robart *et al.*, 2007). It was suggested that the intron recognizes the stem-loop motif adopted by the transiently single-stranded DNA either at a replication fork or a transcription bubble. The intron reverse-splices downstream of the motif, and the IEP uses the newly synthesized DNA strand to prime reverse-transcription. This mobility pathway explains the exclusive location of group IIC introns downstream of transcriptional terminators, and rationalizes the existence of the same group IIC intron inserted downstream of different terminators in the same host (Robart *et al.*, 2007).

### 1.4.3 Retrotransposition

Group II introns can also invade ectopic sites resembling their homing site by retrotransposition (Figure 1.3B, pathway b), although this process occurs at a relatively low frequency (Lambowitz and Zimmerly, 2004). As in the case of retrohoming, the first step of retrotransposition consists of reverse-splicing of the intron RNA into the target site (Figure 1.3B, pathway b, step 1). Retrotransposition sites typically show a good match for IBS1 but a reduced match for IBS2 and for residues required for protein recognition (Cousineau *et al.*, 2000; Ichiyanagi *et al.*, 2002; Lambowitz and Zimmerly, 2004). Retrotransposition occurs independently from the IEP's endonuclease activity, suggesting intron insertion into transiently single-stranded DNA (Cousineau *et al.*, 2000; Ichiyanagi *et al.*, 2002). Intron insertion is biased towards the template for lagging-strand DNA synthesis (Cousineau *et al.*, 2000; Ichiyanagi *et al.*, 2002). This suggests a pathway where the RT uses nascent DNA strand at the replication fork to prime reverse-transcription (Figure 1.3B, pathway b, step 3) (Ichiyanagi *et al.*, 2002). The aptitude of group II introns to invade non-cognate alleles probably constitutes a precious means of dispersal for these selfish elements in nature. Moreover, genetic diversity can be promoted by retrotransposition of an intron followed by recombination between the original copy and the newly inserted one (Sellem *et al.*, 1993).

## 1.5 Evolution of group II introns

As genetic elements with a dual feature, self-splicing and retromobility, group II introns are at the center of many evolutionary theories. Their origin is still the subject of ongoing debates. They are believed to be particularly prone to lateral transfer between species and even across kingdoms. They are also considered as the progenitors of nuclear eukaryotic introns and the five spliceosomal RNAs (snRNAs), as well as non-LTR retroelements. In this sense, they are believed to have played a significant role in shaping contemporary eukaryotic genomes.

### 1.5.1 Origin of group II introns

Multiple scenarios for the origin of group II introns were suggested due to their dual feature. Phylogenetic analyses revealed a strong coevolution pattern between group II introns and the ORFs they harbour, suggesting an early association between the RT and the ribozyme (Fontaine *et al.*, 1997; Toor *et al.*, 2001). The progenitor of group II introns may either consist of a primordial splicing element that was invaded by a retroelement harbouring a reverse-transcriptase, or a retroelement that developed splicing properties to minimize detrimental effects of its integration on the host (Curcio and Belfort, 1996; Belfort *et al.*, 2002). In any case, it is well accepted that group II introns originated in bacteria, rationalizing their actual distribution in bacteria and organelles (Ferat and Michel, 1993; Zimmerly *et al.*, 2001; Lambowitz and Zimmerly, 2004).

The "ORF-invasion" theory suggests that an ancestral group II intron was invaded by a mobile retroelement, thereby acquiring a reverse-transcriptase ORF

(Lambowitz and Belfort, 1993; Curcio and Belfort, 1996). This scenario is attractive as it parallels the suggested ORF-acquisition of the unrelated group I introns (Bell-Pedersen *et al.*, 1990; Loizos *et al.*, 1994). This theory accounts for the exclusive location of group II intron ORFs within domain IV and for the existence of ORF-less group II introns (Lambowitz and Belfort, 1993; Curcio and Belfort, 1996). It is also supported by the finding that the IEP binds the intron primarily in domain IV, which would be a remnant of the original interaction of the RT with the primordial retroelement that invaded the ancestral intron in domain IV (Wank *et al.*, 1999; Belfort *et al.*, 2002). However, the strong coevolution pattern observed between each intron RNA structural class and ORF subclasses would imply independent ORF-invasion events of different ancestral subclasses of group II introns (Toor *et al.*, 2001).

On the other hand, the "retroelement ancestor hypothesis" suggests that a primordial retroelement developed splicing properties. This theory rationalizes the coevolution observed between intron RNA structural classes and ORF subclasses. It is also supported by the discovery of ORF remnants in some ORF-less introns (Toor *et al.*, 2001; Robart and Zimmerly, 2005). However, it does not rationalize the appearance of a complex ribozyme structure to achieve splicing while the reaction could be achieved by protein enzymes (Lambowitz and Zimmerly, 2004).

### 1.5.2 Dissemination of group II introns

The distribution of group II introns between different hosts suggests that these elements were subject to extensive lateral (or horizontal) transfer (Lambowitz and Belfort, 1993). The retroelement nature of these intervening

sequences further rationalized this hypothesis. Lateral transfer was inferred to explain the existence of highly similar introns in different locations in various organisms, and why some introns at specific sites were more conserved than their flanking exons (Lambowitz and Belfort, 1993). Moreover, phylogenetic analyses on bacterial intron-encoded ORFs and intron RNA structures revealed that specific intron classes are not restrained to particular bacterial groups, suggesting that substantial lateral transfer occurs among bacteria (Zimmerly *et al.*, 2001). The following sub-sections list the reported cases of lateral transfer of group II introns.

#### 1.5.2.1 Intra-kingdom lateral transfer

Sequence analyses and comparison unravelled several cases of alleged lateral transfer of group II introns between related species. Instances of lateral transfer of a group II intron in bacteria include the Ll.LtrB intron from the Gram-positive bacterium *Lactococcus lactis*, several *Escherichia coli* group II introns, and the RmIntI intron from the alfalfa symbiont *Sinorhizobium meliloti*.

Identical copies of the Ll.LtrB group II intron were originally found in the relaxase gene of two highly similar conjugative elements of *L. lactis*, the pRS01 plasmid and the chromosomal sex factor (see section 1.6) (Mills *et al.*, 1996; Shearman *et al.*, 1996). The high degree of similarity between the host conjugative elements suggests that the two intron copies were vertically inherited from a common ancestor (Le Bourgeois P. *et al.*, 2000). However, another highly similar group II intron was discovered in the relaxase gene of the pAH90 conjugative plasmid of a different *L. lactis* strain. This intron is 99% identical to

Ll.LtrB; however, its host relaxase gene has only 33% and 29% identity with the relaxase genes of pRS01 and the sex factor, respectively (O'Sullivan *et al.*, 2001). This discrepancy in sequence homology between introns and flanking exons is a strong indication that the intron was laterally acquired either by pAH90 or by the ancestor of pRS01 and the sex factor. Interestingly, Ll.LtrB was shown to recognize a conserved motif of relaxase genes and to have a propensity to target genes from this family (Staddon *et al.*, 2004).

A thorough survey of the ECOR collection of *E. coli* strains confirmed the presence of five introns, termed *E.c.*I1-5 (Ferat *et al.*, 1994; Dai and Zimmerly, 2002b). Four of these introns, *E.c.*I1-4, are located in various IS elements and the fifth one is located into a virulence plasmid of *E. coli* 0157:H7. The distribution pattern of these introns revealed that they are mostly vertically inherited, but also suggests some cases of possible lateral transfer for introns *E.c.*I1, *E.c.*I3, *E.c.*I4 and fragments of *E.c.*I5 (Dai and Zimmerly, 2002b). It was suggested that these introns were, and probably still are, disseminated via lateral transfer of their host IS elements, potentially followed by mobilization of the introns in their new host (Dai and Zimmerly, 2002b).

Finally, the RmInt1 intron was initially identified in *S. meliloti* (Martinez-Abarca *et al.*, 1998). A systematic search for this intron by Southern hybridization in other related bacteria also interacting with plants revealed that RmInt1 was also found in some *Rhizobium* and *Agrobacterium* species (Fernandez-Lopez *et al.*, 2005). The distribution of the identified intron copies suggests that these organisms acquired the intron by lateral transfer (Fernandez-Lopez *et al.*, 2005).

Some instances of intra-kingdom lateral transfer of organellar group II introns were also reported. Lateral transfer of a group II intron from the mitochondrial *coxI* gene was alleged between recent ancestors of two related yeast species, *Kluyveromyces lactis* and *Saccharomyces cerevisiae* (Hardy and Clark-Walker, 1991). Sequence comparison between the two species revealed that intron similarity (96%) is greater than flanking exon similarity (87.6%) (Hardy and Clark-Walker, 1991). Subsequently, lateral transfer of a group II intron, again from the mitochondrial *coxI* gene, was reported between two algal species: the diatom (yellow algae) *Thalassiosira nordenskioeldii* and the haptophyte *Pavlova lutheri* (Ehara *et al.*, 2000). Lateral transfer was inferred by the high discrepancy of phylogenetic trees based on the intron ORF and the *coxI* exon (Ehara *et al.*, 2000). Finally, the second intron from the mitochondrial *nad1* gene of *Gnetum*, a gymnosperm (seed-bearing plant), was suggested to have been acquired along with its flanking exons from an asterid (a subclass of angiosperms, flowering plants) (Won and Renner, 2003). As both exons are also highly similar, it is likely that this case of transfer did not involve only the intron but was probably a result of DNA transfer (Won and Renner, 2003).

#### 1.5.2.2 Inter-kingdom lateral transfer

Hollander and Kuck showed that a mitochondrial intron from the green alga *Scenedesmus obliquus* was successfully expressed and correctly spliced in *E. coli*, despite its lack of an intron-encoded maturase. This work provided a preliminary rationale that group II introns could be active in a different cellular environment, and strengthened the argument that these introns can be laterally

transferred and maintain their splicing ability (Hollander and Kuck, 1999). In parallel, instances of inter-kingdom laterally transfer were discovered. The first reported case was transfer of a group II intron from the large ribosomal subunit gene from a *Calothrix* cyanobacterium to the mitochondria of *Porphyra purpurea*, a red algae (Burger *et al.*, 1999). Lateral transfer was suggested on the basis of sequence similarity between the introns, as well as phylogenetic analyses of their ORF which placed the *P. purpurea* intron ORF in the same clade as cyano-bacterial intron ORFs (Burger *et al.*, 1999). Subsequently, two studies reported the putative lateral transfer of related group II introns from cyanobacteria to chloroplasts. A group II intron from the *psbA* gene, present only in *Euglena myxocylindracea* and absent from its close sister species, was suggested to have been acquired through lateral transfer from a cyanobacterial donor. Phylogenetic analyses of the intron-encoded ORF supports this hypothesis (Sheveleva and Hallick, 2004). Similarly, the chloroplast *psbA1* gene of a *Chlamydomonas* species, a green alga unrelated to *Euglena*, was suggested to have acquired its group II intron from a cyanobacteria through lateral transfer (Odom *et al.*, 2004). This intron is closely related to the one reported in *Euglena* by Sheveleva and Hallick and has a higher GC content than the regular *Chlamydomonas* genome (45% vs. 34% respectively) (Odom *et al.*, 2004).

Finally, several related group II introns were discovered in the archaeal genomes of *Methanosarcina acetivorans* and the closely related *Methanosarcina mazei* (Dai and Zimmerly, 2003; Toro, 2003). Phylogenetic analyses of the introns ORFs and intron RNA structures suggested that these introns were closely related to *Pseudomonas putida* and *E. coli* group II introns (IEPs having 44% and

55% amino acid identity, respectively) (Dai and Zimmerly, 2003). Moreover, analysis of the ORF showed significant sequence alignment with the *S. meliloti* group II intron RmInt1 ORF (Toro, 2003). Therefore, it was suggested that either archaea acquired their introns from a bacterial host (Toro, 2003) or that the transfer could have gone either way (Dai and Zimmerly, 2003).

#### 1.5.2.3 Lateral gene transfer routes in bacteria

Microbial organisms can exchange genetic material through three major means: i) DNA transformation, whereby a micro-organism uptakes naked DNA from the environment; ii) phage transduction, whereby phages encapsidate foreign DNA during packaging and subsequently transfer this foreign DNA when infecting a new host; iii) bacterial conjugation, by means of which genetic material can be transferred between different bacterial cells encoding a particular transfer apparatus, known as the mating pore (Ochman *et al.*, 2000). Conjugation can promote dissemination of DNA to different species and across kingdoms in certain specific cases (Ochman *et al.*, 2000).

Interestingly, the majority of group II introns in bacteria were found associated with other mobile elements such as IS sequences, plasmids (some of which are conjugative), integrons and transposons (Zimmerly *et al.*, 2001; Dai and Zimmerly, 2002a; Dai *et al.*, 2003). The spread of these mobile elements to different hosts, along with the retromobile character of group II introns, could promote the dissemination of these intervening sequences between different species and even across kingdoms (Dai and Zimmerly, 2002a).

### 1.5.3 Evolutionary descendants of group II introns

Group II introns are largely considered as the ancestors of nuclear eukaryotic introns and the five small nuclear RNAs (snRNAs) part of the spliceosome. Moreover, these retromobile elements may have given birth to eukaryotic non-Long Terminal Repeat (non-LTR) retroelements.

#### 1.5.3.1 Group II introns and spliceosomal introns

Nuclear introns are processed by a complex machinery, the spliceosome, composed of more than two hundred proteins and five small nuclear RNAs (snRNAs, U1, U2, U4, U5 and U6) (Valadkhan, 2007). The similarity of splicing pathways of group II and nuclear introns, both occurring via two transesterification reactions and producing intron lariats, gave rise to the hypothesis of an evolutionary relationship between the two classes of intervening sequences (Sharp, 1985; Cech, 1986). This hypothesis was further strengthened by the similarity in consensus sequences of intron boundaries (GUGYG...AY for group II introns and GURAGU...AG for nuclear introns of higher eukaryotes; GUAGUGU...AG for yeast nuclear introns) (Michel *et al.*, 1989; Jacquier, 1990).

This evolutionary theory suggests that group II introns originated in bacteria (Cavalier-Smith, 1991) and were therefore present in the ancestors of organelles (Roger and Doolittle, 1993). Group II introns would then have invaded the eukaryotic nucleus by transfer of organellar DNA (Palmer and Logsdon, Jr., 1991), and spread by virtue of their reverse-splicing and mobility aptitude (Cavalier-Smith, 1991). Then, they would have degenerated into spliceosomal introns, with the loss of the self-splicing capacity associated with structural

features and the rise of a centralized splicing machinery based on RNA fragments derived from group II introns (Sharp, 1985; Cech, 1986; Cavalier-Smith, 1991; Sharp, 1991). Degeneration of self-splicing group II introns towards minimal introns requiring a complex machinery for splicing was suggested to have stopped their spread in the nucleus (Sharp, 1985; Cavalier-Smith, 1991).

Substantial evidence has since strengthened this theory. It has been demonstrated that the first nucleophilic attack in nuclear intron splicing was initiated by the 2'-OH of a bulged adenosine found near the 3' end of the intron, which is identical to the first step of group II intron splicing (Guthrie, 1991). Moreover, extensive evidence points to the fact that intron removal by the spliceosome is fundamentally an RNA-catalyzed reaction achieved by the five snRNAs, which share striking similarities to portions of group II introns (detailed in section 1.5.3.3), and that the protein components of the splicing machinery have little if any direct role in catalysis (Villa *et al.*, 2002; Nilsen, 2005; Valadkhan, 2007). Such a finding reinforces the argument that nuclear introns and their splicing machinery derived from a ribozyme. The similarity of splicing pathways and intron boundaries on one hand, and the striking parallels observed between snRNAs and portions of group II introns on the other hand are compelling evidences that group II introns constitute the ancestral ribozyme that gave rise to nuclear introns and the RNA components of their splicing machinery (Villa *et al.*, 2002; Valadkhan, 2007).

### 1.5.3.2 Group II introns and snRNAs

Some group II introns are fragmented into two or more pieces and retained the ability to splice. *Trans*-splicing is mostly observed in mitochondria of angiosperms (flowering plants) and chloroplasts of algae and land plants (Bonen, 1993; Michel and Ferat, 1995). The majority of *trans*-splicing group II introns are fragmented in domain IV, and there are two reported cases of fragmentation in domain III (Kohchi *et al.*, 1988; Bonen, 1993; Michel and Ferat, 1995; Qiu and Palmer, 2004). Moreover, two instances of introns fragmented into three pieces were reported, in the *psaA* chloroplast gene of the green alga *Chlamydomonas reinhardtii* (Goldschmidt-Clermont *et al.*, 1991) and the mitochondrial *nad5* gene of the angiosperm plant *Oenothera berteriana* (Knoop *et al.*, 1997). These tripartite introns are both fragmented in domain I upstream of the EBS regions and in domain IV. Comparative structural analysis of the *C. reinhardtii psaA* intron from multiple *Chlamydomonas* species suggests that yet a fourth RNA piece might participate in domain I folding (Turmel *et al.*, 1995). The discovery of this tripartite intron prompted the hypothesis that the five snRNAs evolved from "five easy pieces" of group II introns (Sharp, 1991). However, this theory was awaiting support from structural and functional comparative studies between group II intron- and spliceosome-mediated catalysis (Sharp, 1991).

Many structural and functional resemblances between the five snRNAs and portions of group II introns were described. Domain I is responsible for splice site selection in group II intron splicing. 5' exon recognition is achieved through base-pair interactions between the IBS sequences at the 3' end of exon 1 and the EBS sequences in domain I (see section 1.3.1, Figure 1.2). Similarly, 3' exon

recognition is achieved through EBS3-IBS3 or the δ-δ' base-pair interactions between domain I and exon 2 (Figure 1.2). Sub-domain d3 of domain I, which harbours the EBS and δ sequences, is functionally similar to its putative counterpart U5, which also recognizes the intron-exon junctions in spliceosome-mediated splicing through base-pairing with the exons (Newman and Norman, 1992). It has been shown that a slightly modified U5 can complement subdomain Id3 in group II intron catalysis *in vitro* (Hetzer *et al.*, 1997).

The domain V hairpin shares striking similarities with a conserved intramolecular stem-loop in U6. It has been shown that both domain V and U6 coordinate the same type of reactions during splicing (Peebles *et al.*, 1995; Konforti *et al.*, 1998). NMR studies showed that the intramolecular stem-loop of U6 in the U2/U6 complex is structurally identical to domain V (Sashital *et al.*, 2004; Seetharaman *et al.*, 2006). Notably, it was shown that domain V can be added *in trans* and rescue splicing of a domain V-deleted group II intron *in vitro* (Jarrell *et al.*, 1988). Moreover, a slightly modified domain V was shown to functionally replace U6 in the minor spliceosome, further strengthening the putative evolutionary link between these two RNA elements (Shukla and Padgett, 2002).

Other resemblances between snRNAs and portions of group II introns include the similarity of U1 with the 5' portion of group II intron domain I. This portion recognizes the 5' end of the intron through the ε – ε' and λ – λ' base-pair interactions (see section 1.3.1, Figure 1.2), similar to the role of U1 in the spliceosome (Jacquier and Michel, 1990; Steitz, 1992; Nilsen, 1994). Finally,

binding of the U2 snRNA to the distal 3' end of nuclear introns causes bulging of the branch point nucleotide in a structure similar to domain VI of group II introns (Parker *et al.*, 1987; Jacquier, 1990).

Taken together, these similarities support the theory that the five snRNAs evolved from group II intron fragments. Interestingly, certain classes of splicing elements are considered as derivatives of group II introns and are seen as "evolutionary intermediates" between group II and nuclear introns/snRNAs. These include *trans*-splicing group II introns as well as the "spliced leader RNA" (SL RNA) incorporated at the 5' end of messenger RNAs in trypanosomes and nematodes (Bonen, 1993). The spliced leader RNA was proposed to be an intermediate stage of spliceosome evolution where the 5' end of the ancestral group II intron would still be bound to exon 1 (Bruzik and Steitz, 1990; Bonen, 1993). On the other hand, a highly degenerate form of group II introns, so-called group III introns, is found in *Euglena* species. These introns are composed of a portion of domain I, domain VI and some also contain domain V. Despite the lack of critical structural elements, group III introns achieve accurate splicing as lariats (Michel and Ferat, 1995). It was proposed that these particular splicing elements require an undetermined *trans*-acting factor(s) for splicing and are intermediates between group II and nuclear introns (Christopher and Hallick, 1989; Copertino *et al.*, 1991).

#### 1.5.3.3 Group II introns and non-LTR retroelements

Group II intron-encoded reverse-transcriptases share exclusive structural motifs with eukaryotic non-Long Terminal Repeat (non-LTR) retroelements RTs

(Malik *et al.*, 1999; Zimmerly *et al.*, 2001; Blocker *et al.*, 2005). Non-LTR retroelements are found in a wide range of eukaryotic organisms such as mammals, plants, invertebrates and fungi. They differ from LTR retroelements by the absence of flanking repeats and a poly(A) stretch (Schmidt T., 1999). They use a mechanism similar to group II intron mobility to invade DNA; they inflict an endonuclease-mediated cut site to the target DNA, and use the nicked strand to prime reverse-transcription (TPRT, see section 1.4.2) (Kazazian, Jr. and Moran, 1998; Boeke, 2003). These similarities prompted the hypothesis that group II introns are the ancestors of non-LTR elements (Zimmerly *et al.*, 1995b; Eickbush, 1999). It was suggested that upon invasion of the eukaryotic nucleus, some group II introns lost their ORFs and evolved towards spliceosome-dependant introns and some lost their ribozyme structure and splicing ability and evolved towards non-LTR retroelements (Eickbush, 1999; Robart and Zimmerly, 2005).

## 1.6 The *Lactococcus lactis* group II intron: Ll.LtrB

Two nearly identical conjugative elements were identified in related strains of the Gram-positive bacterium *Lactococcus lactis*: the pRS01 conjugative plasmid (48.4 kb), isolated in *L. lactis* ML3 (Anderson and McKay, 1984) and the chromosomal sex factor, an Integrative and Conjugative Element (ICE) isolated in *L. lactis* 712 (Gasson *et al.*, 1992). These two elements were found to harbour a group II intron termed Ll.LtrB (pRS01) and intL (sex factor) in their respective relaxase genes (Mills *et al.*, 1996; Shearman *et al.*, 1996). Since then, Ll.LtrB was shown to actively splice and mobilize in its host and is a master model system to study bacterial group II introns.

### 1.6.1 A splicing and mobile bacterial group II intron

The Ll.LtrB group II intron was discovered by detecting an unexpected RT motif lying in essential genes for conjugative transfer. This 2492 nt-long intron contains an ORF in domain IV coding for a 599 amino acid-protein termed LtrA (pRS01) or MobR (sex factor) (Mills *et al.*, 1996; Shearman *et al.*, 1996). Ll.LtrB was the first bacterial group II intron shown to splice *in vivo* in its natural host, *L. lactis* (Mills *et al.*, 1996; Shearman *et al.*, 1996) and in a heterologous bacterial host, *E. coli* (Matsuura *et al.*, 1997). It was also shown to self-splice *in vitro* (Matsuura *et al.*, 1997). Moreover, Ll.LtrB was the first bacterial group II intron shown to invade its homing site both in *L. lactis* and in *E. coli* (Mills *et al.*, 1997; Cousineau *et al.*, 1998). This proficiency in splicing and mobility *in vivo* makes Ll.LtrB a master model system. It allowed the generation of several experimental systems to investigate multiple aspects of group II intron biology. It also allowed the development of new genetic tools based on this active retroelement.

Homing of Ll.LtrB was shown to proceed predominantly via an RNA intermediate (i.e. retrohoming) (Cousineau *et al.*, 1998). This was demonstrated by the insertion of an autocatalytic group I intron, phage T4 *td* intron, into Ll.LtrB. Absence of *td* from the mobility products attested that during the mobility process an RNA intermediate was formed, from which the *td* intron was spliced out prior to reverse-transcription. It was suggested that mobility occurred via complete reverse-splicing of the intron followed by LtrA-mediated reverse-transcription. Finally, Ll.LtrB retrohoming was shown to be independent of the RecA recombination system, both in *L. lactis* and in *E. coli* (Cousineau *et al.*, 1998).

The final steps of mobility were shown to require host DNA repair mechanisms (Smith *et al.*, 2005). Ll.LtrB was also the first bacterial group II intron shown to invade ectopic sites in its host by retrotransposition (Cousineau *et al.*, 2000). This process showed to be independent of LtrA's endonuclease activity in *L. lactis*, but not in *E. coli* (Cousineau *et al.*, 2000; Ichiyanagi *et al.*, 2002; Coros *et al.*, 2005).

Ll.LtrB can be retargeted to invade specific sequences by changing its EBS sites accordingly (Mohr *et al.*, 2000). This feature was used to build "targetrons" capable of disrupting sequences randomly or specifically in multiple heterologous hosts including Gram-negative bacteria (Karberg *et al.*, 2001; Zhong *et al.*, 2003; Perutka *et al.*, 2004; Jones, III *et al.*, 2005; Yao and Lambowitz, 2007), Gram-positive bacteria (Staddon *et al.*, 2006; Yao *et al.*, 2006; Chen *et al.*, 2007) and even human cells (Guo *et al.*, 2000). Moreover, foreign genetic information can be inserted within domain IV of Ll.LtrB (Matsuura *et al.*, 1997; Cousineau *et al.*, 1998) and given its minimal host requirements for the mobility process, Ll.LtrB can be used as a gene delivery vector (Frazier *et al.*, 2003; Plante and Cousineau, 2006; Rawsthorne *et al.*, 2006).

### 1.6.2 The intron-encoded protein, LtrA

Splicing of Ll.LtrB *in vivo* was demonstrated to be LtrA-dependant both in *L. lactis* and *E. coli* (Mills *et al.*, 1996; Shearman *et al.*, 1996; Matsuura *et al.*, 1997; Cousineau *et al.*, 1998). LtrA has all three functional domains usually found within group II intron-encoded proteins: reverse-transcriptase, maturase and endonuclease (Matsuura *et al.*, 1997). All three functions were demonstrated

biochemically (Matsuura *et al.*, 1997; Saldanha *et al.*, 1999) and were shown to be essential for intron retrohoming in *L. lactis* and *E. coli* (Cousineau *et al.*, 1998).

*ltrA* transcription is initiated from a constitutive internal promoter located within Ll.LtrB domain II (Figure 1.4), in contrast to the majority of group II IEPs (Zhou *et al.*, 2000; Chen *et al.*, 2005). Upon translation, two LtrA proteins bind the Ll.LtrB RNA to form a RNP complex (Saldanha *et al.*, 1999; Rambo and Doudna, 2004). The protein autoregulates its level of expression by binding the intron domain IVa as a primary binding site, which results in occlusion of its own Shine-Dalgarno sequence (Wank *et al.*, 1999; Singh *et al.*, 2002; Watanabe and Lambowitz, 2004). In addition, the protein binds the intron RNA at secondary sites located in domains I, V and VI (Matsuura *et al.*, 2001) and its binding was shown to stabilize critical tertiary interactions allowing splicing *in vivo* (Noah and Lambowitz, 2003).

### 1.6.3 An intron interrupting an essential gene for conjugation

Ll.LtrB interrupts the relaxase gene of both the pRS01 plasmid and the sex factor (Mills *et al.*, 1996; Shearman *et al.*, 1996). Relaxase initiates the DNA transfer process by inflicting a site- and strand-specific nick at the origin of transfer (*oriT*) of the conjugative element by transesterification (Byrd and Matson, 1997). Relaxase remains covalently bound to one DNA strand and is translocated through the conjugative pore from the donor to the recipient cell, "dragging" the DNA along. Finally, by reversal of its initial transesterification reaction, relaxase ligates the two ends of the transferred DNA strand, which is then used as a template for second-strand synthesis (Llosa *et al.*, 2002). Therefore, relaxase is

essential for the onset of DNA conjugation. Splicing of Ll.LtrB from the relaxase transcripts of both pRS01 and the sex factor was shown to be essential for transfer of these host elements, likely by allowing ligation of the two exons and expression of a functional relaxase enzyme (Mills *et al.*, 1996; Shearman *et al.*, 1996).

The intimate relationship between intron splicing and pRS01/sex factor conjugation was exploited to design an elegant genetic assay. In this sytem, conjugation of a mobilizable plasmid, i.e. a plasmid harbouring an *oriT*, was used as a read-out for intron splicing efficiency (Klein *et al.*, 2004). Quantitative RT-PCR analysis showed that conjugation efficiency of the mobilizable plasmid is correlated with the amount of ligated relaxase exons produced upon Ll.LtrB splicing. This system thus constituted a sensitive assay to assess Ll.LtrB splicing efficiency on a $10^6$-fold scale (Klein *et al.*, 2004).

### 1.7 Thesis objective

Group II introns are at the center of many evolutionary theories. The objective of the present work is to provide experimental support to some of theses theories. We focused our efforts on two main questions; i) does association of bacterial group II introns with mobile elements promote their dissemination and lateral transfer? ii) did spliceosomal introns and snRNAs evolve from fragmentation of an ancestral group II intron?

We chose to address these questions using experimental approaches. The proficiency of splicing and mobility of Ll.LtrB in its natural host, combined with the simplicity of genetic manipulations in *L. lactis*, made this intron a perfect

candidate for a model system. In an effort to be biologically relevant and representative of natural phenomena, we designed *in vivo* experimental systems in Ll.LtrB's natural host, *L. lactis*. Moreover, Ll.LtrB is found within three conjugative elements and splicing of the intron naturally controls transfer efficiency of its host elements by controlling the amount of relaxase produced. These key features allowed the design of unique experimental systems to address some evolutionary questions.

We constructed several conjugation/mobility assays to address if the association of group II introns with mobile elements promotes their dissemination by lateral transfer. We used a mobilizable plasmid (pLE12) harbouring a portion of pRS01 spanning the *oriT* and the relaxase gene interrupted by Ll.LtrB. We assessed conjugative transfer of pLE12 between different *L. lactis* strains and to *Enterococcus faecalis*, another low-GC Gram-positive bacterium. Subsequently, we monitored Ll.LtrB mobility efficiency in the recipient cells (Chapter Two). We also monitored transfer of the chromosomal sex factor between *L. lactis* strains and subsequent mobility of the Ll.LtrB intron it carries into recipient cells (Chapter Three). Finally, we monitored transfer of the sex factor and pRS01 to *E. faecalis* and subsequent Ll.LtrB mobility in the new host (Chapter Four).

To address the theory of evolution of nuclear introns and snRNAs from group II introns, we assessed the potential of *trans*-splicing of group II introns. We first developed an *in vivo* conjugation-based assay to assess Ll.LtrB *trans*-splicing when fragmented at specific locations (Chapter Five). Then, we built a Tn5-based genetic screen to assess all possible fragmentation sites of Ll.LtrB that

allowed *trans*-splicing. We monitored whether these sites were consistent with fragmentation of an ancestral group II intron towards the snRNAs (Chapter Six).

## 1.8 References

**Aizawa, Y., Xiang, Q., Lambowitz, A.M. and Pyle, A.M.** (2003) The pathway for DNA recognition and RNA integration by a group II intron retrotransposon. *Mol. Cell* **11:** 795-805.

**Anderson, D.G. and McKay, L.L.** (1984) Genetic and physical characterization of recombinant plasmids associated with cell aggregation and high-frequency conjugal transfer in *Streptococcus lactis* ML3. *J. Bacteriol.* **158:** 954-962.

**Bachl, J. and Schmelzer, C.** (1990) Effect of deletions at structural domains of group II intron bl1 on self-splicing in vitro. *J. Mol. Biol.* **212:** 113-125.

**Belfort, M., Derbyshire V., Parker M.M., Cousineau, B. and Lambowitz, A.M.** (2002) Mobile introns: Pathways and proteins. In *Mobile DNA II*, N.L.Craig, R.Craigie, Gellert M., and A.M.Lambowitz, eds. (Washington, D.C.: ASM Press), pp. 761-783.

**Bell-Pedersen, D., Quirk, S., Clyman, J. and Belfort, M.** (1990) Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. *Nucleic Acids Res.* **18:** 3763-3770.

**Berget, S.M., Moore, C. and Sharp, P.A.** (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. USA* **74:** 3171-3175.

**Blocker, F.J., Mohr, G., Conlan, L.H., Qi, L., Belfort, M. and Lambowitz, A.M.** (2005) Domain structure and three-dimensional model of a group II intron-encoded reverse transcriptase. *RNA*. **11:** 14-28.

**Boeke, J.D.** (2003) The unusual phylogenetic distribution of retrotransposons: a hypothesis. *Genome Res.* **13:** 1975-1983.

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

**Boudvillain, M., de Lencastre A. and Pyle, A.M.** (2000) A tertiary interaction that links active-site domains to the 5' splice site of a group II intron. *Nature* **406:** 315-318.

**Boudvillain, M. and Pyle, A.M.** (1998) Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis. *EMBO J.* **17:** 7091-7104.

**Boulanger, S.C., Faix, P.H., Yang, H., Zhuo, J., Franzen, J.S., Peebles, C.L. and Perlman, P.S.** (1996) Length changes in the joining segment between domains 5 and 6 of a group II intron inhibit self-splicing and alter 3' splice site selection. *Mol. Cell Biol.* **16:** 5896-5904.

**Bruzik, J.P. and Steitz, J.A.** (1990) Spliced leader RNA sequences can substitute for the essential 5' end of U1 RNA during splicing in a mammalian in vitro system. *Cell* **62:** 889-899.

**Burger, G., Saint-Louis, D., Gray, M.W. and Lang, B.F.** (1999) Complete sequence of the mitochondrial DNA of the red alga *Porphyra purpurea.* Cyanobacterial introns and shared ancestry of red and green algae. *Plant Cell* **11:** 1675-1694.

**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.

**Chanfreau, G. and Jacquier, A.** (1996) An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J.* **15:** 3466-3476.

**Chen, Y., Caruso, L., McClane, B., Fisher, D. and Gupta, P.** (2007) Disruption of a toxin gene by introduction of a foreign gene into the chromosome of *Clostridium perfringens* using targetron-induced mutagenesis. *Plasmid.* Epub ahead of print

**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.

**Christopher, D.A. and Hallick, R.B.** (1989) *Euglena gracilis* chloroplast ribosomal protein operon: a new chloroplast gene for ribosomal protein L5 and

description of a novel organelle intron category designated group III. *Nucleic Acids Res.* **17:** 7591-7608.

**Copertino, D.W., Christopher, D.A. and Hallick, R.B.** (1991) A mixed group II/group III twintron in the *Euglena gracilis* chloroplast ribosomal protein S3 gene: evidence for intron insertion during gene evolution. *Nucleic Acids Res.* **19:** 6491-6497.

**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.

**Costa, M., Deme, E., Jacquier, A. and Michel, F.** (1997) Multiple tertiary interactions involving domain II of group II self-splicing introns. *J. Mol. Biol.* **267:** 520-536.

**Costa, M. and Michel, F.** (1999) Tight binding of the 5' exon to domain I of a group II self-splicing intron requires completion of the intron active site. *EMBO J.* **18:** 1025-1037.

**Costa, M. and Michel, F.** (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* **14:** 1276-1285.

**Costa, M., Michel, F. and Westhof, E.** (2000) A three-dimensional perspective on exon binding by a group II self-splicing intron. *EMBO J.* **19:** 5007-5018.

**Cousineau, B., Lawrence, S., Smith, D. and Belfort, M.** (2000) Retrotransposition of a bacterial group II intron. *Nature* **404:** 1018-1021. Erratum in (2001) *Nature* **414:** 84.

**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.

**Curcio, M.J. and Belfort, M.** (1996) Retrohoming: cDNA-mediated mobility of group II introns requires a catalytic RNA. *Cell* **84:** 9-12.

**Dai, L., Toor, N., Olson, R., Keeping, A. and Zimmerly, S.** (2003) Database for mobile group II introns. *Nucleic Acids Res.* **31:** 424-426.

**Dai, L. and Zimmerly, S.** (2002a) Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res.* **30:** 1091-1102.

**Dai, L. and Zimmerly, S.** (2002b) The dispersal of five group II introns among natural populations of Escherichia coli. *RNA.* **8:** 1294-1307.

**Dai, L. and Zimmerly, S.** (2003) ORF-less and reverse-transcriptase-encoding group II introns in archaebacteria, with a pattern of homing into related group II intron ORFs. *RNA.* **9:** 14-19.

**Del Campo M., Tijerina P., Bhaskaran H., Mohr S., Yang Q., Jankowsky E., Russell R. and Lambowitz A.M.** (2007) Do DEAD-box proteins promote group II intron splicing without unwinding RNA? *Mol. Cell.* **28:** 159-166.

**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.

**Dib-Hajj, S.D., Boulanger, S.C., Hebbar, S.K., Peebles, C.L., Franzen, J.S. and Perlman, P.S.** (1993) Domain 5 interacts with domain 6 and influences the second transesterification reaction of group II intron self-splicing. *Nucleic Acids Res.* **21:** 1797-1804.

**Ehara, M., Watanabe, K.I. and Ohama, T.** (2000) Distribution of cognates of group II introns detected in mitochondrial *coxI* genes of a diatom and a haptophyte. *Gene* **256:** 157-167.

**Eickbush, T.H.** (1999) Mobile introns: retrohoming by complete reverse splicing. *Curr. Biol.* **9:** R11-R14.

**Eskes, R., Liu, L., Ma, H., Chao, M.Y., Dickson, L., Lambowitz, A.M. and Perlman, P.S.** (2000) Multiple homing pathways used by yeast mitochondrial group II introns. *Mol. Cell Biol.* **20:** 8432-8446.

**Fauron, C., Casper, M., Gao, Y. and Moore, B.** (1995) The maize mitochondrial genome: dynamic, yet functional. *Trends Genet.* **11:** 228-235.

**Fedorova, O. and Pyle, A.M.** (2005) Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3. *EMBO J.* **24:** 3906-3916.

**Fedorova, O. and Zingler, N.** (2007) Group II introns: structure, folding and splicing mechanism. *Biol. Chem.* **388:** 665-678.

**Ferat, J.L., LeGouar M. and Michel, F.** (1994) Multiple group II self-splicing introns in mobile DNA from *Escherichia coli*. *C. R. Acad. Sci. III* **317:** 141-148.

**Ferat, J.L. and Michel, F.** (1993) Group II self-splicing introns in bacteria. *Nature* **364:** 358-361.

**Fernandez-Lopez, M., Munoz-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.

**Fontaine, J.M., Goux, D., Kloareg, B. and Loiseaux-de Goër, S.** (1997) The reverse-transcriptase-like proteins encoded by group II introns in the mitochondrial genome of the brown alga *Pylaiella littoralis* belong to two different lineages which apparently coevolved with the group II ribozyme lineages. *J. Mol. Evol.* **44:** 33-42.

**Frazier, C.L., San Filippo J., Lambowitz,A.M. and Mills, D.A.** (2003) Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: generation of stable insertions without selection. *Appl. Environ. Microbiol.* **69:** 1121-1128.

**Gasson, M.J., Swindell, S., Maeda, S. and Dodd, H.M.** (1992) Molecular rearrangement of lactose plasmid DNA associated with high-frequency transfer and cell aggregation in *Lactococcus lactis* 712. *Mol. Microbiol.* **6:** 3213-3223.

**Gaur, R.K., McLaughlin, L.W. and Green, M.R.** (1997) Functional group substitutions of the branchpoint adenosine in a nuclear pre-mRNA and a group II intron. *RNA.* **3:** 861-869.

**Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M. and Rochaix, J.D.** (1991) A small chloroplast RNA may be required for *trans*-splicing in *Chlamydomonas reinhardtii. Cell* **65:** 135-143.

**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.

**Guo, H., Zimmerly, S., Perlman, P.S. and Lambowitz, A.M.** (1997) Group II intron endonucleases use both RNA and protein subunits for recognition of specific sequences in double-stranded DNA. *EMBO J.* **16:** 6835-6848.

**Guthrie, C.** (1991) Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science* **253:** 157-163.

**Hardy, C.M. and Clark-Walker, G.D.** (1991) Nucleotide sequence of the COX1 gene in *Kluyveromyces lactis* mitochondrial DNA: evidence for recent horizontal transfer of a group II intron. *Curr. Genet.* **20:** 99-114.

**Harris-Kerr, C.L., Zhang, M. and Peebles, C.L.** (1993) The phylogenetically predicted base-pairing interaction between alpha and alpha' is required for group II splicing *in vitro*. *Proc. Natl. Acad. Sci. USA* **90:** 10658-10662.

**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.

**Hollander, V. and Kuck, U.** (1999) Group II intron splicing in *Escherichia coli*: phenotypes of cis-acting mutations resemble splicing defects observed in organelle RNA processing. *Nucleic Acids Res.* **27:** 2339-2344.

**Ichiyanagi, 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.

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

**Jacquier, A. and Jacquesson-Breuleux, N.** (1991) Splice site selection and role of the lariat in a group II intron. *J. Mol. Biol.* **219:** 415-428.

**Jacquier, A. and Michel, F.** (1987) Multiple exon-binding sites in class II self-splicing introns. *Cell* **50:** 17-29.

**Jacquier, A. and Michel, F.** (1990) Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns. *J. Mol. Biol.* **213:** 437-447.

**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.

**Jones, J.P. III, Kierlin, M.N., Coon, R.G., Perutka, J., Lambowitz, A.M. and Sullenger, B.A.** (2005) Retargeting mobile group II introns to repair mutant genes. *Mol. Ther.* **11:** 687-694.

**Karberg, M., Guo, H., Zhong, J., Coon, R., Perutka, J. and Lambowitz, A.M.** (2001) Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nat. Biotechnol.* **19:** 1162-1167.

**Kazazian, H.H. Jr. and Moran, J.V.** (1998) The impact of L1 retrotransposons on the human genome. *Nat. Genet.* **19:** 19-24.

**Klein, J.R., Chen, Y., Manias, D.A., Zhuo, J., Zhou, L., Peebles, C.L. and Dunny, G.M.** (2004) A conjugation-based system for genetic analysis of group II intron splicing in *Lactococcus lactis*. *J. Bacteriol.* **186:** 1991-1998.

**Knoop, V.** (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.* **46:** 123-139.

**Knoop, V., Altwasser, M. and Brennicke, A.** (1997) A tripartite group II intron in mitochondria of an angiosperm plant. *Mol. Gen. Genet.* **255:** 269-276.

**Koch, J.L., Boulanger, S.C., Dib-Hajj, S.D., Hebbar, S.K. and Perlman, P.S.** (1992) Group II introns deleted for multiple substructures retain self-splicing activity. *Mol. Cell Biol.* **12:** 1950-1958.

**Kohchi, T., Umesono, K., Ogura, Y., Komine, Y., Nakahigashi, K., Komano, T., Yamada, Y., Ozeki, H. and Ohyama, K.** (1988) A nicked group II intron and *trans*-splicing in liverwort, *Marchantia polymorpha*, chloroplasts. *Nucleic Acids Res.* **16:** 10025-10036.

**Konforti, B.B., Abramovitz, D.L., Duarte, C.M., Karpeisky, A., Beigelman, L. and Pyle, A.M.** (1998) Ribozyme catalysis from the major groove of group II intron domain 5. *Mol. Cell* **1:** 433-441.

**Kwakman, J.H., Konings, D., Pel, H.J. and Grivell, L.A.** (1989) Structure-function relationships in a self-splicing group II intron: a large part of domain II of the mitochondrial intron aI5 is not essential for self-splicing. *Nucleic Acids Res.* **17:** 4205-4216.

**Lambowitz, A.M.** (1989) Infectious introns. *Cell* **56:** 323-326.

**Lambowitz, A.M. and Belfort, M.** (1993) Introns as mobile genetic elements. *Annu. Rev. Biochem.* **62:** 587-622.

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

**Lang, B.F., Laforest, M.J. and Burger, G.** (2007) Mitochondrial introns: a critical view. *Trends Genet.* **23:** 119-125.

**Le Bourgeois, P., Daveran-Mingot, M.L. and Ritzenthaler, P.** (2000) Genome plasticity among related ++*Lactococcus* strains: identification of genetic events associated with macrorestriction polymorphisms. *J. Bacteriol.* **182:** 2481-2491.

**Lehmann, K. and Schmidt, U.** (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit Rev. Biochem. Mol. Biol.* **38:** 249-303.

**Li-Pook-Than, J. and Bonen, L.** (2006) Multiple physical forms of excised group II intron RNAs in wheat mitochondria. *Nucleic Acids Res.* **34:** 2782-2790.

**Liu, Q., Green, J.B., Khodadadi, A., Haeberli, P., Beigelman, L. and Pyle, A.M.** (1997) Branch-site selection in a group II intron mediated by active recognition of the adenine amino group and steric exclusion of non-adenine functionalities. *J. Mol. Biol.* **267:** 163-171.

**Llosa, M., Gomis-Ruth, F.X., Coll, M. and de la Cruz, F.F.** (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.* **45:** 1-8.

**Loizos, N., Tillier, E.R. and Belfort, M.** (1994) Evolution of mobile group I introns: recognition of intron sequences by an intron-encoded endonuclease. *Proc. Natl. Acad. Sci. USA* **91:** 11983-11987.

**Malik, H.S., Burke, W.D. and Eickbush, T.H.** (1999) The age and evolution of non-LTR retrotransposable elements. *Mol. Biol. Evol.* **16:** 793-805.

**Martinez-Abarca, F., Barrientos-Duran, A., Fernandez-Lopez, 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.

**Martinez-Abarca, F., Garcia-Rodriguez, 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.

**Martinez-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.

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

**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.

**Meng, Q., Wang, Y. and Liu, X.Q.** (2005) An intron-encoded protein assists RNA splicing of multiple similar introns of different bacterial genes. *J. Biol. Chem.* **280:** 35085-35088.

**Meunier, B., Tian, G.-L., Macadre, C., Slonimski, P.P. and Lazowska, J.** (1990) Group II introns transpose in yeast mitochondria. In *Structure, function and biogenesis of energy transfer systems*, E.Quagliariello, S.Papa, F.Palmieri, and C.Saccone, eds. (Amsterdam: Elsevier), pp. 169-174.

**Michel, F. and Dujon, B.** (1983) Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear-encoded members. *EMBO J.* **2:** 33-38.

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

**Michel, F., Jacquier, A. and Dujon, B.** (1982) Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie* **64:** 867-881.

**Michel, F. and Lang, B.F.** (1985) Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. *Nature* **316:** 641-643.

**Michel, F., Umesono, K. and Ozeki, H.** (1989) Comparative and functional anatomy of group II catalytic introns--a review. *Gene* **82:** 5-30.

**Mills, D.A., Manias, D.A., McKay, L.L. and Dunny, G.M.** (1997) Homing of a group II intron from *Lactococcus lactis* subsp. *lactis* ML3. *J. Bacteriol.* **179:** 6107-6111.

**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.

**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.

**Mohr, G., Smith, D., Belfort, M. and Lambowitz, A.M.** (2000) Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences. *Genes Dev.* **14:** 559-573.

**Munoz-Adelantado, E., San Filippo J., Martinez-Abarca, F., Garcia-Rodriguez, 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.

**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.

**Nilsen, T.W.** (2005) Spliceosome assembly in yeast: one ChIP at a time? *Nat. Struct. Mol. Biol.* **12:** 571-573.

**Noah, J.W. and Lambowitz, A.M.** (2003) Effects of maturase binding and Mg2+ concentration on group II intron RNA folding investigated by UV cross-linking. *Biochemistry* **42:** 12466-12480.

**O'Sullivan, D., Ross, R.P., Twomey, D.P., Fitzgerald, G.F., Hill, C. and Coffey, A.** (2001) Naturally occurring lactococcal plasmid pAH90 links

bacteriophage resistance and mobility functions to a food-grade selectable marker. *Appl. Environ. Microbiol.* **67:** 929-937.

**Ochman, H., Lawrence, J.G. and Groisman, E.A.** (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405:** 299-304.

**Odom, O.W., Shenkenberg, D.L., Garcia, J.A. and Herrin, D.L.** (2004) A horizontally acquired group II intron in the chloroplast *psbA* gene of a psychrophilic *Chlamydomonas*: *in vitro* self-splicing and genetic evidence for maturase activity. *RNA*. **10:** 1097-1107.

**Palmer, J.D. and Logsdon, J.M. Jr.** (1991) The recent origins of introns. *Curr. Opin. Genet. Dev.* **1:** 470-477.

**Parker, R., Siliciano, P.G. and Guthrie, C.** (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49:** 229-239.

**Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A. and Cheng, H.L.** (1986) A self-splicing RNA excises an intron lariat. *Cell* **44:** 213-223.

**Peebles, C.L., Zhang, M., Perlman, P.S. and Franzen, J.S.** (1995) Catalytically critical nucleotide in domain 5 of a group II intron. *Proc. Natl. Acad. Sci. USA* **92:** 4422-4426.

**Perutka, J., Wang, W., Goerlitz, D. and Lambowitz, A.M.** (2004) Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. *J. Mol. Biol.* **336:** 421-439.

**Plante, I. and Cousineau, B.** (2006) Restriction for gene insertion within the *Lactococcus lactis* Ll.LtrB group II intron. *RNA* **12:** 1980-1992.

**Qin, P.Z. and Pyle, A.M.** (1997) Stopped-flow fluorescence spectroscopy of a group II intron ribozyme reveals that domain 1 is an independent folding unit with a requirement for specific Mg2+ ions in the tertiary structure. *Biochemistry* **36:** 4718-4730.

**Qin, P.Z. and Pyle, A.M.** (1998) The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol.* **8:** 301-308.

**Qiu, Y.L. and Palmer, J.D.** (2004) Many independent origins of *trans*-splicing of a plant mitochondrial group II intron. *J. Mol. Evol.* **59:** 80-89.

**Rambo, R.P. and Doudna, J.A.** (2004) Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry* **43:** 6486-6497.

**Rawsthorne, H., Turner, K.N. and Mills, D.A.** (2006) Multicopy integration of heterologous genes, using the lactococcal group II intron targeted to bacterial insertion sequences. *Appl. Environ. Microbiol.* **72:** 6088-6093.

**Robart, A.R., Seo, W. and Zimmerly, S.** (2007) Insertion of group II intron retroelements after intrinsic transcriptional terminators. *Proc. Natl. Acad. Sci. USA* **104:** 6620-6625.

**Robart, A.R. and Zimmerly, S.** (2005) Group II intron retroelements: function and diversity. *Cytogenet. Genome Res.* **110:** 589-597.

**Roger, A.J. and Doolittle, W.F.** (1993) Molecular evolution. Why introns-in-pieces? *Nature* **364:** 289-290.

**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.

**Sashital, D.G., Cornilescu, G., McManus, C.J., Brow, D.A. and Butcher, S.E.** (2004) U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat. Struct. Mol. Biol.* **11:** 1237-1242.

**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.

**Sellem, C.H., Lecellier, G. and Belcour, L.** (1993) Transposition of a group II intron. *Nature* **366:** 176-178.

**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.

**Sheveleva, E.V. and Hallick, R.B.** (2004) Recent horizontal intron transfer to a chloroplast genome. *Nucleic Acids Res.* **32:** 803-810.

**Schmidt T.** (1999) LINEs, SINEs and repetitive DNA: non-LTR retrotransposons in plant genomes. *Plant Mol. Biol.* **40:** 903-910.

**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.

**Sigel, R.K., Sashital, D.G., Abramovitz, D.L., Palmer, A.G., Butcher, S.E. and Pyle, A.M.** (2004) Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif. *Nat. Struct. Mol. Biol.* **11:** 187-192.

**Singh, R.N., Saldanha, R.J., D'Souza, L.M. and Lambowitz, A.M.** (2002) Binding of a group II intron-encoded reverse transcriptase/maturase to its high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation. *J. Mol. Biol.* **318:** 287-303.

**Skelly, P.J., Hardy, C.M. and Clark-Walker, G.D.** (1991) A mobile group II intron of a naturally occurring rearranged mitochondrial genome in *Kluyveromyces lactis*. *Curr. Genet.* **20:** 115-120.

**Smith, D., Zhong, J., Matsuura, M., Lambowitz, A.M. and Belfort, M.** (2005) Recruitment of host functions suggests a repair pathway for late steps in group II intron retrohoming. *Genes Dev.* **19:** 2477-2487.

**Staddon, J.H., Bryan, E.M., Manias, D.A., Chen, Y. and Dunny, G.M.** (2006) Genetic characterization of the conjugative DNA processing system of enterococcal plasmid pCF10. *Plasmid* **56:** 102-111.

**Staddon, J.H., Bryan, E.M., Manias, D.A. and Dunny, G.M.** (2004) Conserved target for group II intron insertion in relaxase genes of conjugative elements of gram-positive bacteria. *J. Bacteriol.* **186:** 2393-2401.

**Steitz, J.A.** (1992) Splicing takes a holliday. *Science* **257:** 888-889.

**Su, L.J., Waldsich, C. and Pyle, A.M.** (2005) An obligate intermediate along the slow folding pathway of a group II intron ribozyme. *Nucleic Acids Res.* **33:** 6674-6687.

**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.

**Toor, N., Robart, A.R., Christianson, J. and Zimmerly, S.** (2006) Self-splicing of a group IIC intron: 5' exon recognition and alternative 5' splicing events implicate the stem-loop motif of a transcriptional terminator. *Nucleic Acids Res.* **34:** 6461-6471.

**Toro, N.** (2003) Bacteria and Archaea Group II introns: additional mobile genetic elements in the environment. *Environ. Microbiol.* **5:** 143-151.

**Toro, N., Jimenez-Zurdo, J.I. and Garcia-Rodriguez, F.M.** (2007) Bacterial group II introns: not just splicing. *FEMS Microbiol. Rev.* **31:** 342-358.

**Turmel, M., Choquet, Y., Goldschmidt-Clermont, M., Rochaix, J.D., Otis, C. and Lemieux, C.** (1995) The *trans*-spliced intron 1 in the *psaA* gene of the *Chlamydomonas* chloroplast: a comparative analysis. *Curr. Genet.* **27:** 270-279.

**Valadkhan, S.** (2007) The spliceosome: a ribozyme at heart? *Biol. Chem.* **388:** 693-697.

**van der Veen, R., Arnberg, A.C., van der Horst G., Bonen, L., Tabak, H.F. and Grivell, L.A.** (1986) Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* **44:** 225-234.

**Villa, T., Pleiss, J.A. and Guthrie, C.** (2002) Spliceosomal snRNAs: Mg(2+)-dependent chemistry at the catalytic core? *Cell* **109:** 149-152.

**Waldsich, C. and Pyle, A.M.** (2007) A folding control element for tertiary collapse of a group II intron ribozyme. *Nat. Struct. Mol. Biol.* **14:** 37-44.

**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.

**Watanabe, K. and Lambowitz, A.M.** (2004) High-affinity binding site for a group II intron-encoded reverse transcriptase/maturase within a stem-loop structure in the intron RNA. *RNA.* **10:** 1433-1443.

**Won, H. and Renner, S.S.** (2003) Horizontal gene transfer from flowering plants to *Gnetum*. *Proc. Natl. Acad. Sci. USA* **100:** 10824-10829.

**Xiang, Q., Qin, P.Z., Michels, W.J., Freeland, K. and Pyle, A.M.** (1998) Sequence specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets. *Biochemistry* **37:** 3839-3849.

**Yao, J. and Lambowitz, A.M.** (2007) Gene targeting in gram-negative bacteria by use of a mobile group II intron ("Targetron") expressed from a broad-host-range vector. *Appl. Environ. Microbiol.* **73:** 2735-2743.

**Yao, J., Zhong, J., Fang, Y., Geisinger, E., Novick, R.P. and Lambowitz, A.M.** (2006) Use of targetrons to disrupt essential and nonessential genes in *Staphylococcus aureus* reveals temperature sensitivity of Ll.LtrB group II intron splicing. *RNA.* **12:** 1271-1281.

**Zhang, L. and Doudna, J.A.** (2002) Structural insights into group II intron catalysis and branch-site selection. *Science* **295:** 2084-2088.

**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.

**Zhou, L., Manias, D.A. and Dunny, G.M.** (2000) Regulation of intron function: efficient splicing in vivo of a bacterial group II intron requires a functional promoter within the intron. *Mol. Microbiol.* **37:** 639-651.

**Zimmerly, S., Guo, H., Eskes, R., Yang, J., Perlman, P.S. and Lambowitz, A.M.** (1995a) A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* **83:** 529-538.

**Zimmerly, S., Guo, H., Perlman, P.S. and Lambowitz, A.M.** (1995b) Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82:** 545-554.

**Zimmerly, S., Hausner, G. and Wu, X.** (2001) Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res.* **29:** 1238-1250.

## 1.9 Figure legends

### Figure 1.1

Splicing of group II introns. **A.** *Cis*-splicing of group II introns. The 2'OH of the branch point present in domain VI of the intron (circled A) performs a nucleophilic attack on the exon 1-intron junction (step 1), generating a 2'-5' linkage. Then, the 3'-OH of the released exon 1 performs a second nucleophilic attack on the intron-exon 2 junction (step 2), releasing the intron lariat and ligating the flanking exons. **B.** Schematic of a fragmented gene undergoing intron *trans*-splicing. The two intron fragments assemble and catalyze intron splicing and exon ligation. Intron, red; exons, blue; branch point, circled A.

### Figure 1.2

Secondary structure of group II introns. Schematic of the characteristic secondary structure of group II introns, which consists of six helical domains (I to VI) radiating from a central wheel. Domain I sub-domains are depicted (Ia to Id). Tertiary interactions are colour-coded and indicated as pairs of Greek letters. Variations among group IIA, IIB and IIC classes are indicated. IBS, intron binding sites; EBS, exon binding sites. Modified from Fedorova O. and Zingler N., 2007

**Figure 1.3**

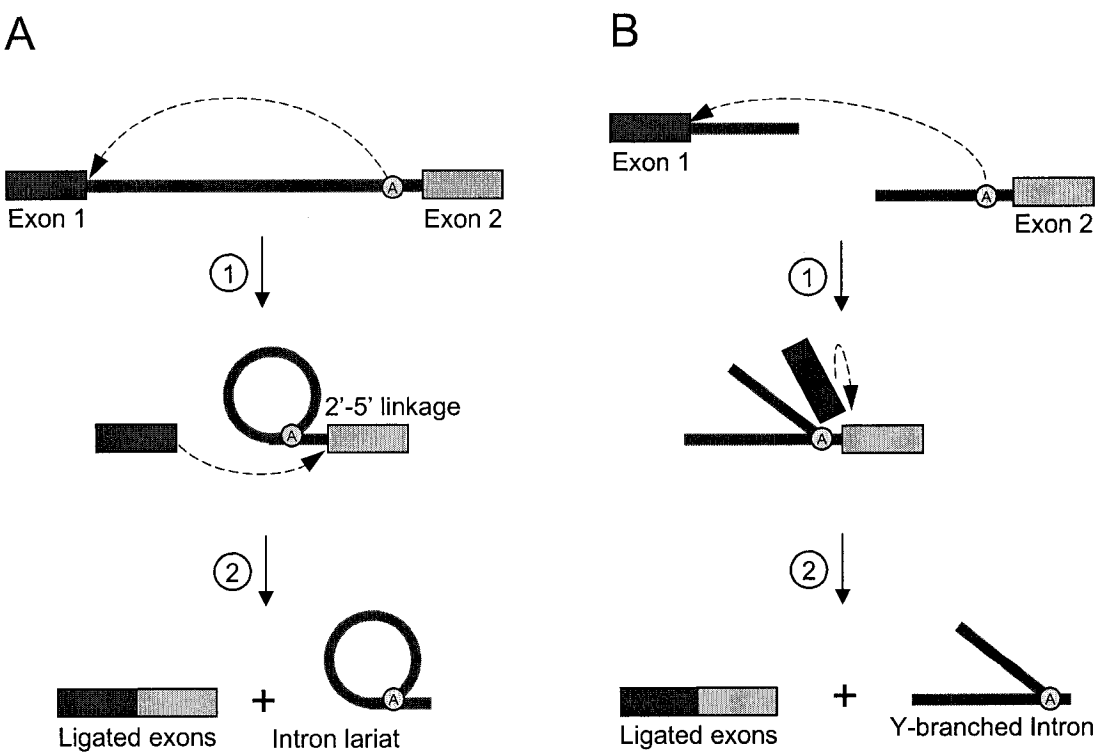
Group II intron encoded protein and mobility pathways. **A.** Domains of Intron-Encoded Proteins. The putative four domains of intron-encoded proteins (RT, Mat, D and Endo) are schematized. RT subdomains are also depicted (RT 0 to 7). Some IEPs lack the D and Endo domains. Adapted from Toro *et al.*, 2007. **B.** Group II intron mobility pathways. Retrohoming (**a**) is initiated upon recognition of the homing-site DNA by the intron RNP (intron RNA+IEP). The intron reverse-splices into its target site (step 1). The endonuclease domain of the IEP nicks the complementary strand nine or ten nucleotides upstream of the insertion site (Pathway **a**, step 2). The intron is then reverse-transcribed by the IEP (TPRT) (step 3). Finally, the DNA repair functions of the host cell degrade the intron RNA and finalize the retrohoming process generating a mobility product (step 4). Group II introns that encode IEPs lacking an endonuclease domain use an alternate pathway for retrohoming (Pathway **a'**). In this pathway, reverse-transcription is primed by nascent DNA at a replication fork. Retrotransposition of group II introns into ectopic sites also occurs independently from the IEP's

endonuclease activity (Pathway **b**). RNA, dotted lines; DNA, solid lines; exons, blue; intron, red; nascent DNA, green; IEP, orange ovale. Adapted from Ichiyanagi *et al.*, 2002.

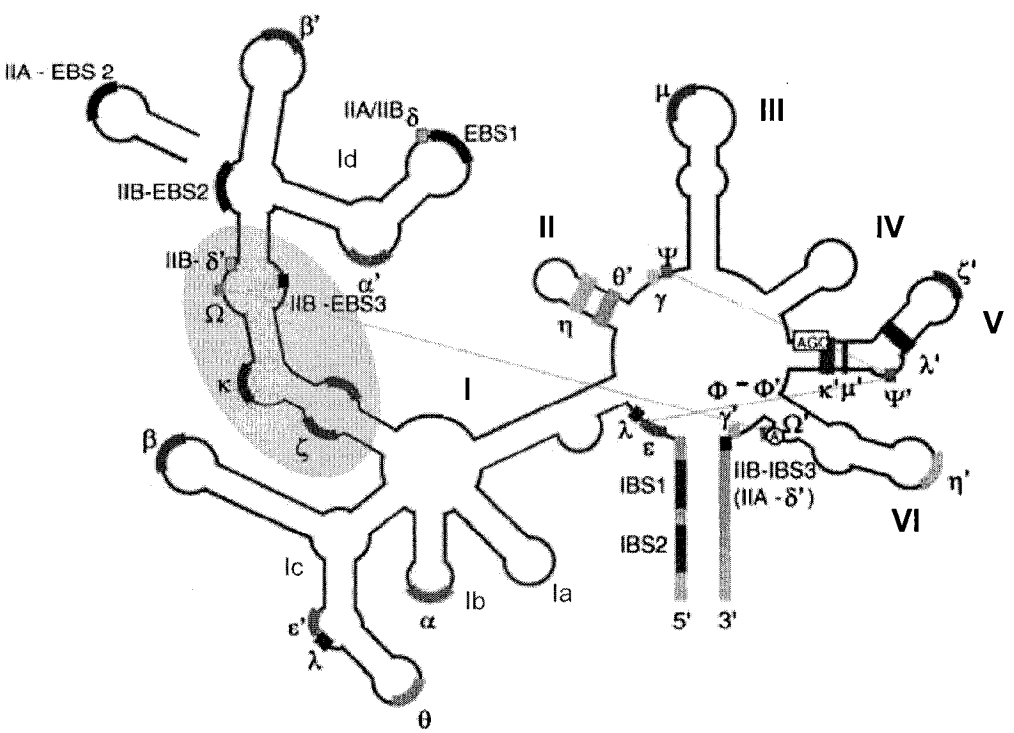
**Figure 1.4**

Proposed Ll.LtrB secondary structure. The six domains of Ll.LtrB are indicated (I to VI). Exons are boxed in blue. The *ltrA* ORF is boxed in yellow and LtrA's enzymatic domains are schematized. LtrA's promoter region is indicated in yellow lines (-35, -10 sequences; +1 is the transcription initiation site). Pairs of Greek letters indicate tertiary interactions. EBS, exon binding sites; IBS, intron binding sites; RT, reverse transcriptase; Mat, maturase; Endo, endonuclease.

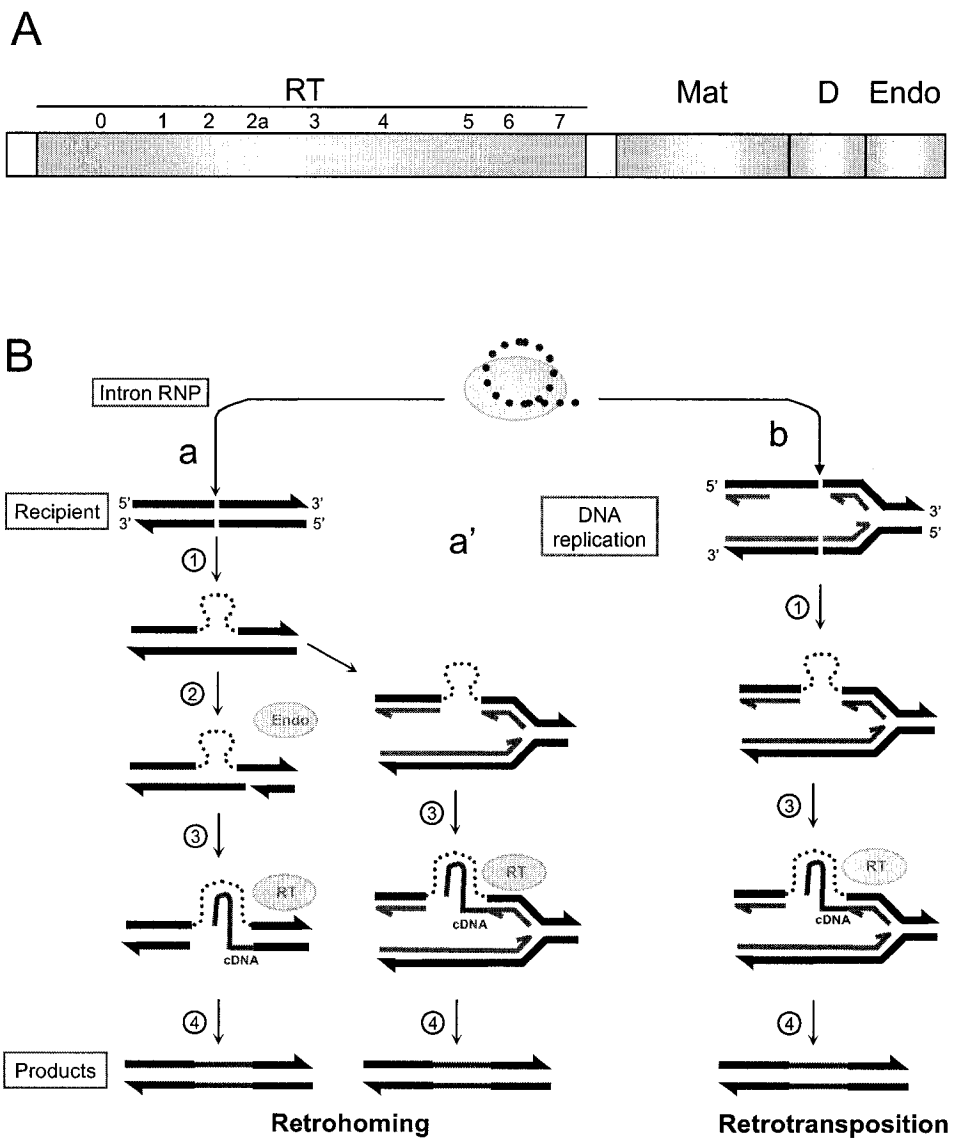
# Figure 1.1


A
Exon 1
Exon 2
1
2'-5' linkage
2
Ligated exons
+
Intron lariat
B
Exon 1
Exon 2
1
2
Ligated exons
+
Y-branched Intron

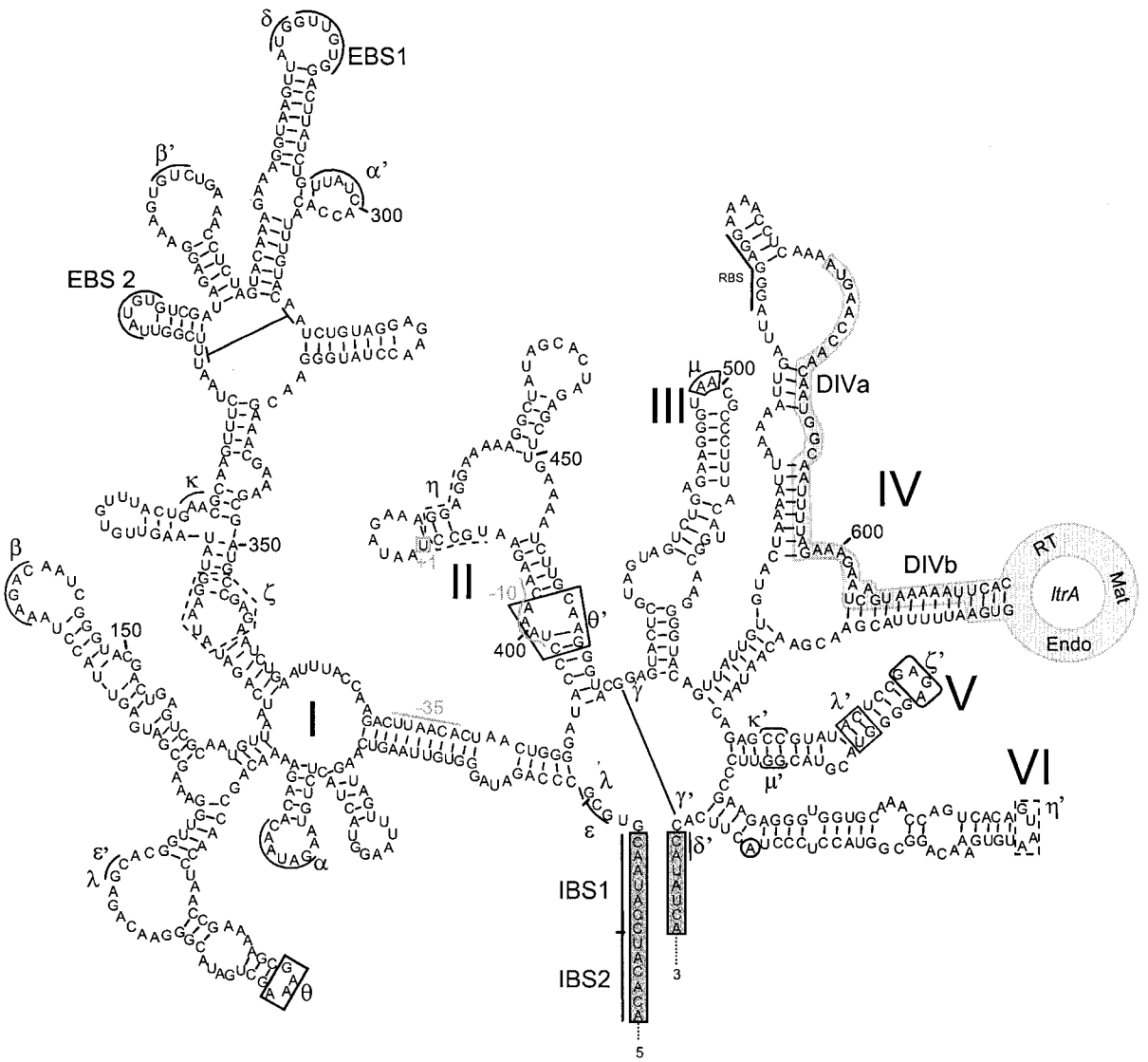
# Figure 1.2


β'
IIA - EBS 2
IIA/IIB δ
EBS1
Id
IIB-EBS2
IIB- δ'
α'
Ω
IIB -EBS3
κ
I
ζ
β
Ic
Ib
Ia
α
ε'
λ
θ
μ
III
II
IV
ζ'
V
θ'
Ψ
γ
η
AGC
λ'
Φ - Φ'
κ'μ'
Ψ'
λ
ε
γ'
Ω'
IBS1
IIB-IBS3
(IIA -δ')
η'
VI
IBS2
5'
3'

# Figure 1.3


A
RT
Mat
D
Endo
0
1
2
2a
3
4
5
6
7
B
Intron RNP
a
b
Recipient
5'
3'
3'
5'
a'
DNA replication
5'
3'
3'
5'
①
②
③
④
Endo
RT
cDNA
Products
Retrohoming
Retrotransposition

# Figure 1.4


EBS1
EBS 2
RBS
DIVa
DIVb
RT
Mat
ltrA
Endo
IBS1
IBS2
I
II
III
IV
V
VI
-35
-10
150
300
350
400
450
500
600
5
3

# CHAPTER TWO

## Conjugation mediates transfer of the Ll.LtrB group II intron between different bacterial species

### 2.1 Preface

It is largely accepted that group II introns readily spread among members of a given species and are laterally transferred between different organisms (Lambowitz and Belfort, 1993). Particularly, the association of group II introns with other mobile elements in bacteria prompted the hypothesis that these intervening sequences take advantage of their host mobile elements to be laterally transferred and disseminated (Dai and Zimmerly, 2002a). This proposition was strengthened by the pattern of distribution of group II introns and their host mobile elements within the same bacterial species (Dai and Zimmerly, 2002b; Fernandez-Lopez et al., 2005). However, it was never demonstrated that group II introns indeed benefit from their association with mobile elements to be disseminated between species.

The following chapter describes the conjugation/mobility assays that we designed to address if the Ll.LtrB intron from *L. lactis* can be disseminated between *L. lactis* strains and to other bacterial species by conjugation. Ll.LtrB being naturally associated with conjugative elements, this route of dissemination was the most biologically relevant and straight forward to test. This study focused on

Ll.LtrB dissemination by a mobilizable plasmid, pLE12. The pLE12 plasmid contains the sex factor/pRS01 origin of transfer (*oriT*) and is thereby recognized and transferred by the conjugative machinery of the chromosomal sex factor of the donor cell. We monitored transfer of pLE12 among *L. lactis* species, and to another low-GC Gram-positive bacterium *Enterococcus faecalis*. We also monitored intron mobility in the new host upon conjugative transfer of its carrying plasmid, either by retrohoming into its homing site present on a resident plasmid, or by retrotransposition into the chromosome of the new host. We studied intron dissemination from the pLE12 mobilizable plasmid, instead of the original host elements (pRS01 and sex factor), to facilitate the assessment of different intron mutants. We analyzed retrohoming of intron mutants for various splicing and mobility attributes to decipher which mobility pathway was used.

Molecular Microbiology (2004) 51(5), 1459–1469 doi:10.1046/j.1365-2958.2003.03923.x

# Conjugation mediates transfer of the Ll.LtrB group II intron between different bacterial species

**Kamila Belhocine, Isabelle Plante and Benoit Cousineau***
*Department of Microbiology and Immunology, McGill University, Montréal, Québec, Canada H3A 2B4.*

In: ***Molecular Microbiology*** (2004) **51**:1459-1469


## 2.2 Summary

Some self-splicing group II introns (ribozymes) are mobile retroelements. These retroelements, which can insert themselves into cognate intronless alleles or ectopic sites by reverse splicing, are thought to be the evolutionary progenitors of the widely distributed eukaryotic spliceosomal introns. Lateral or horizontal transmission of introns (i.e., between species), although never experimentally demonstrated, is a well-accepted model for intron dispersal and evolution. Horizontal transfer of the ancestral bacterial group II introns may have contributed to the dispersal and wide distribution of spliceosomal introns present in modern eukaryotic genomes. Here, the Ll.LtrB group II intron from the Gram-positive bacterium *Lactococcus lactis* was used as a model system to address the dissemination of introns in the bacterial kingdom. We report the first experimental demonstration of horizontal transfer of a group II intron. We show that the Ll.LtrB

group II intron, originally discovered on a *L. lactis* conjugative plasmid (pRS01), and within a chromosomally-located sex factor in *L. lactis* 712, invades new sites using both retrohoming and retrotransposition pathways following its transfer by conjugation. Ll.LtrB lateral transfer is shown among different *L. lactis* strains (intra-species) (retrohoming and retrotransposition) and between *L. lactis* and *Enterococcus faecalis* (inter-species) (retrohoming). These results shed light on long-standing questions about intron evolution and propagation, and demonstrate that conjugation is one of the mechanisms by which group II introns are, and probably were, broadly disseminated between widely diverged organisms.

## 2.3 Introduction

Self-splicing group II introns are large autocatalytic RNAs (ribozymes) (Belfort *et al.*, 2002; Lambowitz and Belfort, 1993; Saldanha *et al.*, 1993; Michel and Ferat, 1995). Some group II introns harbouring an intron-encoded protein (IEP) are also mobile elements. These retroelements are capable of inserting themselves into both cognate intronless alleles (homing sites; HSs) and ectopic sites (non-homologous sites) by retrohoming and retrotransposition, respectively (Belfort *et al.*, 2002). The Ll.LtrB group II intron was initially discovered in the industrially important Gram-positive bacterium, *Lactococcus lactis* (Mills *et al.*, 1996; Shearman *et al.*, 1996). Ll.LtrB is the first bacterial group II intron that was shown to splice and to be mobile *in vivo* (Mills *et al.*, 1996, 1997; Shearman *et al.*, 1996). The mobility pathways of Ll.LtrB for both retrohoming (Cousineau *et al.*, 1998) and retrotransposition (Cousineau *et al.*, 2000; Ichiyanagi *et al.*, 2002) were studied using genetic systems in both *L. lactis* and *Escherichia coli*. Retrohoming is

a very efficient process compared to the insertion of the intron at ectopic sites through the retrotransposition pathway (Cousineau *et al.*, 1998; 2000; Ichiyanagi *et al.*, 2002).

The Ll.LtrB group II intron (2.5 kb) encodes, within the loop region of domain IV, a protein of 599 amino acids called LtrA (1.8 kb) (Mills *et al.*, 1996). The LtrA protein carries three functional domains (reverse transcriptase, maturase, endonuclease) essential in promoting Ll.LtrB mobility via retrohoming (Matsuura *et al.*, 1997; Cousineau *et al.*, 1998). The retrohoming and retrotransposition pathways of the Ll.LtrB group II intron are duplicative processes that, similarly to other retroelements (e.g. retrotransposons and retroviruses), proceed through an RNA intermediate (Cousineau *et al.*, 1998, 2000; Ichiyanagi *et al.*, 2002). Retrohoming occurs through a target DNA-primed reverse transcription mechanism (TPRT). The first step in retrohoming is splicing of the intron from the pre-mRNA, which absolutely requires the maturase activity of its intron-associated protein LtrA *in vivo*. Active ribonucleoprotein particles (RNPs; intron RNA + LtrA) are liberated from the pre-mRNA following splicing of the intron and ligation of the two flanking exons. These RNPs then identify uninterrupted homing sites, which the intron invades by complete reverse splicing into the sense strand (mRNA-like strand) of double-stranded DNA. After insertion of the intron RNA, the endonuclease activity of LtrA cuts the bottom strand 9 nt upstream of the intron insertion site to complete the staggered double strand cut. The liberated 3'-OH of the DNA antisense strand is then recognized by the RT domain of LtrA and is used as a primer to synthesize a full cDNA copy of the intron. The final steps of the retrohoming pathway are mainly supported by the host DNA repair mechanisms

(Cousineau *et al.*, 1998). Even though host DNA repair seems to be involved in establishing intron insertion, the retrohoming pathway of Ll.LtrB was shown to be completely independent of the major RecA-dependent homologous recombination pathway (Mills *et al.*, 1997; Cousineau *et al.*, 1998).

From an evolutionary perspective, group II introns are also very fascinating. They are considered to be the ancestors of the spliceosome-dependent eukaryotic nuclear introns (Sharp, 1991). Indeed, group II and spliceosomal introns share numerous striking similarities. Both intron types are excised as lariats and use the same splicing mechanism (Saldanha *et al.*, 1993; Michel and Ferat, 1995). Moreover, the small nuclear RNAs (snRNAs) that are part of the spliceosome machinery are proposed to be structurally similar to specific domains of group II introns (Sharp, 1991). This theory suggests that fragmentation and degeneracy of ancestral group II introns gave rise to the current eukaryotic spliceosome machinery (Sharp, 1991). On the other hand, the phylogenetic distribution and relative abundance of group II versus spliceosomal introns are quite different. The non-mobile and highly abundant spliceosomal introns (> 16% of the human genome) are exclusively found in the nucleus of eukaryotes. Group II introns, some of which are mobile elements, are found in bacteria and in eukaryotic organelles derived from bacteria, such as fungal and plant mitochondria, as well as plant chloroplasts (Lambowitz and Belfort, 1993; Saldanha *et al.*, 1993; Michel and Ferat, 1995; Belfort *et al.*, 2002). A better understanding of the pathways supporting group II intron propagation, within and between species, should help to explain the idiosyncratic distribution of these two types of introns and the prominence of the spliceosomal introns in eukaryotes. Inter-species dissemination of mobile group II

introns, through lateral transfer, may have contributed to the wide distribution of eukaryotic introns. Horizontal transfer of group II introns between organisms, although never experimentally demonstrated, is a well-accepted model of intron evolution. Horizontal transfer was proposed to explain the presence of closely related introns at different locations (genes or species) and to rationalize why, in specific cases, introns are more conserved than their flanking exons (Lambowitz and Belfort, 1993). The recent study of group II intron distribution in bacterial genomes also suggests substantial horizontal transfer of these introns within the bacterial kingdom (Zimmerly *et al.*, 2001).

Here we describe *L. lactis* genetic assays for lateral transfer of group II introns between bacteria using a biologically relevant and genetically tractable experimental setting. The Ll.LtrB group II intron was originally found to interrupt a relaxase gene (*ltrB*) within two very similar *L. lactis* mobilizable elements: a conjugative plasmid (pRS01) (Mills *et al.*, 1996), and a sex factor embedded within the chromosome of the *L. lactis* 712 strain (Shearman *et al.*, 1996). We thus explored the possibility that conjugation is a basis for the lateral transfer of group II introns promoting their dissemination to new homologous (retrohoming) and non-homologous (retrotransposition) sites both within and between bacterial species.

## 2.4 Results

### 2.4.1 Conjugation/mobility assay in *L. lactis*

To determine whether the Ll.LtrB group II intron could be transferred following conjugation of its host elements between different *L. lactis* strains, we

built a conjugation/mobility assay (Figure 2.1). We used the pLE12 plasmid (16 kb) (Mills *et al.*, 1996; Cousineau *et al.*, 1998), which is smaller and easier to handle than either the original Ll.LtrB-containing conjugative plasmid pRS01 (48.4 kb) (Mills *et al.*, 1994; 1996) or the chromosomally located sex factor (50 kb) (Shearman *et al.*, 1996). The pLE12 plasmid contains only two (Tra1-2) of the four conjugative transfer regions (Tra1-Tra4) identified in pRS01 (Mills *et al.*, 1994) (Figure 2.1A). The Ll.LtrB intron interrupts a relaxase gene (*ltrB*; Figure 2.1A) absolutely required for pRS01 conjugation; hence splicing of the intron is essential for plasmid transfer (Mills *et al.*, 1996; Zhou *et al.*, 2000). The Tra1-2 region from pRS01 also contains, between *ltrD* and *ltrE*, a conjugation transfer origin (*oriT*) recognized by the relaxase to initiate conjugative transfer (Figure 2.1A) (Mills *et al.*, 1998).

We first analyzed the conjugation efficiency of pLE12 between two *L. lactis* strains. The donor strain (LM0230) contained the pLE12 vector and the recipient strain (MMS372) the pMNHS plasmid, harbouring the intron homing site (HS) (LM0230/pLE12 X MMS372/pMNHS). Using milk-plate conjugation assays, we noticed, as expected from previous studies (Mills *et al.*, 1994), that the Tra1-2 region is not sufficient to significantly promote conjugation of its host plasmid. Only a small increase in conjugation was observed between the backbone plasmid pLE1 and pLE12 (Table 2.2, 3.3 fold). Moreover, the overall conjugation efficiencies were relatively low compared to what was previously observed for pRS01 transfer ($10^{-6}$ vs $10^{-3}$) (Mills *et al.*, 1994) suggesting that some conjugative factors, likely those expressed from the Tra3 and Tra4 regions, were missing.

In order to study Ll.LtrB transfer from pLE12, but in the context of its

complete conjugative element (pRS01 or sex factor), we analyzed pLE12 transfer from *L. lactis* donor strains harbouring a chromosomal copy of the sex factor. These strains provided, by complementation, all the conjugation machinery necessary for pLE12 conjugative transfer. The *L. lactis* NZ9800 strain was derived from MG1363, a strain harbouring a chromosomal copy of the sex factor (Shearman *et al.*, 1996). We confirmed the presence of the sex factor within the chromosome of the NZ9800 strain, and its absence from the LM0230 and MMS372 strains by PCR and by Southern blots (data not shown). Using NZ9800 as the donor strain and either LM0230 or MMS372 as the recipient strain, we observed, in both cases, a significant increase in conjugation efficiencies between pLE1 ($10^{-8}$) and pLE12 ($10^{-3}$) (Table 2.2). These experiments showed that the sex factor within the NZ9800 chromosome complements pLE12 transfer by conjugation and that the transfer of pLE1 is significantly increased by the presence of the Tra1-2 region from pRS01 (100,000 fold, $10^{-8}$ to $10^{-3}$) (Table 2.2), possibly by the recognition of the transfer origin (*oriT*). Accordingly, we saw a dramatic drop in the conjugation efficiency of the pLE12 derivative harbouring a splicing-deficient intron (pLE12Mat$^-$; no relaxase (LtrB)) from the NZ9800Δ*ltrB*::Tet$^R$ strain ($10^{-8}$) but not from NZ9800 ($10^{-3}$) (Table 2.2). NZ9800 thus provides the relaxase from its chromosome and drives the conjugation of the pLE12Mat$^-$ plasmid that is unable to produce its own relaxase. In fact, all the intron mutants, including the non-splicing variants (Mat$^-$, ΔORF, ΔD5), show efficient conjugation levels from the NZ9800 strain (Table 2.2). A modest but reproducible drop (≈ 10 fold) in conjugation efficiency was also noticed for pLE12 when only the plasmid-encoded relaxase

gene is present (NZ9800Δ*ltrB*) (Table 2.2).

To further verify that the recipient cells harbouring both plasmids (transconjugants) acquired pLE12 by conjugation, we performed 6 independent conjugation assays (NZ9800/pLE12 X LM0230/pMNHS) (Figure 2.1B). For one half of the assays, a cocktail of DNase I and RNase was applied onto the mating plates (Trieu-Cuot *et al.*, 1998). Conjugation efficiencies were identical when supplemented or non-supplemented plates were used (Table 2.2), confirming that pLE12 acquisition by the recipient cells is achieved through conjugation.

### 2.4.2 Transfer of the Ll.LtrB group II intron between *L. lactis* strains

To study whether the newly transferred Ll.LtrB intron was still expressed, active, and proficient in invading its homing site following its conjugative transfer, we recovered and analyzed the plasmid mix from 10 independent transconjugants (Figure 2.1B). The presence of the homing site-containing plasmid pMNHS within the recipient cells gave us the opportunity to directly assess Ll.LtrB mobility from pLE12 after its mobilization by conjugation. These plasmids were first analyzed by agarose gel electrophoresis, where both donor (D) and recipient (R) plasmids could be seen and distinguished (Figure 2.2A) confirming that the $Cam^R/Spc^R$ recipient cells were indeed transconjugants. We also noticed the presence of an additional characteristic band, corresponding to an intron-interrupted recipient plasmid or mobility event (pMNHS + Ll.LtrB) (Figure 2.2A) (Cousineau *et al.*, 1998). This additional band and the donor plasmid pLE12 were both absent from recipient cells carrying only the recipient plasmid (LM0230/pMNHS) (Figure 2.2A). We

confirmed that the additional band corresponds to Ll.LtrB mobility events by Southern blot, using a Ll.LtrB-specific probe (Figure 2.2B). The $^{32}$P-labeled intron probe annealed to the intron-containing recipient plasmid (M) and to both the donor plasmid (D; pLE12) and positive control (M; pMNHS + Ll.LtrB) and did not hybridize to the uninterrupted recipient plasmid (R; pMNHS).

Using patch-hybridization assays (Figure 2.1D) (Cousineau *et al.*, 1998), we scored the mobility efficiency of the Ll.LtrB intron for the 10 independent mobility assays and calculated an average efficiency of 10.5% (Table 2.3). The mobility efficiency of the Ll.LtrB intron after conjugation of pLE12 is slightly higher (2.4-fold) than previously observed when the pLE12 and pMNHS plasmids were co-transformed in the same strain (LM0230) (Cousineau *et al.*, 1998).

### 2.4.3 Retrohoming of Ll.LtrB after conjugation

To determine if Ll.LtrB invades its homing site via retrohoming, we constructed a series of Ll.LtrB mutants (Figure 2.3) (Cousineau *et al.*, 1998, 2000) in pLE12 and calculated resulting intron mobility efficiencies following conjugation from NZ9800 to LM0230/pMNHS recipients (Table 2.3). Two of these constructs were artificial twintrons (Figure 2.3; pLE12I, pLE12IK) where the phage T4 group I *td* intron was present within domain IV of Ll.LtrB, downstream of LtrA. This autocatalytic group I intron splices only at the RNA level. The ratio of mobility products compared to the total amount of recipient plasmids were calculated by patch-hybridization assays for the different Ll.LtrB intron variants. Mobility products were identified using an intron specific probe (gpII) while a group I splice junction probe (gpI SJ) was used to assess if the mobility products

had lost the group I intron (gpI loss). As previously observed (Cousineau *et al.*, 1998), the twintron constructs, carrying additional sequences, are not as proficient as the wild-type intron (Table 2.3). However, mobility events obtained after conjugation of the Ll.LtrB twintron constructs (pLE12I, pLE12IK) showed conclusively that Ll.LtrB invades its homing site using an RNA intermediate, since all the mobility events had lost the retromobility indicator *td* intron (Table 2.3, columns gpISJ and gpI loss) (Cousineau *et al.*, 1998, 2000; Ichiyanagi *et al.*, 2002). If Ll.LtrB mobility had proceeded through a DNA-based pathway, the group I intron would still have been present within Ll.LtrB after its insertion. Specific mutants of the intron-encoded protein (IEP) reverse transcriptase (RT$^-$), maturase (Mat$^-$), and LtrA$^-$ ($\Delta$ORF)), and of the autocatalytic core of the intron ($\Delta$D5) (Matsuura et al., 1997) were also analyzed (Figure 2.3). These mutations reduced the mobility efficiency of Ll.LtrB below our detection limit for this assay (>10 fold; $<10^{-2}$) (Table 2.3). However, we detected some mobility of the endonuclease mutant (Endo$^-$). Its efficiency was reduced 9.5-fold compared to the wild-type Ll.LtrB. Interestingly, this Ll.LtrB variant, which is also known to be deficient in RT activity (San Filippo and Lambowitz, 2002), was previously shown to be virtually immobile (59-fold lower) in *E. coli* (Cousineau *et al.*, 1998). The analysis of this Ll.LtrB mutant in its original host environment (*L. lactis* vs *E. coli*) during and/or following conjugation of its carrying plasmid could account for the difference observed in mobility efficiencies. Nevertheless, this result is not completely unexpected, since some bacterial group II introns harbour IEPs lacking an endonuclease domain (Martinez-Abarca and Toro, 2000c; Dai and Zimmerly,

2002a), one of which, the RmInt1 intron from *Sinorhizobium meliloti*, has already been shown to be mobile (Martinez-Abarca and Toro, 2000a; 2000b).

We also evaluated the mobility efficiency of some of the Ll.LtrB constructs after their transfer by conjugation from NZ9800 to MMS372, a *recA*$^-$ isogenic strain of LM0230. We noticed that although the conjugation levels are comparable (Table 2.2), the Ll.LtrB retrohoming efficiency was higher when using MMS372 as the recipient strain (Table 2.3, pLE12 (WT) and pLE12Endo$^-$). A similar mobility increase of Ll.LtrB in MMS372 (*recA*$^-$) cells compared to LM0230 (*recA*$^+$) was previously observed when pLE12 and pMNHS were co-transformed in these same strains (Mills *et al.*, 1997; Cousineau *et al.*, 1998). These data show that the homologous recombination system is not involved in Ll.LtrB mobility during and/or following conjugation. We also noticed that the mobility efficiency of the RT (RT$^-$), and maturase (Mat$^-$) mutants of LtrA did not vary between these strains.

While scoring the mobility efficiency of the different Ll.LtrB constructs following conjugation, we observed unexpectedly high mobility levels for some of the 10 independent events examined (Table 2.3, underlined values). These events, often much more proficient than the wild-type intron, were observed in different conditions for all constructs studied (Figure 2.3 and Table 2.3). The detailed mechanism promoting these mobility events is currently under investigation.

Taken together, these results show that Ll.LtrB is retrohoming to its new location during and/or following conjugation of its host plasmid, and that the three functions of LtrA (RT, maturase, endonuclease) are involved in its mobility pathway. The complementation/conjugation system where the relaxase enzyme

(LtrB) expressed from the chromosome of the donor strain gave us the opportunity to study the transfer of different intron mutants, between *L. lactis* strains, after conjugation of their host plasmid. The level of LtrB produced from the chromosome is sufficient to sustain the transfer of our pLE12 variants to wild-type conjugation levels (Table 2.2; $10^{-3}$).

### 2.4.4 Ll.LtrB retrotransposition after its transfer by conjugation

In order to determine if the Ll.LtrB group II intron was able to invade ectopic sites following its transfer by conjugation, we performed conjugation assays between *L. lactis* strains (NZ9800/pGNIK X NZ9800Δ*ltrB*::$Cam^R$) where the recipient strain was plasmid free (Figure 2.1C). We then looked, after conjugation of the donor plasmid, for chromosomal insertions of Ll.LtrB in recipient cells.

In this conjugation/retrotransposition assay (Figure 2.1C), the donor plasmid pGNIK (pG+host5 based) contains a temperature sensitive origin of replication (Ts). This plasmid replicates normally in cells grown at the permissive temperature (30°C) but cannot replicate and gets diluted out of cultures grown at 37°C. This feature allowed us to select for retrotransposition events ($Kan^R$) following the loss of the intron-carrying plasmid. The Ll.LtrB variant we used for this assay harboured the *td* group I intron and the $Kan^R$ gene, and was only flanked by small portions of both exons (Figure 2.3, pGNIK). Although pGNIK is a non-conjugative plasmid, we were nevertheless able to obtain transconjugants (NZ9800Δ*ltrB*/pGNIK, $Cam^R/Erm^R$). The conjugation frequencies were comparable, when the mating plates were supplemented ($3.3 \pm 0.6 \times 10^{-8}$) or not

($2.6 \pm 0.4 \times 10^{-8}$) with the DNase I/RNase cocktail (3 assays each). Despite the low conjugation efficiencies, the DNase I/RNase control suggests that pGNIK is indeed transferred between the *L. lactis* strains by conjugation.

We selected for Ll.LtrB insertions within the *L. lactis* chromosome of transconjugants following loss of the donor plasmid upon temperature shift. Kanamycin-resistant colonies were obtained even if intron expression was not induced with nisin. Using Southern blots (Ll.LtrB probe) we showed that these independently isolated recipient strains contained only one Ll.LtrB insertion per genome at five different sites. Moreover, using the *td* group I intron splice junction probe we found that the *td* intron was precisely spliced out in all cases (Figure 2.4A). The absence of the *td* intron from these chromosomal insertions suggests that Ll.LtrB invaded the chromosome of the recipient cell by retrotransposition (ectopic sites, RNA intermediate). Taking advantage of the kanamycin gene present within the newly inserted introns, we cloned and sequenced the five independent Ll.LtrB chromosomal insertion sites (Figure 2.4B). The sequences confirmed that these Ll.LtrB insertions were retrotransposition events and that they were new insertion sites never observed in previous retrotransposition studies (Cousineau *et al.*, 2000; Ichiyanagi *et al.*, 2002). The homology between these sites and the wild-type homing site are, as expected, confined to a short 13-17 base pair region spanning the intron insertion sites. This suggests that the intron invaded these target sites through reverse splicing (Cousineau *et al.*, 2000; Ichiyanagi *et al.*, 2002). As anticipated from previous retrotransposition studies using this Ll.LtrB variant (pGNIK) (Cousineau *et al.*, 2000; Ichiyanagi *et al.*, 2002), the intron was inserted, in all five cases, in the same orientation as the interrupted genes.

These data demonstrate that the Ll.LtrB group II intron can be widely spread, following its transfer by conjugation, invading non-homologous sites within the chromosome of the recipient strain by retrotransposition. Ll.LtrB is thus not restrained to move exclusively by retrohoming to a very specific site but can also invade a multitude of sites within its new host's chromosome. Moreover, we showed that Ll.LtrB can be transferred even when present on a non-conjugative plasmid. This situation is probably restricted to donor cells housing some minimal transfer functions allowing cell-cell contacts and the formation of a mating channel.

### 2.4.5 Lateral transfer of the Ll.LtrB intron between Gram-Positive bacteria

To further demonstrate the biological relevance of Ll.LtrB lateral transfer, we asked whether pLE12 could relocate from *L. lactis* to other bacterial species by conjugation. We thus looked at the conjugation of pLE12 from *L. lactis* to *Enterococcus faecalis,* another low-GC Gram-positive bacterium, by filter mating assays (Sasaki *et al.*, 1998). Using the JH2-2 *E. faecalis* strain as the recipient strain (NZ9800/pLE12 X JH2-2/pMNHS), we obtained transconjugants at a frequency of $10^{-7}$ (Table 2.2). Although the conjugation frequencies from *L. lactis* to *E. faecalis* ($10^{-7}$) are much lower than those observed between *L. lactis* strains ($10^{-3}$), the DNase I/RNase control assay confirmed that pLE12 was transferred to *E. faecalis* by conjugation (Table 2.2).

Using the pLE12 and pLE12IK constructs, we detected lateral transfer of the Ll.LtrB group II intron by retrohoming with efficiencies of 6 ± 1% and 1.2 ± 0.2% respectively. Again, the mobility efficiency of the group I/Kan$^R$-containing

construct is lower than the wild-type intron and they are both less efficient in *E. faecalis* than in *L. lactis*. The presence of *L. lactis* promoters could explain the mobility difference in the two cell backgrounds, with lower expression of the *ltrB* and/or *ltrA* gene leading to lower levels of active RNPs in *E. faecalis*. However, the absence of the group I intron from the great majority of the mobility events (67%) demonstrates that the intron is invading its homing site via the retrohoming pathway in *E. faecalis*. These results confirm that plasmid conjugation can support the lateral or horizontal transfer of Ll.LtrB in the Gram-positive branch of the bacterial kingdom.

## 2.5 Discussion

In this study we developed conjugation/retrohoming and conjugation/retrotransposition assays to analyze the lateral transfer of the Ll.LtrB group II intron from *L. lactis*. We have presented the first experimental demonstration that bacterial group II introns can be laterally transferred in the bacterial kingdom using both retrohoming (*L. lactis* to *L. lactis* and *L. lactis* to *E. faecalis*) and retrotransposition (*L. lactis* to *L. lactis*) pathways during and/or following their transfer by conjugation. This work demonstrates that group II introns can invade either resident plasmids or the chromosome of the recipient strain following their transfer by conjugation. We also showed that group II introns can be laterally transferred by conjugation even when present on non-conjugative plasmids if the donor cells house some minimal transfer functions allowing cell-cell contacts and the formation of a mating channel.

The data presented are biologically relevant for the following reasons. First,

Ll.LtrB was originally discovered in *L. lactis* on a natural conjugative plasmid (pRS01) and within a chromosomally located sex factor (Mills *et al.*, 1996; Shearman *et al.*, 1996). These two elements were shown to be actively transferred by conjugation between *L. lactis* strains and Gram-positive bacteria (Mills *et al.*, 1996; Shearman *et al.*, 1996). Moreover, it was proposed that pRS01 can be transferred by conjugation from *L. lactis* to other Gram-positive genera like streptococci and lactobacilli (Mills *et al.*, 1996), indicating a possible vast dissemination of Ll.LtrB by conjugation within the Gram-positive bacterial branch. Second, in our conjugation/mobility assays, natural levels of active RNPs were produced since the expression of the *ltrB* gene, interrupted by the intron, and *ltrA* present within the intron were under the control of their native promoters (not overexpressed). Third, Ll.LtrB can be efficiently disseminated following its transfer by conjugation, through retrohoming, not only between closely related *L. lactis* strains, but also from *L. lactis* to other Gram-positive bacteria (*E. faecalis*). Fourth, Ll.LtrB can also invade multiple chromosomal sites by retrotransposition into a recipient cell following its transfer by conjugation. Fifth, Ll.LtrB splicing directly controls the expression level of the relaxase gene (*ltrB*), essential for conjugative transfer, thus controlling the conjugation of its host element and at the same time its own dissemination and survival.

Taken together, these results suggest that conjugation of these broad host range mobile elements (pRS01 and sex factor) should happen in nature and may support a wider dissemination of Ll.LtrB than we demonstrated using pLE12 in a constrained laboratory setting. Plasmid pRS01 and the *L. lactis* sex factor can thus be considered as infectious elements driving the horizontal transfer and spread of

their associated mobile intron.

Since conjugation is the most efficient way to transfer genetic information between widely diverged bacterial species and even across phyla (Lambowitz and Belfort, 1993), conjugation may have been and probably still is an important means of intron dispersal.

Interestingly, the great majority of group II introns found in bacteria are associated with various mobile elements, such as IS elements, transposons, conjugative plasmids, pathogenicity islands, and virulence plasmids (Martinez-Abarca and Toro, 2000; Dai and Zimmerly, 2002a). The results presented in this study suggest that these introns could also be disseminated and spread by horizontal transfer in the bacterial kingdom, following the transfer of their host mobile elements. The association of some group II introns with other mobile elements may have been a means of survival for these introns and could explain why they are still present and unexpectedly highly represented within contemporary bacterial genomes.

## 2.6 Experimental procedures

### 2.6.1 Strains and plasmids

The strains and plasmids used in this study are described in Table 2.1. *L. lactis* and *E. faecalis* strains were grown without shaking in M17 supplemented with 0.5% glucose (GM17) (30°C) and BHI (brain heart infusion) media (37°C) respectively. *E. coli* strains (DH5α, DH10β), used for cloning and mobility scoring, were grown with shaking at 37°C in LB broth. The milk plates used in conjugation assays were made of 5% dry milk (Carnation), 1% dextrose, and 1.5% agar. The

pLE12 plasmid consists of the Tral-2 region from pRS01 (7.5 kb, *Pst* I, Figure 2.1A) cloned into the pLE1 vector at its unique *Pst* I site. The pMNHS plasmid (pMN1343) contains a 271 base pairs homing site (exon 1: 179 bp , exon 2: 92 bp) (*Hin*d III) inserted at the unique *Hin*d III restriction site in the pDL278 vector. The pLE12-based constructs harbouring different Ll.LtrB mutants were made by replacement cloning using either *Bsr*G I and *Bsi*W I or *Bsr*G I and *Kpn* I restriction enzymes. Ll.LtrB mutants are as follows: RT$^-$, reverse transcriptase mutant (YADD to YAAA) (Matsuura *et al.*, 1997); Mat$^-$, maturase defective ($SC_{463}$ to LA) (Matsuura *et al.*, 1997; Cousineau *et al.*, 1998); Endo$^-$, endonuclease domain deletion (amino acids 543-599) (Matsuura *et al.*, 1997); ΔORF, amino acids 40-572 in LtrA were replaced by TR (RT$^-$, Mat$^-$, Endo$^-$) (Matsuura *et al.*, 1997); ΔD5, domain V of the intron was deleted (non-splicing) (Matsuura *et al.*, 1997). The pGNIK plasmid is a pG$^+$host5 based construct (Ts) carrying the Ll.LtrB intron (*Hin*d III, *td* intron, Kan$^R$) under the control of the nisin inducible promoter. Selective medium contained the following concentrations of antibiotics: chloramphenicol, 5 or 10μg/ml; spectinomycin, 300μg/ml; erythromycin, 300μg/ml; and kanamycin, 20μg/ml; fusidic acid 25μg/ml; tetracycline, 3μg/ml.

### 2.6.2 Conjugation assays

*L. lactis* strains (donor, recipient) were diluted (0.4 or 0.8 ml/10 ml) from overnight (ON) saturated cultures and grown for 7 hours at 30°C with the appropriate antibiotics. Cells were recovered by centrifugation, the pellets were mixed (1:1), spread on milk plates, and incubated at 30°C for 12 hours. The cell

mixtures were recovered with PBS 1X (1 ml) and serial dilutions were made to score donor ($Cam^R$), recipient ($Spc^R$), and transconjugant ($Cam^R/Spc^R$) cells (Mills *et al.*, 1996).

Conjugation assays from *L. lactis* (NZ9800) to *E. faecalis* (JH2-2) were done on GM17 plates (filter mating) at 37°C (Sasaki *et al.*, 1998). The identity of the recipient strain (JH2-2) containing both donor and recipient plasmids was confirmed by its resistance to fusidic acid (25 µg/ml). The conjugation efficiencies (3 assays) were calculated as the number of transconjugants/donor cells ($Cam^R$-$Spc^R/Cam^R$).

In order to perform conjugation controls, a cocktail of DNase I and RNase (100 U/ml each) (Trieu-Cuot *et al.*, 1998) was applied onto the mating plates and the conjugation efficiencies calculated as previously described. The typical conjugation control experiment uses only DNase I, but since the Ll.LtrB intron is a retroelement, we also controlled for the possible non-conjugative uptake of active RNP particles (intron RNA + intron-encoded protein (IEP)).

### 2.6.3 Mobility assay (patch-hybridization)

The plasmid mixtures (donor, recipient, and mobility products) from 10 independent transconjugants were prepared (same conjugation assay), retransformed into *E. coli* (DH5α), and plated on LB/Spc plates to select for recipient plasmids interrupted or not by the intron (Figure 2.1D). In order to calculate the percentage of recipient plasmids that received the intron (mobility products), 100 isolated colonies ($Spc^R$) were patched for each independent assay

(LB/Spc plates). The patches were lifted on nylon membranes and hybridized with the appropriate 5'-$^{32}$P end labeled probes for homing (Ll.LtrB probe, 5'-GTAT-GGCTATGCCCGGAATAC-3' (3' end of the intron) or 5'-CCGTGCTCTG-TTCCCGTATCAGC-3' (5' end of the intron)) and retrohoming (*td* intron splice junction probe, 5'-ATTAAACGGTAGACCCAAGAAAAC-3') (Cousineau *et al.*, 1998). The *td* intron splice junction probe recognizes the two *td* ligated exons (12 nt each exon) flanking it and gives a positive signal only upon group I intron loss in mobility products (Belfort *et al.*, 1990).

### 2.6.4 Retrotransposition assay

The conjugative transfer of pGNIK from NZ9800 to NZ9800Δ*ltrB*::Cam$^R$ was performed as described above. Following conjugation, a transconjugant colony (NZ9800Δ*ltrB*/pGNIK, Cam$^R$/Erm$^R$) was grown ON in GM17 at 37°C to prevent replication of the pGNIK Ts-vector. Two successive ON cultures at 37°C (1/1000) were done followed by a final ON culture in GM17/Kan at 37°C (1/1000) to select for Ll.LtrB chromosomal insertions (retrotransposition events (RTP)). The cells were then diluted, plated on GM17/Kan at 30°C, and isolated colonies were picked and grown ON (GM17/Kan, 30°C). These cultures were split for plasmid and genomic DNA isolation. Plasmid preparations confirmed that pGNIK was lost during growth at 37°C. Genomic DNAs were digested (*Spe* I or *Nco* I) and hybridized with the *td* splice junction probe and showed only one signal at different positions (Figure 2.4A). This result was confirmed, again by Southern blot (*Spe* I), using a Ll.LtrB specific probe (5'- CCGTGCTCTGTTCCCGTATCAGC-3'). To isolate and characterize the RTP events, genomic DNA was digested with either

*Spe* I, *Hin*d III, or *Pvu* II, cloned in the pBS vector, and selected for on LB/Kan plates (DH5α or DH10β). Both 5' and 3' junctions between Ll.LtrB and its five independent insertion sites were obtained by sequencing using intron specific primers (5'-CCGTGCTCTGTTCCCGTATCAGC-3', 5'-CAGAGCCGTATACT-CCGAG-3') (Figure 2.4B). Genomic DNAs from the RTP events were also digested with *Pst* I and hybridized with an exon 2 probe (5'-GTGAGAGTTACCT-GGAGACT-3') to confirm the identity of the recipient genomic DNA (NZ9800Δ*ltrB*::Cam$^R$) (data not shown).

## 2.7 Acknowledgements

We thank M. Belfort for kindly providing the NZ9800Δ*ltrB*::Tet$^R$ strain. We also thank M. Belfort, K. Fiola, R. Lease, J. Wyse, and S. Zimmerly for providing comments on the manuscript, and A. Villeneuve for technical assistance. B.C. is a CIHR New Investigator Scholar and a McGill William Dawson Scholar. This work was supported by an NSERC grant to B.C.

## 2.8 References

**Anderson, D.G. and McKay, L.L.** (1984) Genetic and physical characterization of recombinant plasmids associated with cell aggregation and high-frequency conjugal transfer in *Streptococcus lactis* ML3. *J. Bacteriol.* **158:** 954-962.

**Belfort, M., Ehrenman, K. and Chandry, P.S.** (1990) Genetic and molecular

analysis of RNA splicing in *Escherichia coli*. *Methods Enzymol.* **181:** 521-539.

**Belfort, M., Derbyshire, V., Parker, M.M., Cousineau, B. and Lambowitz, A. M.** (2002) Mobile introns: pathways and proteins. in *Mobile DNA II*, Craig, N. L., Craigie, R., Gellert, M. and Lambowitz, A. M. (eds). Washington DC: American Society for Microbiology Press, pp. 761-783.

**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.

**Cousineau, B., Lawrence, S., Smith, D. and Belfort, M.** (2000) Retrotransposition of a bacterial group II intron. *Nature* **404:** 1018-1021. See correction, *Nature* **414:** 84 (2001). Erratum in (2001) *Nature* **414:** 84.

**Dai, L. and Zimmerly, S.** (2002a) Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res.* **30:** 1091-1102.

**Dai, L. and Zimmerly, S.** (2002b) The dispersal of five group II introns among natural populations of Escherichia coli. *RNA*. **8:** 1294-1307. Added in thesis

**Fernandez-Lopez, M., Munoz-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. Added in thesis

**Ichiyanagi, K., Beauregard, A., Lawrence, S., Smith, D., Cousineau, B. and Belfort, M.** (2002) Retrotransposition of the Ll.LtrB group II intron procedes predominantly via reverse splicing into DNA targets. *Mol. Microbiol.* **46:** 1259-1272.

**Jacob, A.E. and Hobbs, S.J.** (1974) Conjugal transfer of plasmid-borne multiple antibiotic resistance in *Streptococcus faecalis* var. zymogenes. *J. Bacteriol.* **117:** 360-372.

**Kuipers, O.P., Beerthuyzen, M.M., Siezen, R.J. and De Vos, W.M.** (1993) Characterization of the nisin gene cluster nisABTCIPR of *Lactococcus lactis*. Requirement of expression of the *nisA* and *nisI* genes for development of immunity. *Eur. J. Biochem.* **216:** 281-291.

**Lambowitz, A.M. and Belfort, M.** (1993) Introns as mobile genetic elements. *Ann. Rev. Biochem.* **62:** 587-622.

**Martinez-Abarca, F., Garcia-Rodriguez, F.M. and Toro, N.** (2000a) Homing of a bacterial group II intron with an intron-encoded protein lacking a recognizable endonuclease domain. *Mol. Microbiol.* **35:** 1405-1412.

**Martinez-Abarca, F. and Toro, N.** (2000b) RecA-independent ectopic transposition *in vivo* of a bacterial group II intron. *Nucleic Acids Res.* **28:**

4397-4402.

**Martinez-Abarca, F. and Toro, N.** (2000c) Group II introns in the bacterial world. *Mol. Microbiol.* **38:** 917-926.

**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.

**McKay, L.L., Baldwin, K.A. and Walsh, P.M.** (1980) Conjugal transfer of genetic information in group N streptococci. *Appl. Environ. Microbiol.* **40:** 84-91.

**Michel, F. and Ferat, J.-L.** (1995) Structure and activities of group II introns. *Ann. Rev. Biochem.* **64:** 435-461.

**Mills, D.A., Choi, C. K., Dunny, G. M. and McKay, L.L.** (1994) Genetic analysis of regions of the *Lactococcus lactis* subsp. *lactis* plasmid pRS01 involved in conjugative transfer. *Appl. Environ. Microbiol.* **60:** 4413-4420.

**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.

**Mills, D.A., Manias, D.A., McKay, L.L. and Dunny, G.M.** (1997) Homing of a group II intron from *Lactococcus lactis* subsp. *lactis* ML3. *J. Bacteriol.* **179:**

6107-6111.

**Mills, D., Phister, T.G., Dunny, G.M. and McKay, L.L.** (1998) An origin of transfer (*oriT*) on the conjugative element pRS01 from *Lactococcus lactis* subsp. *lactis* ML3. *Appl. Environ. Microbiol.* **64:** 1541-1544.

**Saldanha, R., Mohr, G., Belfort, M. and Lambowitz, A.M.** (1993) Group I and group II introns. *FASEB Journal* **7:** 15-24.

**San Filippo, J. and Lambowitz, A.M.** (2002) Characterization of the C-terminal DNA-binding/DNA endonuclease region of a group II intron-incoded protein. *J. Mol. Biol.* **324:** 933-951.

**Sasaki, Y., Taketomo, N. and Sasaki, T.** (1998) Factors affecting transfer frequency of pAMß1 from *Streptococcus faecalis* to *Lactobacillus plantarum. J. Bacteriol.* **170:** 5939-5942.

**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.

**Trieu-Cuot, P., Carlier, C. and Courvalin, P.** (1998) Conjugative plasmid transfer from *Enterococcus faecalis* to *Escherichia coli. J. Bacteriol.* **170:** 4388-4391.

**Zhou, L., Manias, D.A. and Dunny, G.M.** (2000) Regulation of intron function:

efficient splicing *in vivo* of a bacterial group II intron requires a functional promoter within the intron. *Mol. Microbiol.* **37:** 639-651.

**Zimmerly, S., Hausner, G. and Wu, X.** (2001) Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res.* **29:** 1238-1250.

## 2.9 Figure Legends

### Figure 2.1

Conjugation and mobility assays. **A.** Schematic of the Tral-2 region from the *L. lactis* conjugative plasmid pRS01 (adapted from Mills *et al.*, 1996). The *Pst* I fragment (7.5 kb, Tral-2 region) from pRS01 present in the pLE12 vector, and the Ll.LtrB group II intron are shown. The conjugative origin of transfer (*oriT*) is also represented (black circle). **B.** Conjugation/retrohoming assay between two *L. lactis* strains. The first step (1) represents the transfer of pLE12 by conjugation from the donor to the recipient strain. The second step (2) shows the invasion (retrohoming) of the homing site (E1/E2) present on the recipient plasmid (pMNHS) by the Ll.LtrB intron expressed from the donor plasmid (pLE12). **C.** Conjugation/retrotransposition assay between two *L. lactis* strains. The first step (1) represents the transfer of pGNIK by conjugation from the donor to the recipient strain. The second step (2) illustrates Ll.LtrB invasion of the recipient strain's chromosome by retrotransposition (group I'). The loss of the intron-carrying plasmid (pGNIK, temperature sensitive (Ts)) from transconjugants, following a temperature shift, allowed the selection of Ll.LtrB chromosomal insertions ($Kan^R$).

The chromosome is only depicted for the recipient strain. **D.** Scoring of intron mobility efficiency (patch-hybridization) after conjugation. The plasmid mix (donor, recipient, and mobility products) from 10 independent transconjugants are recovered by plasmid preparation (1) and retransformed in *E. coli* DH5α cells (LB/Spc) (2). 100 individual colonies, containing either mobility events or the uninterrupted recipient plasmid, are patched on LB/Spc plates (3) and transferred onto a nylon membrane (4). The nylon membrane is then hybridized with an intron-specific probe to calculate the ratio of recipient plasmids that received the intron. Ll.LtrB group II intron, gray; E1 and E2, Ll.LtrB exons.

**Figure 2.2**

Lateral transfer of Ll.LtrB. **A.** Agarose gel (0.5%) containing the plasmid mix (undigested) recovered from 10 recipients ($Spc^R$) and 10 transconjugants ($Spc^R/Cam^R$) following pLE12 conjugation (NZ9800/pLE12 X LM0230/pMNHS, Fig 1B). Presence of the donor (D), recipient (R), and mobility product (M) are indicated. **B.** Southern blot of the agarose gel shown in panel A (Cousineau *et al.*, 1998).

**Figure 2.3**

Different constructs of the Ll.LtrB group II intron. Ll.LtrB group II intron, gray; phage T4 group I *td* intron, black; Ll.LtrB exons, E1 and E2; wild type Ll.LtrB group II intron, WT.

**Figure 2.4**

Retrotransposition sites. **A.** Southern blot on digested genomic DNA (*Nco* I) from five independent $Kan^R$/pGNIK$^-$ NZ9800Δ*ltrB*::$Cam^R$ colonies (1-5) and from NZ9800Δ*ltrB*::$Cam^R$ (NZΔL, negative control). The 32P-labelled *td* splice junction probe revealed only one signal per lane all at different positions (1-5). **B.** Sequences of five independent Ll.LtrB retrotransposition sites isolated from the chromosome of the NZ9800Δ*ltrB*::$Cam^R$ strain are shown and compared to the WT retrohoming site (HS). White nucleotides on black background represent identity with HS or potential G-U base pairs (asterisk) with intron RNA. The arrowhead indicates the Ll.LtrB insertion site. The interrupted genes and the potential interactions between the intron RNA and its substrate at the insertion sites are illustrated (IBS1/EBS1, IBS2/EBS2, δ/δ'). Intron binding sites, IBS1 and IBS2; exon binding sites, EBS1 and EBS2.

## 2.10 Tables

**Table 2.1.** Bacterial strains and plasmids

| Strain or plasmid | Relevant characteristics[a] | Source, description, and/or reference |
|---|---|---|
| *L. lactis* strains | | |
| LM0230 | Plasmid free, $Rec^+$ | Donor and recipient strain for conjugation assays[b] |
| MMS372 | Plasmid free, $Rec^-$ | Recipient strain for conjugation assays, isogenic to LM0230[c] |
| NZ9800 (Ll.LtrB intron)[d] | Plasmid free, $Rec^+$ | Donor strain for conjugation assays, chromosomal sex-factor |
| NZ9800Δ*ltrB* (*Hind* III)[e] | $Cam^R$, plasmid free, $Rec^+$ | The Ll.LtrB intron was deleted with parts of its flanking exons |
| NZ9800Δ*ltrB* (*Hind* III)[f] | $Tet^R$, plasmid free, $Rec^+$ | The Ll.LtrB intron was deleted with parts of its flanking exons |
| *E. faecalis* strain | | |
| JH2-2 | Plasmid free, $Rif^R$, $Fus^R$ | Recipient strain for conjugation assays (*L. lactis* to *E. faecalis*)[g] |
| Plasmids | | |
| pLE1 | $Cam^R$, 8.7 kb | Backbone shuttle plasmid (Gram+/Gram-)[h] |
| pLE12 | $Cam^R$, 16.2 kb | pLE1 containing the Tra1-2 region from pRS01 (Ll.LtrB; 7.5 kb)[i] |
| pLE12I | $Cam^R$, 16.6 kb | Ll.LtrB contains the *td* group I intron in domain IV[j] |
| pLEIK | $Cam^R$, 17.7 kb | Ll.LtrB contains the *td* group I intron and $Kan^R$ gene in domain IV |
| pLE12RT- | $Cam^R$, 16.2 kb | Ll.LtrB carrying LtrA with a mutated reverse transcriptase domain |
| pLE12 Mat- | $Cam^R$, 16.2 kb | Ll.LtrB carrying LtrA with a mutated maturase domain |
| pLE12Endo- | $Cam^R$, 16.0 kb | Ll.LtrB carrying LtrA missing the endonuclease domain |
| pLE12ΔORF | $Cam^R$, 14.6 kb | Ll.LtrB carrying a large deletion within LtrA ($RT^-$, $Mat^-$, $Endo^-$) |
| pLE12ΔD5 | $Cam^R$, 16.2 kb | Ll.LtrB missing domain V, the LtrA protein is intact |
| pMNHS | $Spc^R$, 6.9 kb | Shuttle plasmid pDL278 containing the Ll.LtrB homing site (*Hind* III, 271 bp)[k] |
| pMNHS/$Erm^R$ (*Kpn* 1) | $Erm^R$, $Spc^S$, 7.7 kb | The Erm gene (PCR) was inserted in the middle of the $Spc^R$ gene |
| pGNIK | $Erm^R$, $Kan^R$, 11.2 kb | Ts plasmid, Ll.LtrB intron (*td* intron, $Kan^R$ ) is flanked by short exons (*Hind* III), intron is located downstream from the nisin inducible promoter[e,f] |

[a]$Cam^R$, chloramphenicol resistant ; $Erm^R$, erythromycin resistant ; $Spc^R$, spectinomycin resistant ; $Kan^R$, kanamycin resistant ; $Fus^R$, acid fusidic resistant ; $Tet^R$, tetracyclin resistant ; $Rif^R$, rifampicin resistant ; Rec, host recombination status.
[b]McKay *et al.*, 1980 ; [c]Anderson and McKay, 1984 ; [d]Kuipers *et al.*, 1993 ; [e]Cousineau *et al.*, 2000 ; [f]Ichiyanagi *et al.*, 2002 ; [g]Jacob and Hobbs, 1974 ; [h]Mills *et al.*, 1994 ; [i]Mills *et al.*, 1996 ; [j]Cousineau *et al.*, 1998 ; [k]Mills *et al.*, 1997.

**Table 2.2.** Conjugation efficiency of Ll.LtrB-carrying plasmids

| Donor plasmid | LM0230 to MMS372 ($10^{-6}$) | NZΔ*ltrB*[c] to MMS372 ($10^{-4}$) | NZ9800 to LM0230 ($10^{-3}$) | NZ9800 to MMS372 ($10^{-3}$) | NZ9800 to JH2-2 ($10^{-7}$) |
|---|---|---|---|---|---|
| pLE1 | 1.6 ± 0.3 | --- | 4.9 ±1.2* | --- | --- |
| pLE12 | 5.2 ± 0.3 | 2.2 ± 0.6 | 2.3 ± 0.6 | 2.3 ± 0.5 | 1.6 ± 0.3 |
| pLE12[a] | | | 2.7 ± 0.6 | | 6.0 ± 4.2 |
| pLE12I | | | 3.7 ± 0.8 | 4.8 ± 1.1 | |
| pLE12IK[b] | | | 3.7 ± 0.1 | | 5.5 ± 0.8 |
| pLE12RT$^-$ | | | 1.4 ± 0.5 | 2.5 ± 0.5 | |
| pLE12Mat$^-$ | | 7.7 ± 2.9* | 2.7 ± 0.4 | 5.5 ± 0.7 | |
| pLE12Endo$^-$ | | | 2.4 ± 0.9 | 1.9 ± 0.1 | |
| pLE12ΔORF | | | 4.3 ± 0.4 | | |
| pLE12ΔD5 | | | 3.9 ± 1.0 | | |

Efficiencies are an average of 3 independent conjugation assays
All the receiving cells contained the pMNHS ($Spc^R$) recipient plasmid
*The conjugation efficiency is $10^{-8}$
[a]DNase I/RNase conjugation control assays
[b]For these assays, the pMNHS recipient plasmid is $Spc^S$-$Erm^R$
[c]NZΔ*ltrB*=NZ9800Δ*ltrB*: :$Tet^R$
NZΔ*ltrB*, NZ9800, LM0230(*recA*$^+$), and MMS372(*recA*$^-$) are *L. lactis* strains, and JH2-2 is an *E. faecalis* strain
Dashed line (---) = non detected; no transconjugants (3 assays) (<$10^{-8}$)

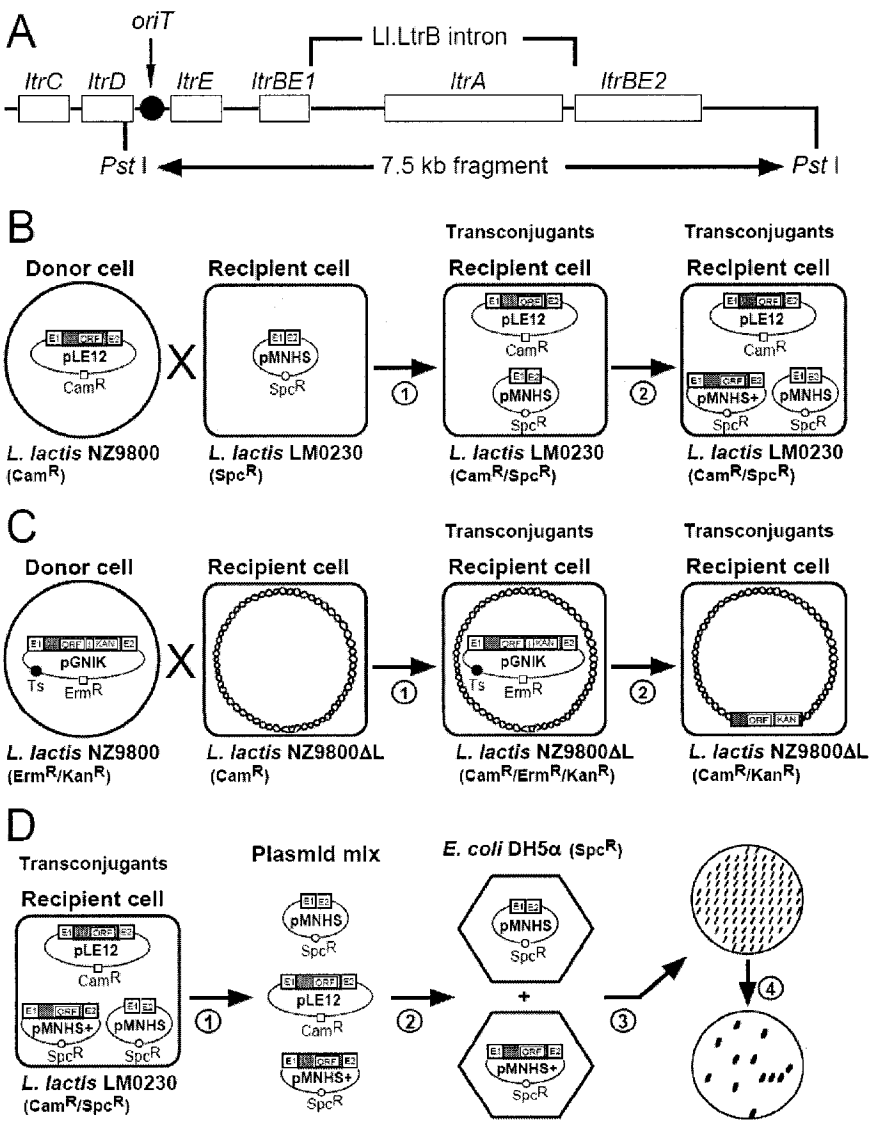
**Table 2.3.** Efficiency of Ll.LtrB mobility from pLE12 derivatives following their transfer by conjugation

| | NZ9800/plasmid × LM0230(*recA*$^+$)/pMNHS | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| No. | WT | I | | | IK | | | Mat$^-$ | RT$^-$ | Endo$^-$ | ΔORF | ΔD5 |
| | | gpI SJ | gpII | gpI loss | gpI SJ | gpII | gpI loss | | | | | |
| 1 | 10 | 1 | 1 | 100 | 5 | 5 | 100 | 0 | 0 | 2 | 0 | 0 |
| 2 | 10 | 4 | 4 | 100 | 1 | 1 | 100 | 2 | 0 | 0 | 0 | 0 |
| 3 | 10 | 7 | 7 | 100 | 2 | 2 | 100 | 0 | 0 | 0 | 0 | 0 |
| 4 | 11 | 2 | 2 | 100 | 2 | 2 | 100 | 0 | 0 | 1 | 0 | 0 |
| 5 | 8 | 2 | 2 | 100 | 4 | 4 | 100 | 0 | 0 | 2 | 0 | 4 |
| 6 | 8 | 2 | 2 | 100 | 4 | 4 | 100 | 0 | 0 | 0 | 0 | 0 |
| 7 | 10 | 4 | 4 | 100 | 5 | 5 | 100 | 55 | 0 | 96 | 0 | 0 |
| 8 | 13 | 3 | 3 | 100 | 1 | 1 | 100 | 0 | 0 | 10 | 0 | 0 |
| 9 | 16 | 3 | 3 | 100 | 4 | 4 | 100 | 1 | 0 | 0 | 0 | 0 |
| 10 | 9 | 4 | 4 | 100 | 1 | 1 | 100 | 0 | 1 | 4 | 47 | 0 |
| Mean | 10.5% | | 3.2% | | | 2.9% | | <1% | <1% | 1.1% | <1% | <1% |
| SEM | ±0.8 | | ±0.5 | | | ±0.5 | | | | ±0.5 | | |

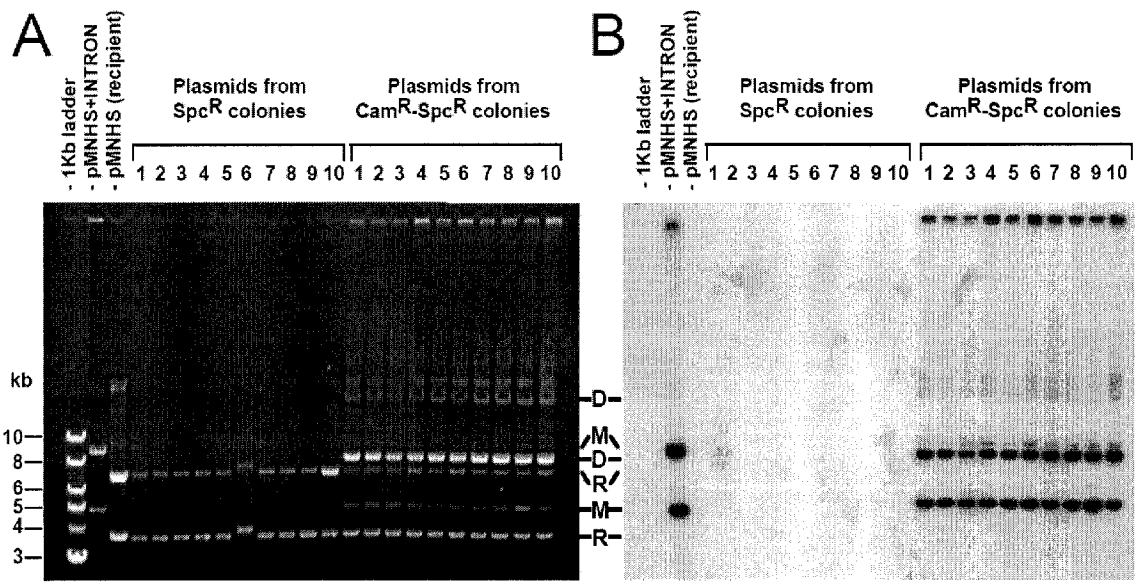
| | NZ9800/plasmid × MMS372(*recA*$^+$)/pMNHS | | | | | | |
|---|---|---|---|---|---|---|---|
| No. | WT | I | | | Mat$^-$ | RT$^-$ | Endo$^-$ |
| | | gpI SJ | gpII | gpI loss | | | |
| 1 | 100 | 14 | 60 | – | 0 | 44 | 1 |
| 2 | 14 | 19 | 19 | 100 | 86 | 0 | 4 |
| 3 | 50 | 20 | 20 | 100 | 100 | 0 | 1 |
| 4 | 35 | 21 | 22 | 95 | 0 | 91 | 4 |
| 5 | 26 | 16 | 17 | 94 | 0 | 0 | 1 |
| 6 | 19 | 15 | 15 | 100 | 0 | 0 | 3 |
| 7 | 94 | 13 | 13 | 100 | 0 | 0 | 0 |
| 8 | 42 | 24 | 24 | 100 | 0 | 0 | 28 |
| 9 | 22 | 19 | 85 | – | 0 | 0 | 1 |
| 10 | 13 | 15 | 15 | 100 | 100 | 24 | 4 |
| Mean | 27.6% | | 18.1% | | <1% | <1% | 2.1% |
| SEM | ±4.8 | | ±1.3 | | | | ±0.5 |

The mobility efficiencies of the Ll.LtrB intron variants from the donor plasmids (Fig. 3) were calculated by patch hybridization assays (*Experimental procedures*) analysing 100 colonies for each of the 10 independent transconjugants (nos 1–10).
For conjugation assays with pLE12IK, the pMNHS used was Spc$^R$-Erm$^R$.
Mobility efficiencies from supermobiles are underlined and not included in the means.
gpI SJ, number of colonies revealed with the *td* group I intron splice junction probe (ligated exons); gpII, number of colonies revealed with the group II intron probe; gpI loss, percentage of mobility events lacking the group I intron.

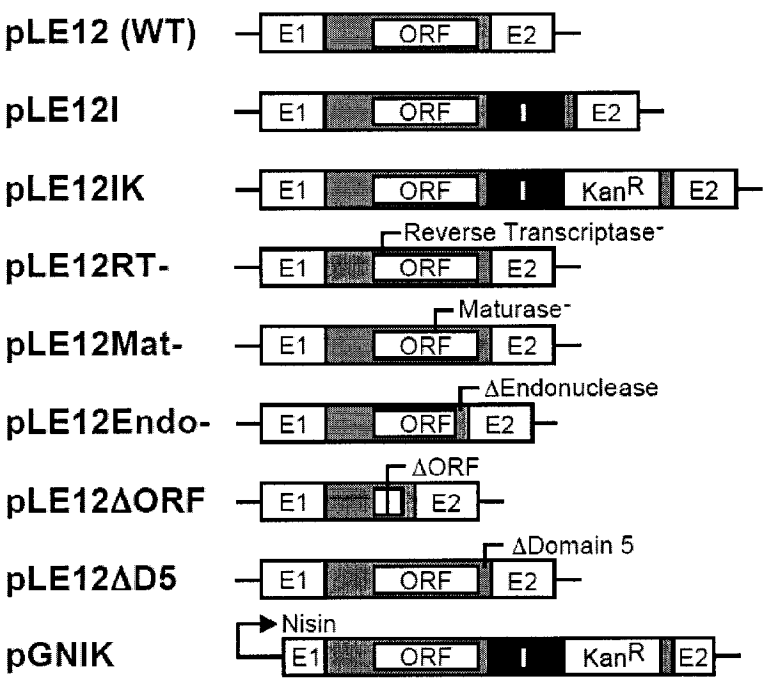
# Figure 2.1


A
oriT
LI.LtrB intron
ltrC
ltrD
ltrE
ltrBE1
ltrA
ltrBE2
Pst I
7.5 kb fragment
Pst I
B
Donor cell
Recipient cell
Transconjugants
Recipient cell
Transconjugants
Recipient cell
pLE12
CamR
pMNHS
SpcR
pMNHS+
L. lactis NZ9800
(CamR)
L. lactis LM0230
(SpcR)
L. lactis LM0230
(CamR/SpcR)
L. lactis LM0230
(CamR/SpcR)
C
Donor cell
Recipient cell
Transconjugants
Recipient cell
Transconjugants
Recipient cell
pGNIK
Ts
ErmR
KAN
L. lactis NZ9800
(ErmR/KanR)
L. lactis NZ9800ΔL
(CamR)
L. lactis NZ9800ΔL
(CamR/ErmR/KanR)
L. lactis NZ9800ΔL
(CamR/KanR)
D
Transconjugants
Recipient cell
Plasmid mix
E. coli DH5α (SpcR)
L. lactis LM0230
(CamR/SpcR)

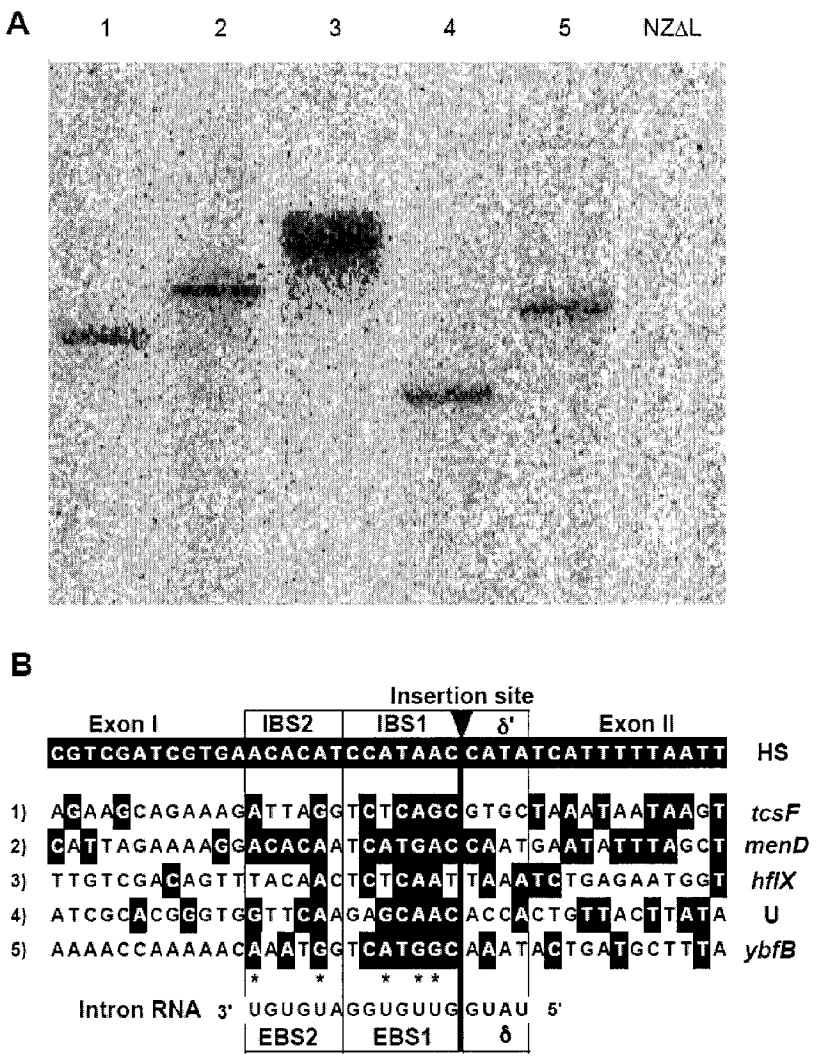
# Figure 2.2


A
- 1Kb ladder
- pMNHS+INTRON
- pMNHS (recipient)
Plasmids from Spc^R colonies
Plasmids from Cam^R-Spc^R colonies
1 2 3 4 5 6 7 8 9 10 1 2 3 4 5 6 7 8 9 10
kb
10
8
6
5
4
3
D
M
D
R
M
R
B
- 1Kb ladder
- pMNHS+INTRON
- pMNHS (recipient)
Plasmids from Spc^R colonies
Plasmids from Cam^R-Spc^R colonies
1 2 3 4 5 6 7 8 9 10 1 2 3 4 5 6 7 8 9 10

## Figure 2.3


pLE12 (WT)
E1
ORF
E2
pLE12I
I
pLE12IK
KanR
pLE12RT-
Reverse Transcriptase-
pLE12Mat-
Maturase-
pLE12Endo-
ΔEndonuclease
pLE12ΔORF
ΔORF
pLE12ΔD5
ΔDomain 5
pGNIK
Nisin

# Figure 2.4


A
1
2
3
4
5
NZΔL
B
Insertion site
Exon I
IBS2
IBS1
δ'
Exon II
CGTCGATCGTGAACACATCCATAACCATATCATTTTTAATT
HS
1) AGAAGCAGAAAGATTAGGTCTCAGCGTGCTAAATAATAAGT tcsF
2) CATTAGAAAAGGACACAATCATGACCAATGAATATTTAGCT menD
3) TTGTCGACAGTTTACAACTCTCAATTTAAATCTGAGAATGGT hflX
4) ATCGCACGGGTGGTTCAAGAGCAACACCACTGTTACTTATA U
5) AAAACCAAAAACAAATGGTCATGGCAAATACTGATGCTTTA ybfB
Intron RNA 3' UGUGUAGGUGUUGGUAU 5'
EBS2
EBS1
δ

# CHAPTER THREE

## Conjugative transfer of the *Lactococcus lactis* chromosomal sex factor promotes dissemination of the Ll.LtrB group II intron

### 3.1 Preface

In the previous chapter, we demonstrated that Ll.LtrB can be disseminated when present on a mobilizable plasmid, provided that the host cell harbours minimal transfer functions. Following conjugative transfer, Ll.LtrB can invade DNA sites in the new host either by retrohoming or by retrotransposition. Unusually high retrohoming efficiencies were observed for wild-type introns as well as for introns mutated for critical splicing and/or mobility functions (Table 2.3, underlined values). The absence of the *td* splice-junction from these events and their independence from critical splicing functions first suggested a DNA-based mobility pathway. The following chapter describes the study we undertook to explain these unusual mobility efficiencies and how they originated.

JOURNAL OF BACTERIOLOGY, Feb. 2005, p. 930–939 Vol. 187, No. 3
0021-9193/05/$08.00+0 doi:10.1128/JB.187.3.930–939.2005


# Conjugative Transfer of the *Lactococcus lactis* Chromosomal Sex Factor Promotes Dissemination of the Ll.LtrB Group II Intron

Kamila Belhocine, Karen K. Yam, and Benoit Cousineau*
*Department of Microbiology and Immunology, McGill University, Montréal, Québec, Canada*

In: ***Journal of Bacteriology*** (2005) **187**:930-939


## 3.2 Abstract

The Ll.LtrB group II intron from the low-G+C Gram-positive bacterium *Lactococcus lactis* was the first bacterial group II intron shown to splice and mobilize *in vivo*. This retroelement interrupts the relaxase gene (*ltrB*) of three *L. lactis* conjugative elements: plasmids pRS01 and pAH90 and the chromosomal sex factor. Conjugative transfer of a plasmid harbouring a segment of the pRS01 conjugative plasmid including the Ll.LtrB intron allows dissemination of Ll.LtrB among *L. lactis* strains and lateral transfer of this retroelement from *L. lactis* to *Enterococcus faecalis*. Here we report the dissemination of the Ll.LtrB group II intron among *L. lactis* strains following conjugative transfer of the native chromosomally embedded *L. lactis* sex factor. We demonstrated that Ll.LtrB dissemination is highly variable and often more efficient from this integrative and conjugative element than from an engineered conjugative plasmid. Cotransfer among *L. lactis* strains of both Ll.LtrB-containing elements, the conjugative plasmid and the sex factor, was detected and shown to be synergistic. Moreover,

following their concurrent transfer, both mobilizable elements supported the spread of their respective copies of the Ll.LtrB intron. Our findings explain the unusually high efficiency of Ll.LtrB mobility observed following conjugation of intron-containing plasmids.

## 3.3 Introduction

Group II introns are large ribozymes that splice autocatalytically from their pre-mRNAs (Lambowitz and Belfort, 1993; Saldanha et al., 1993; Michel and Ferat, 1995; Belfort et al., 2002). Some self-splicing group II introns are also mobile retroelements that invade new DNA sites in a duplicative process using an RNA intermediate, like retrotransposons and retroviruses (Belfort et al., 2002). They can reinsert either in cognate intronless alleles (homing site [HS]) by retrohoming or in non-homologous sites by retrotransposition (Cousineau et al., 1998, 2000; Belfort et al., 2002; Ichiyanagi et al., 2002, 2003). Mobile group II introns harbour a multifunctional open reading frame (ORF) that is directly involved in their mobility processes (Belfort et al., 2002). Group II introns are found in eubacteria (Martinez-Abarca and Toro, 2000; Dai and Zimmerly, 2002), archaea (Dai and Zimmerly, 2003; Toro, 2003), and eukaryotic organelles derived from bacteria such as fungal and plant mitochondria and plant chloroplasts (Lambowitz and Belfort, 1993; Saldanha et al., 1993; Michel and Ferat, 1995; Belfort et al., 2002). Horizontal transfer of group II introns between organisms is a well-accepted model of intron dissemination and evolution. This model suggests that group II introns are not only mobile within cells but can also be transferred

between species, where they can invade new sites (Lambowitz and Belfort, 1993; Belhocine et al., 2004).

The *Lactococcus lactis* LtrB group II intron (Ll.LtrB) is the first bacterial group II intron that was shown to splice and mobilize *in vivo* (Mills et al., 1996, 1997; Shearman et al., 1996). *L. lactis*, an industrially important low-G+C Gram-positive bacterium, is extensively used in the dairy industry. Ll.LtrB mobility via the retrohoming (Cousineau et al., 1998) and retrotransposition pathways (Cousineau et al., 2000; Ichiyanagi et al., 2002, 2003) was studied in both *L. lactis* and *Escherichia coli*. The Ll.LtrB group II intron (2.5 kb) harbours an ORF called *ltrA* (599 amino acids) that exhibits reverse transcriptase, endonuclease, and RNA maturase activities (Matsuura et al., 1997). These three functions are essential for retrohoming of Ll.LtrB to intronless alleles (Cousineau et al., 1998). Following translation, LtrA binds to its harbouring intron within the pre-mRNA as a dimer (Rambo and Doudna, 2004). The maturase function of LtrA promotes splicing of Ll.LtrB and concurrent ligation of its flanking exons. Following intron excision, the LtrA dimer remains bound to the intron RNA lariat as a ribonucleoprotein particle (RNP) (intron RNA lariat plus two LtrA proteins). Upon recognition of the homing site by these RNPs, the intron RNA reverse splices into the sense strand of its double-stranded DNA target. The antisense strand is then nicked nine nucleotides downstream from the intron insertion site by the endonuclease domain of LtrA. Using the 3' end generated by endonuclease cleavage as a primer, LtrA reverse transcribes the intron RNA by a process called target-primed reverse transcription (TPRT). The final steps of the Ll.LtrB retrohoming pathway are thought to be carried out by host DNA repair mechanisms independent of the

RecA-dependent homologous recombination pathway (Mills et al., 1997; Cousineau et al., 1998).

The Ll.LtrB group II intron interrupts the relaxase gene (*ltrB*) that is present in three *L. lactis* mobilizable elements: two conjugative plasmids, pRS01 (48.4 kb) (Mills et al., 1996) and pAH90 (26.5 kb) (O'Sullivan et al., 2001), and the chromosomally embedded sex factor (50 kb) (Shearman et al., 1996). The relaxase enzyme functions by nicking the plasmid at its origin of transfer (*oriT*) to initiate conjugation; hence, splicing of the Ll.LtrB intron is essential for relaxase production and plasmid transfer (Mills et al., 1996; Zhou et al., 2000; Belhocine et al., 2004; Klein et al., 2004). The pRS01 plasmid and the chromosomal sex factor are very similar and were probably derived from a common ancestor (Burrus et al., 2004). They are considered integrative and conjugative elements (ICEs) because they excise by site-specific recombination into a circular form, self transfer by conjugation, and integrate into the host genome (Burrus et al. 2004).

We previously demonstrated that transfer of an intron-harbouring conjugative plasmid among different *L. lactis* strains and from *L. lactis* to *Enterococcus faecalis* supports intron dissemination and lateral transfer within the recipient cells (Belhocine et al., 2004). This plasmid contained the conjugative transfer regions Tra1-2 from pRS01 (7.5 kb) harbouring the *oriT*, *ltrE*, and *ltrB* genes; the latter was interrupted by the Ll.LtrB intron. Following its transfer via plasmid-based conjugation, the Ll.LtrB intron was shown to invade either its recognition site (homing site [HS]) harboured on a resident plasmid by retrohoming or different non-homologous sites present within the chromosome of the recipient cells by retrotransposition (Belhocine et al., 2004). While studying

intron dissemination by plasmid conjugation, we consistently observed an unusually high efficiency of Ll.LtrB mobility in some analyzed isolates. Moreover, we observed very efficient Ll.LtrB mobility for some intron variants carrying mutations that inactivated any of the three catalytic activities of LtrA (reverse transcriptase, maturase, and endonuclease), although these LtrA functions were demonstrated to be essential for Ll.LtrB retrohoming (Cousineau et al., 1998).

Why were there unusually high levels of Ll.LtrB mobility products in some transconjugant isolates following conjugation of intron-carrying plasmids? Using genetic as well as conjugation/retrohoming assays, we describe the dissemination of the Ll.LtrB group II intron following transfer of the chromosomally embedded *L. lactis* sex factor between *L. lactis* strains. We show that (i) the chromosomal sex factor can be cotransferred along with a conjugative plasmid, leading to the dissemination of the Ll.LtrB intron it conveys; (ii) cotransfer of a conjugative plasmid and the sex factor is synergistic; (iii) Ll.LtrB dissemination is highly variable and often more efficient from the sex factor than from a vector carrying a portion of the pRS01 conjugative plasmid. This work reveals the nature and origin of the unusually high Ll.LtrB mobility efficiencies observed following conjugation of intron-containing plasmids (Belhocine et al., 2004).

## 3.4 Materials and methods

### 3.4.1 Strains and plasmids

*L. lactis* strains [NZ9800, NZ9800Δ*ltrB*::Tet$^r$, and MMS372(*recA*)] were grown without shaking in M17 medium supplemented with 0.5% glucose (GM17) at 30°C. The chromosomes of the NZ9800 and NZ9800Δ*ltrB*::Tet$^r$ strains contain a copy of the conjugative sex factor, while the MMS372 strain is sex factor free. *E. coli* strains (DH5α, DH10β), used for cloning and mobility scoring, were grown shaken at 37°C in Luria-Bertani (LB) broth. Milk plates used in conjugation assays were made of 5% dry milk (Carnation), 1% dextrose, and 1.5% agar. pLE12I plasmid consists of the Tral-2 regions from pRS01 cloned into the pLE1 vector at its unique PstI site, where the *td* group I intron with portions of its exons was subsequently inserted into the Ll.LtrB intron downstream from *ltrA* (Cousineau et al., 1998; Belhocine et al., 2004). pMNHS plasmid (pMN1343) contains a 271-bp homing site (exon 1, 179 bp; exon 2, 92 bp) (HindIII) inserted at the unique HindIII restriction site in the pDL278 vector (Mills et al., 1997; Cousineau et al., 1998; Belhocine et al., 2004). pMNHS-CR plasmid contains, in both exons, polymorphic sites that do not interfere with Ll.LtrB mobility (E1, –7 and –30/35; E2, +7 and +25). The modified homing site was isolated from pLHS-CR plasmid (XbaI) (Cousineau et al., 1998) and cloned blunt into the pDL278 vector (HindIII). Selective medium contained the following concentrations of antibiotics: chloramphenicol (Cam), 10 μg/ml; spectinomycin (Spc), 300 μg/ml; kanamycin (Kan), 20 μg/ml; tetracycline (Tet), 3 μg/ml.

### 3.4.2 Conjugation assays

*L. lactis* strains (donor and recipient) were diluted (0.4 or 0.8 ml/10 ml) from cultures saturated overnight and were grown for 7 h at 30°C with appropriate antibiotics. Cells were recovered by centrifugation and the pellets were mixed (1:1), spread on milk plates containing DNase I and RNase I (100 U of each/ml) (Trieu-Cuot *et al.*, 1998; Belhocine *et al.*, 2004), and incubated at 30°C for 12 h. Typical conjugation control experiments used only DNase I, but because the Ll.LtrB intron is a retroelement, we also controlled for possible non-conjugative uptake of active RNP particles (intron RNA lariat plus two LtrA proteins). Cell mixtures were recovered with 1x PBS (1 ml), and serial dilutions were made to score donor ($Cam^r$), recipient ($Spc^r$), and transconjugant ($Cam^r/Spc^r$) cells (Mills et al., 1996; Belhocine et al., 2004). Conjugation efficiencies (three assays) were calculated as the number of transconjugants ($Cam^r/Spc^r$)/donor cell ($Cam^r$). In sex factor conjugation assays, the identity of the recipient strain [MMS372(*recA*)] was confirmed by its resistance to Kan and its UV sensitivity.

### 3.4.3 Mobility assay (colony patch hybridization)

The plasmid mix (donor plasmids, recipient plasmids, and mobility products) from 10 independent transconjugant cells was prepared (same conjugation assay) and retransformed into *E. coli* (DH5 α), and bacteria were plated on LB/Spc plates to select for cells containing recipient plasmids that were or were not interrupted by the intron (Figure 3.1A). To calculate the intron mobility efficiency measured by the percentage of recipient plasmids that received the intron (mobility products), 100 isolated colonies ($Spc^r$) were patched

for each independent assay (LB/Spc plates). The patches were lifted on nylon membranes and hybridized with the appropriate 5'-end-labeled probe ($^{32}P$) (Figure 3.1B). The group II probe (gpII; 5'-CCGTGCTCTGTTCCCGTATCAGC-3') (5' end of the intron) (Cousineau et al., 1998) is general and recognizes the three types of mobility events (Figure 3.1B). Three other probes (gpI SJ, gpI, and WT gpII) are specific, and each recognizes only one type of mobility event. The *td* group I intron splice junction probe (gpI SJ; 5'-ATTAAAC-GGTAGACCCAAGAAAAC-3') (Cousineau et al., 1998) recognizes the two *td* ligated exons (12 nucleotides each) flanking it and gives a positive signal only when the group I intron is absent in Ll.LtrB mobility products (Belfort et al., 1990; Cousineau et al., 1998, 2000; Belhocine et al., 2004). This probe is used as a retromobility indicator. The group I probe (gpI; 5'-GGAGATATAG-TCTGCTCTGCA-3') hybridizes within the *td* group I intron and reveals mobility events still harbouring it. The wild-type group II probe (WT gpII; 5'-AAACA-CAAGTGAATTTTTACGA-3') spans the region where the *td* group I intron is inserted just downstream from the LtrA stop codon (engineered SalI site). We designed this probe to specifically recognize the wild-type Ll.LtrB intron and to not hybridize with either of the two twintron mobility products.

## 3.5 Results

### 3.5.1 Unusually high efficiency of Ll.LtrB mobility among *L. lactis* strains is consistently observed following conjugation of intron-harbouring plasmids

To study the variable efficiencies of Ll.LtrB mobility observed following its transfer by plasmid conjugation, we performed a typical conjugation assay between NZ9800/pLE12I (donor strain) and MMS372(*recA*)/pMNHS-CR (recipient strain) (Figure 3.2A) (Belhocine et al., 2004). In this assay, the non-autonomously conjugative donor plasmid (pLE12I; $Cam^r$) carried an engineered Ll.LtrB intron harbouring the *td* group I intron from phage T4 flanked by portions of its exons, while the recipient plasmid (pMNHS-CR; $Spc^r$) contained the intron recognition site (HS) (Figure 3.2A) (Cousineau et al., 1998; Belhocine et al., 2004). Conjugation machinery is provided by the sex factor embedded in the chromosome of NZ9800 (Belhocine et al., 2004). After mating, the recipient cells that received the conjugative plasmid (transconjugant; MMS372(*recA*)/pMNHS-CR/pLE12I) were selected on $Cam^r/Spc^r$ plates, while the donor and recipient strains were isolated on $Cam^r$ and $Spc^r$ plates, respectively. The presence of the *td* self-splicing intron within Ll.LtrB (twintron) allowed us to determine whether the intron invaded its homing site by using an RNA intermediate. The occurrence of mobility events missing the group I intron confirms that these products are generated using an Ll.LtrB RNA intermediate (Cousineau et al., 1998, 2000; Ichiyanagi et al., 2002, 2003; Belhocine et al., 2004).

Conjugation efficiency of the pLE12I plasmid was $5.2 \times 10^{-3} \pm 0.4 \times 10^{-3}$ transconjugant per donor cell (three independent assays), as previously observed ($4.8 \times 10^{-3} \pm 1.1 \times 10^{-3}$) (Belhocine et al., 2004). To determine the mobility efficiency of the Ll.LtrB intron following plasmid conjugation, we isolated the plasmids (donor plasmids, recipient plasmids, and mobility products) from 10 independent transconjugant colonies (Figure 3.2A). Analyses by agarose gel

electrophoresis (Figure 3.2B) showed, as previously observed (Belhocine et al., 2004), some variability in Ll.LtrB mobility efficiency among the different isolates. Indeed, intensity of bands corresponding to uninterrupted recipient plasmid (R; pMNHS-CR) and mobility product (M; pMNHS-CR plus intron) varied from lane to lane (Figure 3.2B).

Efficiency of Ll.LtrB mobility was determined by colony patch hybridization assays for each of the 10 plasmid mixes isolated (Figure 3.1A) (Belhocine et al., 2004). In these assays, *E. coli* DH5α cells were transformed with the plasmid mixes recovered from *L. lactis* transconjugant colonies and were plated on LB/Spc plates to select for cells containing recipient plasmids harbouring or not harbouring the intron. For each transconjugant event, 100 colonies were patched onto LB/Spc plates, transferred to nylon membranes, and hybridized with intron-specific $^{32}$P-labeled probe (gpII; Figure 3.1B). Efficiency of Ll.LtrB mobility, defined as the ratio of Ll.LtrB-interrupted recipient plasmids to total recipient plasmids, varied from 13 to 70% for the 10 studied transconjugant events (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, gpII).

Because pLE12I carried the twintron variant of Ll.LtrB, we also assessed which mobility products had lost the group I intron by probing them with the group I splice junction probe (Figure 3.1B, gpI SJ). This probe recognizes the group I intron ligated exons and produces a positive signal only if the *td* group I intron is missing from group II intron mobility products. As previously observed (Belhocine et al., 2004), the majority of the transconjugant cells (7 out of 10) harboured mobility products showing nearly 100% loss of the retromobility indicator (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, gpI SJ, gpI SJ/gpII, nos. 3

and 5 to 10). These results demonstrate that the majority of mobility products resulted from Ll.LtrB invading its homing site through the retrohoming pathway (Cousineau et al., 1998). In sharp contrast, in transconjugant cells that displayed unusually high levels of Ll.LtrB mobility, only a fraction of mobility products appear to have lost the group I intron (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, gpI SJ/gpII, nos. 1, 2, and 4). Levels of bona fide retromobility products (gpII and gpI SJ positive) observed in these three transconjugant cells are in the same range as those observed for the other seven transconjugant cells (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, compare gpI SJ and gpII mobility averages). Similar unusually high frequencies of Ll.LtrB mobility were previously observed in various crosses among different *L. lactis* strains where transferred plasmids were carrying different intron derivatives (Belhocine et al., 2004). Our findings confirm that unusually high efficiencies of Ll.LtrB mobility following plasmid conjugation are a general and reproducible phenomenon.

### 3.5.2 Some Ll.LtrB mobility products in transconjugant cells are not generated from pLE12I plasmid

Does a significant portion of mobility products present in the three isolates showing unusually high mobility efficiency of Ll.LtrB still harbour the *td* group I intron (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, gpI SJ/gpII, no. 1, 2, and 4)? The presence of Ll.LtrB mobility products still harbouring the *td* intron would explain why the group I splice junction probe did not hybridize to this subset of mobility products. However, using a group I intron-specific probe (Figure 3.1B, gpI), we found that almost none of these mobility products contained the *td* group

I intron (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, gpI). This suggests that these inserted Ll.LtrB introns are not derived from donor pLE12I plasmid, because mobility products coming from this plasmid should harbour either the *td* splice junction or the unspliced *td* intron. Only one mobility product still contained the *td* group I intron (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, gpI, no. 10). This mobility product is most likely a retrohoming event in which the *td* group I intron was unable to splice before the Ll.LtrB intron RNA was reverse transcribed (Cousineau et al., 1998).

To demonstrate that these mobility products were not derived from the pLE12I plasmid and that they were probably generated from a wild-type copy of the intron, we designed a probe specific for wild-type Ll.LtrB (Figure 3.1B, WT gpII) that would not hybridize to mobility products generated from donor plasmid, harbouring or lacking the *td* group I intron (Figure 3.1B). This DNA oligonucleotide probe spans the region where the *td* group I intron is inserted in Ll.LtrB just downstream of LtrA (engineered SalI site) (Cousineau et al., 1998). We confirmed that mobility products that are negative for both group I and group I splice junction probes contained wild-type Ll.LtrB introns not expressed from pLE12I plasmid (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, WT gpII). The WT gpII probe gave positive signals only for the three transconjugant isolates where unusual Ll.LtrB mobility efficiency was observed and where group I intron loss from mobility products seemed inefficient (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, WT gpII, gpII, and gpI SJ/gpII, no. 1, 2, and 4). Moreover, mobility products harbouring the wild-type Ll.LtrB intron corresponded in all three cases to gpII-positive, gpI SJ-negative mobility products (Table 3.1,

NZ/pLE12I x MMS/pMNHS-CR; WT gpII plus gpI SJ = gpII). This implies that mobility products derived from pLE12I plasmid efficiently lost the *td* intron and therefore were generated by retrohoming.

To determine whether the wild-type Ll.LtrB mobility products were generated through a DNA-based mobility pathway or through the retrohoming pathway, we performed a coconversion analysis of flanking markers around the intron homing site (Cousineau et al., 1998). DNA-dependent Ll.LtrB mobility predicts that genetic information located within flanking exons could be exchanged between donor and recipient alleles, while precise insertion of the intron through complete reverse splicing (retrohoming) does not promote genetic exchange between homologous exons (Cousineau et al., 1998). The recipient plasmid used in the conjugation assay (Figure 3.2A, pMNHS-CR) contained polymorphic sites in both exons that do not interfere with Ll.LtrB mobility (E1, –7 and –30/35; E2, +7 and +25) (Cousineau et al., 1998). Sequencing analyses of intron-exon junctions (5' and 3') of 10 wild-type Ll.LtrB mobility products (gpII+/gpI SJ–/gpI–/WT gpII+) (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR, no. 1) were performed. All 10 mobility products had retained their polymorphic nucleotides on both sides of the intron insertion site, showing that no marker coconversion occurred during Ll.LtrB insertion. This suggests that wild-type Ll.LtrB intron invades its recognition site on the recipient plasmid by the RNA-based retrohoming pathway.

Using the same four probes (gpII, gpI SJ, gpI, and WT gpII) (Figure 3.1B), we analyzed two previously isolated transconjugant events (no. 1 and 9) along with a control (no. 2) from the same cross that showed an unusual pattern of

Ll.LtrB mobility (Belhocine et al., 2004). Similarly, in isolates 1 and 9, all Ll.LtrB mobility products that are gpII positive and gpI and gpI SJ negative were recognized by the WT gpII probe, showing that they were also not derived from pLE12I donor plasmid (Table 3.1, NZ/pLE12I x MMS/pMNHS, isolates from a previous study (Belhocine *et al.*, 2004)). Taken together, these results suggest that in cases of unusually high efficiency of Ll.LtrB mobility following plasmid conjugation, two different Ll.LtrB dissemination pathways are operating concurrently.

### 3.5.3 Wild-type Ll.LtrB mobility products originated from the chromosome of the donor strain

Because the recipient strain does not contain a chromosomal copy of the Ll.LtrB intron, we hypothesized that the detected wild-type intron originated from the chromosome of the donor strain. To test our hypothesis, we performed a conjugation assay similar to the one described above but where the wild-type chromosomal copy of the Ll.LtrB intron was deleted from the donor strain (NZ9800Δ*ltrB*/pLE12I x MMS372/pMNHS-CR) (Figure 3.2A). Chromosomal deletion of Ll.LtrB removed portions of the flanking exons, most probably destroying the *ltrB* gene and preventing production of the relaxase enzyme (LtrB) that is essential to initiate plasmid conjugation (Mills et al., 1996; Zhou et al., 2000; Belhocine et al., 2004; Klein et al., 2004). However, in our conjugation system, this defect is complemented by expression of the relaxase enzyme from the pLE12I plasmid (Belhocine et al., 2004). As previously observed (Belhocine et al., 2004), because the relaxase is produced only from the conjugative plasmid,

conjugation efficiency was lower than that when the relaxase is also expressed from the chromosome (~100-fold; 7.5 x 10–5 ± 2.7 x $10^{-5}$ versus 5.2 x $10^{-3}$ ± 0.4 x $10^{-3}$). In this particular case, the presence of the group I intron within Ll.LtrB hinders its splicing efficiency and lowers relaxase expression. This explains the more dramatic reduction of pLE12I conjugation efficiency (~100-fold) in the NZ9800Δ*ltrB* background compared to that previously observed for pLE12, where the intron is wild type (~10-fold; 2.2 x $10^{-4}$ ± 0.6 x $10^{-4}$ versus 2.3 x $10^{-3}$ ± 0.6 x $10^{-3}$) (Belhocine et al., 2004).

Following conjugation of pLE12I, the plasmid mix from progeny of 10 independent transconjugant colonies were studied as described above by using the same four probes (gpII, gpI SJ, gpI, and WT gpII) (Figure 3.1B). The gpII probe is general and reveals all types of mobility products, while the other probes are specific for three different subsets of Ll.LtrB mobility products (Figure 3.1B). In contrast to what was observed in the first cross (Figure 3.2B), all plasmid mixes contained the same ratio of mobility products (M) over uninterrupted plasmids (R), suggesting similar efficiencies of Ll.LtrB mobility (Figure 3.2C, compare M and R). This observation was confirmed by colony patch hybridization assays; no mobility products harbouring a wild-type copy of Ll.LtrB were found (Table 3.1, NZΔ*ltrB*/pLE12I x MMS/pMNHS-CR, WT gpII) and no unusually high Ll.LtrB mobility was seen (Table 3.1, NZΔ*ltrB*/pLE12I x MMS/pMNHS-CR). Finally, all mobility products hybridized either with the gpI or gpI SJ probe, confirming that all originated from pLE12I donor plasmid. Mobility products harbouring a wild-type copy of Ll.LtrB observed following plasmid conjugation most probably acquired their introns from the chromosome of the donor strain.

### 3.5.4 Conjugative transfer of the chromosomally located sex factor induces dissemination of the Ll.LtrB group II intron

Can the Ll.LtrB-carrying sex factor embedded within the chromosome of the NZ9800 strain be efficiently transferred, and can conjugation disseminate its copy of Ll.LtrB to other *L. lactis* strains? We performed a conjugation assay in which the donor strain was plasmid free (Figure 3.3A, NZ9800 x MMS372/pMNHS). Because the sex factor and the Ll.LtrB intron are wild type and do not carry a selective marker, we directly probed, via colony patch hybridization, 1,000 colonies of recipient cells (MMS372/pMNHS) for the presence of the group II intron (Figure 3.1B, gpII probe). We found 9 positive patches potentially carrying Ll.LtrB mobility products (9 out of 1,000). Recipient plasmids showed extra bands of variable intensity from lane to lane, corresponding to hypothetical Ll.LtrB mobility products (Figure 3.3C, M). Colony patch hybridization assays (Figure 3.1A) confirmed that all nine isolates contained Ll.LtrB mobility products and that the percentage of interrupted plasmid was highly variable (6 to 96%) (Figure 3.3C, % Mob). Because the recipient strain [MMS372(*recA*); $Kan^r$] is kanamycin resistant and UV sensitive (*recA*), we could rule out the possibility that the recipient plasmid had been transferred from the recipient to donor cell [NZ9800(*recA*+); $Kan^s$]. Indeed, we observed that, like the MMS372 recipient control strain, all nine isolates were kanamycin resistant and UV sensitive, while the NZ9800 donor control strain was kanamycin sensitive and UV resistant (data not shown).

To determine if the sex factor was present in the nine MMS372 recipient cells carrying Ll.LtrB mobility products, we performed Southern blot analyses on genomic DNA using two sex factor-specific probes that recognize both extremities of the element (Godon et al., 1995). We confirmed that the sex factor was present in eight of the nine isolated recipient strains (Figure 3.3C, Sex F) but was absent from the genomic DNA of the control MMS372 strain (data not shown).

To confirm the frequency at which the sex factor is transferred, we twice repeated the cross described above. Recipient cells were first analyzed with a sex factor-specific probe (Godon et al., 1995), where we twice found that 7 recipient cells had received the sex factor (7 out of 1,000). The patches were then re-hybridized with gpII probe; no additional signals appeared, showing that no recipient cells carried Ll.LtrB without harbouring the sex factor. If we exclude isolate 6 from the first cross, because it is not clear whether it harboured the sex factor, the sex factor transfer frequency is high at $7.3 \times 10^{-3} \pm 0.3 \times 10^{-3}$ (three assays) and correlates well with that previously observed (Shearman et al., 1996) and with the transfer frequency of the engineered plasmid pLE12 carrying wild-type Ll.LtrB intron ($2.3 \times 10^{-3} \pm 0.5 \times 10^{-3}$) (Belhocine et al., 2004).

Taken together, these results demonstrate that the *L. lactis* sex factor can be efficiently transferred by conjugation from the chromosome of strain NZ9800 to strain MMS372 and can support Ll.LtrB dissemination.

### 3.5.5 Co-transfer of pLE12I plasmid and sex factor

Knowing that our donor strain contains two proficient mobilizable elements, we hypothesized that the conjugative plasmid and the sex factor can be simultaneously transferred from donor cell to recipient cell during conjugation and that both support the dissemination of their Ll.LtrB copy. To test our hypotheses, we repeated the NZ9800/pLE12I x MMS372/pMNHS cross (Figure 3.3B). Following conjugation, we probed 100 transconjugant cells (MMS372/pMNHS/pLE12I) with a sex factor-specific probe (Godon et al., 1995) and found that approximately one-third (37.6% ± 2.9% [three assays]) of donor plasmid-containing recipient cells had also acquired the sex factor from the chromosome of the donor strain. Co-transfer frequency of these two conjugative elements is thus in the $10^{-3}$ range. Taking into account the transfer frequency of both conjugative elements ($10^{-3}$ range each), and assuming that they are independently transferred, we expected the number of recipient cells containing both elements to be much lower ($10^{-5}$ range, $7.3 \times 10^{-3}$ multiplied by $5.2 \times 10^{-3}$). This result shows that the conjugative transfer of either pLE12I plasmid or sex factor can happen independently but that co-transfer of both elements occurs quite frequently. We thus consider that pLE12I and sex factor co-transfer is synergistic, suggesting that when the conjugative pores are created to allow mobilization of the chromosomal sex factor between donor and recipient cells, transfer of the conjugative plasmid, present in approximately 25 copies (Ichiyanagi et al., 2003) in the cytoplasm of the donor strain, is facilitated.

To analyze Ll.LtrB dissemination from these two elements, we randomly picked 10 colonies corresponding to recipient cells that had acquired both pLE12I

plasmid and sex factor. Plasmid mixes (Figure 3.3D) were retransformed in DH5α, and patches were first hybridized with WT gpII probe (Figure 3.1B) that specifically identified the mobility products generated from the sex factor (Table 3.1, NZ/pLE12I x MMS/pMNHS, WT gpII). The general gpII probe (Figure 3.1B) was then used to reveal all Ll.LtrB mobility products generated from both elements (Table 3.1, NZ/pLE12I x MMS/pMNHS, gpII). The number of Ll.LtrB mobility products that originated from pLE12I plasmid was calculated by subtracting the products generated by sex factor (WT gpII) from the total amount of Ll.LtrB mobility products (gpII) (Table 3.1, NZ/pLE12I x MMS/pMNHS, gpII minus WT gpII). We found that for the majority of isolates (8 of 10), both conjugative elements generated Ll.LtrB mobility products. As previously observed (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR and NZΔ*ltrB*/pLE12I x MMS/pMNHS-CR), the efficiency of Ll.LtrB mobility from the sex factor is broadly variable (19 to 98%) (Table 3.1, NZ/pLE12I x MMS/pMNHS). However, in this assay, average mobility efficiency of the Ll.LtrB intron expressed from the sex factor is significantly higher (Figure 3.3D) (76.2% ± 8.2% versus 44.2% ± 10.0%). On the other hand, Ll.LtrB dissemination from pLE12I showed more variability (0 to 39%) than previously observed (Table 3.1, NZ/pLE12I x MMS/pMNHS) (Belhocine et al., 2004). These results suggest that plasmid and sex factor transfer almost invariably stimulates Ll.LtrB dissemination, leading to the generation of mobility products upon their transfer to a recipient cell. In the two isolates showing no Ll.LtrB mobility products generated from pLE12I, 98% of recipient plasmids are interrupted by Ll.LtrB intron produced from the sex factor. This suggests that the sex factor was first transferred to the recipient cell,

where its copy of Ll.LtrB invaded almost all available homing sites on recipient plasmids before pLE12I was acquired. This scenario would prevent establishment of Ll.LtrB mobility products generated from pLE12I plasmid. Nevertheless, it is difficult to compare this experiment with the previous crosses presented here and earlier (Belhocine et al., 2004) because of the extra steps involved in identifying recipient cells that had acquired both elements before scoring Ll.LtrB mobility efficiencies.

## 3.6 Discussion

Horizontal transfer of group II introns between organisms, a well-accepted model of intron dissemination and evolution (Lambowitz and Belfort, 1993; Zimmerly et al., 2001), was only recently experimentally corroborated (Belhocine et al., 2004). This model was proposed to explain the presence of closely related introns at different locations (genes or species) and to rationalize why, in specific genes, introns are more conserved than their flanking exons (Lambowitz and Belfort, 1993; Zimmerly et al., 2001). Interestingly, the great majority of bacterial group II introns is found associated with other mobile elements, such as transposons, insertion sequence (IS) elements, conjugative plasmids, pathogenicity islands, and virulence plasmids (Martinez-Abarca and Toro, 2000; Dai and Zimmerly, 2002). Furthermore, a phylogenetic study using numerous bacterial intron-encoded proteins suggests that mobile group II introns originated in the bacterial kingdom and that these prokaryotic introns seem to be subject to a high level of horizontal transfer (Zimmerly et al., 2001). Taking these facts into consideration, we hypothesized that these various mobile elements carrying group

II introns promote dissemination of their hitchhikers to new sites following their transfer between cells (Belhocine et al., 2004).

As a first initiative, using plasmid conjugation assays, we showed that the Ll.LtrB group II intron, present within the *L. lactis* pRS01 conjugative plasmid, can be disseminated and laterally transferred following its conjugative transfer (Belhocine et al., 2004). The engineered plasmids used in that study contained a small segment (Tra1-2; 7.5 kb) of the pRS01 conjugative plasmid (48.4 kb) that harboured the relaxase gene interrupted by the Ll.LtrB group II intron. Remaining functions of the conjugation machinery were provided by the integrative and conjugative sex factor embedded in the chromosome of the donor strain (Belhocine et al., 2004). Using that system, Ll.LtrB dissemination was shown among *L. lactis* strains (retrohoming and retrotransposition) and from *L. lactis* to *E. faecalis* (retrohoming) (Belhocine et al., 2004).

Here, we demonstrated that a fraction of Ll.LtrB mobility products, present in transconjugant cells showing an unusually high level of mobility products after plasmid acquisition (Belhocine et al., 2004), was wild type and, thus, not produced from plasmid. We showed that wild-type Ll.LtrB mobility products originated from the chromosome of the donor strain and were generated by the RNA-mediated retrohoming pathway. We then demonstrated that the sex factor present within the chromosome of NZ9800 strain can be efficiently transferred by conjugation to the MMS372 strain and can support Ll.LtrB dissemination at highly variable rates ranging from 6 to 96%. Even if pLE12I contains an Ll.LtrB variant that is approximately 33% less proficient than wild-type Ll.LtrB in transconjugant MMS372 cells (Belhocine et al., 2004),

frequencies of wild-type Ll.LtrB mobility from the sex factor can be much higher (almost 100%) than ever observed in any of our plasmid conjugation assays (~20%) (Table 3.1) (Belhocine et al., 2004) or when these two plasmids (pLE12I and pMNHS) are simply co-transformed in the same strain (11%) (Cousineau et al., 1998). Furthermore, taking into consideration that the sex factor is maintained at 1 copy per cell while pLE-based plasmids are at 25 copies per cell, efficiency of mobility of the Ll.LtrB intron is significantly higher from the sex factor. These results suggest that regulation of production of active Ll.LtrB RNP particles is different between the whole conjugative element and the simplified context of the engineered conjugative plasmids. This finding is biologically relevant, because the Ll.LtrB group II intron can be spread among *L. lactis* strains, using the native chromosomal sex factor (50 kb) as a carrier. It also suggests that spreading of Ll.LtrB among *L. lactis* strains by sex factor conjugative transfer may have previously occurred and could still occur naturally, because these two mobile elements are active in the NZ9800 *L. lactis* laboratory strain.

Following sex factor conjugative transfer, we also found Ll.LtrB mobility products within a recipient cell that did not harbour the sex factor (Figure 3.3C, isolate 6). This implies that the sex factor can transiently express active Ll.LtrB RNP particles and promote intron dissemination even if it is not permanently retained within recipient cells following conjugative transfer. Indeed, it has been shown that sex factor can be lost from its host strains (Gasson et al., 1992). Isolate 6 showed the lowest efficiency of Ll.LtrB mobility among the nine isolates examined in this experiment, perhaps a result of the transient presence of sex factor in the recipient cell. Furthermore, uptake of Ll.LtrB RNP particles directly

from medium is an unlikely means of transmission of this mobile element, because RNase was spread on the mating plates.

### 3.6.1 Unusually high levels of Ll.LtrB mobility products following plasmid conjugation

Figure 3.4 shows different potential outcomes of a typical cross between NZ9800/pLE12I and MMS372/pMNHS. Because the donor strain contains both pLE12I conjugative plasmid and a proficient copy of the integrative and conjugative sex factor embedded within its chromosome, three different types of transconjugant cells can be produced. Acquisition of the intron by recipient cells can occur by conjugation of pLE12I plasmid (1a), sex factor (1b), or both mobilizable elements (1c). We demonstrated that acquisition of pLE12I plasmid by the recipient cell supports Ll.LtrB mobility to relatively low levels (~20%), potentially creating two types of mobility products that do or do not harbour the *td* group I intron (step 2a) (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR and NZΔ*ltrB*/pLE12I x MMS/pMNHS-CR) (Belhocine et al., 2004). On the other hand, acquisition of sex factor promotes spreading of wild-type Ll.LtrB intron at higher and more variable efficiencies (6 to 96%) and creates only one type of mobility product (step 2b) (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR and NZ/pLE12I x MMS/pMNHS). Finally, when pLE12I plasmid and sex factor are co-transferred into the same recipient cell, they can both independently promote mobility of their intron variant with the potential of creating three types of mobility products (step 2c) (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR and NZ/pLE12I x MMS/pMNHS).

Using different probes that are specific for each of the three types of Ll.LtrB mobility products, plus a general probe that recognizes all of them (Figure 3.1B), we demonstrated that the unusually high level of mobility products consists of those generated from both pLE12I plasmid and sex factor (Figure 3.4, step 1c). In our standard conjugation/retrohoming assays, we isolated 10 independent transconjugant cells ($Cam^r/Spc^r$) that had received the donor plasmid, thus selecting for only two (Figure 3.4, 1a and 1c) of the three possible outcomes of the cross (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR) (Belhocine et al., 2004). We typically observed two or three isolates showing unusually high efficiency of Ll.LtrB mobility and reduced *td* loss (Table 3.1, NZ/pLE12I x MMS/pMNHS-CR) (Belhocine et al., 2004). Taking into consideration that we studied the plasmid mix from only 10 transconjugant cells, a relatively small sample, this ratio (20 to 30%) correlates with pLE12I and sex factor co-transfer frequency (37.6% ± 2.9%).

### 3.6.2 Group II intron dissemination through conjugation

Our study demonstrates that *L. lactis* sex factor can efficiently disseminate the Ll.LtrB mobile group II intron between *L. lactis* strains. Our findings support the theory that bacterial group II introns can be spread from cell to cell following mobilization of their host elements (Belhocine et al., 2004). Furthermore, the association of group II introns with various mobile elements in bacteria may have been a means of survival for these introns and could explain why a great majority of them are found within other mobile elements (Belhocine et al., 2004).

This intron dissemination theory is further strengthened by recent studies. It was demonstrated that the Ll.LtrB group II intron has a marked preference to

retrotranspose into plasmids rather than chromosomal target sites (Ichiyanagi et al., 2003). It was proposed that this bias for plasmid invasion is linked to the nature of the retrotransposition pathway and to target accessibility (Ichiyanagi et al., 2003). This propensity to invade plasmids may explain the unexpectedly high representation of bacterial group II introns within plasmids. Moreover, this bias seems advantageous for these mobile elements because invasion of plasmids limits host damage and could also promote their lateral transfer by conjugation if the invaded plasmid is mobilizable. This inclination of group II introns to invade plasmids may ultimately lead to their better chance of survival.

It was recently shown that other relaxase genes present within the enterococcal conjugative plasmid pCF10 and the streptococcal conjugative transposon Tn5252 can also be invaded by the Ll.LtrB intron from *L. lactis* (Staddon et al., 2004). The site invaded by Ll.LtrB lies within a conserved functional domain of the relaxase enzymes (Staddon et al., 2004). If Ll.LtrB jumps into these available target sites in nature, newly inserted copies of the intron could be spread between cells by these new conjugative carriers, potentially increasing its dispersal. Indeed, the enterococcal plasmid pCF10, like pRS01 plasmid and sex factor from *L. lactis*, retained conjugative function following Ll.LtrB acquisition (Staddon et al., 2004). This means that splicing of Ll.LtrB from its newly invaded site is efficient enough to produce sufficient amounts of relaxase enzyme to support plasmid transfer. However, it also implies that active RNP particles are produced upon splicing and that Ll.LtrB could invade new sites following conjugative transfer of pCF10.

Another *L. lactis* mobilizable plasmid (pAH90) was shown to harbour a copy of the Ll.LtrB intron that interrupts its relaxase gene at the same position (O'Sullivan et al., 2001). This Ll.LtrB-invaded plasmid can probably also support the dissemination of its stowaway because, like the other Ll.LtrB-harbouring elements, it retained its conjugative function (O'Sullivan et al., 2001). By targeting a conserved region within relaxase genes that code for essential enzymes involved in initiating the transfer of their conjugative elements, Ll.LtrB ensures its efficient dissemination and survival. This may explain why identical proficient copies of the Ll.LtrB intron, which are probably recent acquisitions, have been found in three different natural conjugative elements of *L. lactis*: chromosomal sex factor, pRS01, and pAH90. Conjugation, the most efficient way to transfer genetic information between bacterial cells and even across phyla (Lambowitz and Belfort, 1993), may have been and probably still is an efficient mode of dissemination, not only for Ll.LtrB but also for other group II introns present on various conjugative elements.

## 3.7 Acknowledgements

We thank M. Belfort for kindly providing NZ9800Δ*ltrB*::Tet$^r$ strain and pLHS-CR plasmid. We also thank N. H. Acheson, J. W. Coulton, K. Fiola, R. A. Lease, and H. Le Moual for providing comments and J. Kashul for editing the manuscript. B.C. is a CIHR New Investigator Scholar and a McGill University William Dawson Scholar. Research in the B.C. laboratory is supported by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada.

## 3.8 References

**Belfort, M., Derbyshire V., Parker M.M., Cousineau B. and Lambowitz A.M.** (2002) Mobile introns: pathways and proteins, p. 761-783. In N. L. Craig, R. Craigie, M. Gellert, and A. M Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D.C.

**Belfort, M., Ehrenman K. and Chandry P. S.** (1990) Genetic and molecular analysis of RNA splicing in *Escherichia coli*. *Methods Enzymol.* **181:** 521-539

**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.

**Burrus, V., Pavlovic G., Decaris B. and Guédon G.** (2002) Conjugative transposons: the tip of the iceberg. *Mol. Microbiol.* **46:** 601-610.

**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.

**Cousineau, B., Lawrence S., Smith D. and Belfort M.** (2000). Retro-transposition of a bacterial group II intron. *Nature* **404:** 1018-1021. Erratum in (2001) *Nature* **414:** 84.

**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.

**Dai, L. and Zimmerly S.** (2003) ORF-less and reverse-transcriptase-encoding group II introns in archaebacteria, with a pattern of homing into related group II intron ORFs. *RNA* **9:** 14-19.

**Gasson, M.J., Swindell S., Maeda S. and Dodd H.M.** (1992) Molecular rearrangements of lactose plasmid DNA associated with high-frequency transfer and cell aggregation in *Lactococcus lactis* 712. *Mol. Microbiol.* **6:** 3213-3223.

**Godon, J.-J., Pillidge C.J., Shearman C.A. and Gasson M.J.** (1995) Molecular analysis of the *Lactococcus lactis* sex factor. *Dev. Biol. Stand.* **85:** 423-430.

**Ichiyanagi, 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.

**Ichiyanagi, 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.

**Klein, J. R., Chen Y., Manias D.A., Zhuo J., Zhou L., Peebles C.L. and Dunny G.M.** (2004) A conjugation-based system for genetic analysis of group II intron splicing in *Lactococcus lactis. J. Bacteriol.* **186:** 1991-1998.

**Lambowitz, A.M. and Belfort M.** (1993). Introns as mobile genetic elements. *Annu. Rev. Biochem.* **62:** 587-622.

**Martinez-Abarca, F. and Toro N.** (2000) Group II introns in the bacterial world. *Mol. Microbiol.* **38:** 917-926.

**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.

**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.

**Mills, D.A., Manias, D.A., McKay, L.L. and Dunny, G.M.** (1997). Homing of a group II intron from *Lactococcus lactis* subsp. *lactis* ML3. *J. Bacteriol.* **179:** 6107-6111.

**O'Sullivan, D., Ross, R.P., Twomey, D.P., Fitzgerald, G.F., Hill, C. and Coffey, A.** (2001). Naturally occurring lactococcal plasmid pAH90 links

bacteriophage resistance and mobility functions to a food-grade selectable marker. *Appl. Environ. Microbiol.* **67:** 929-937.

**Rambo, R. P. and Doudna J. A.** (2004) Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry* **43:** 6486-6497.

**Saldanha, R., Mohr G., Belfort M. and Lambowitz A. M.** (1993) Group I and group II introns. *FASEB J.* **7:** 15-24.

**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.

**Staddon, J.H., Bryan E.M., Manias D.A., and Dunny G.M.** (2004) Conserved target for group II intron insertion in relaxase genes of conjugative elements of gram-positive bacteria. *J. Bacteriol.* **186:** 2393-2401.

**Toro, N.** (2003) Bacteria and archaea group II introns: additional mobile genetic elements in the environment. *Environ. Microbiol.* **5:** 143-151.

**Trieu-Cuot, P., Carlier C. and Courvalin P.** (1998) Conjugative plasmid transfer from *Enterococcus faecalis* to *Escherichia coli*. *J. Bacteriol.* **170:** 4388-4391.

**Zhou, L., Manias D.A. and Dunny G.M.** (2000) Regulation of intron function: efficient splicing in vivo of a bacterial group II intron requires a functional promoter within the intron. *Mol. Microbiol.* **37:** 639-651.

**Zimmerly, S., Hausner G. and Wu X.** (2001) Phylogenetic relationship among group II intron ORFs. *Nucleic Acids Res.* **29:** 1238-1250.

## 3.9 Figure legends

### Figure 3.1

Mobility assay and probes. **A.** Scoring of intron mobility efficiency after conjugation (colony patch hybridization). Plasmid mixes (donor, recipient, and mobility products) from 10 independent transconjugants were recovered (1) and retransformed in *E. coli* DH5α cells (LB/Spc) (2). One-hundred colonies were patched on LB/Spc plates (3) and transferred onto a nylon membrane that was then hybridized with an intron-specific probe (4) to calculate the ratio of recipient plasmids that received the intron. **B.** Different probes used to identify the two categories of Ll.LtrB mobility products. The mobility products observed were generated either from the plasmid (twintron Ll.LtrB) or the sex factor (wild-type Ll.LtrB). The wild-type intron generates only one kind of mobility product, while the twintron potentially generates two types of mobility products, harbouring or lacking the *td* intron. The group II probe (gpII) recognizes the three types of mobility products. The other three probes (gpI SJ, gpI, and WT gpII) are specific and recognize only one type of mobility product. The *td* group I intron splice junction probe (gpI SJ) recognizes the two *td* ligated exons (12 nucleotides for each exon) and gives a positive signal only upon group I intron loss. The group I probe (gpI) hybridizes within the *td* group I intron and reveals twintron mobility products still harbouring it. The wild-type group II probe (WT gpII) recognizes

wild-type Ll.LtrB mobility products and does not hybridize with either of the two twintron mobility products. Ll.LtrB group II intron, black; Ll.LtrB exons, E1 and E2; *td* intron, I; *td* exons, grey.

**Figure 3.2**

Conjugation/retrohoming assays. **A.** Schematic of the assays. The first step (1) represents the transfer of pLE12I by conjugation from donor (NZ9800 or NZ9800Δ*ltrB*) to recipient cell (MMS372). The second step (2) shows the invasion (retrohoming) of some homing sites (E1-E2) present on the recipient plasmid (pMNHS) by the Ll.LtrB intron expressed from donor plasmid (pLE12I). Intron-harbouring recipient plasmids (pMNHS+) are represented in the transconjugant cell. Recipient plasmid used in these assays contained polymorphic sites in both exons (E1 and E2) (pMNHS-CR). Mobility products harbouring the *td* intron are not depicted. Ll.LtrB group II intron, black; Ll.LtrB exons, E1 and E2; *td* intron, I; *td* exons, grey. **B.** and **C.** Agarose (0.5%) gels containing undigested plasmid mixes recovered from the progeny of 10 transconjugant cells ($Spc^r/Cam^r$). Bands corresponding to donor plasmid (D), recipient plasmid (R), and mobility product (M) are highlighted.

**Figure 3.3**

Conjugation/retrohoming assays. **A.** and **B.** Schematic of the assays. The first step (1) represents the transfer of sex factor (A) or both sex factor and pLE12I plasmid (B) by conjugation from donor (NZ9800) to recipient cell (MMS372). The second step (2) shows invasion (retrohoming) of some homing

sites (E1-E2) present on the recipient plasmid (pMNHS) by Ll.LtrB introns expressed either from pLE12I or sex factor. Intron-harbouring recipient plasmids (pMNHS+) are also represented in the transconjugant cell. Chromosomes are depicted as circular double-stranded helices. Ll.LtrB group II intron, black; Ll.LtrB exons, E1 and E2. **C.** and **D.** Agarose (0.5%) gels containing undigested plasmid mixes recovered from transconjugants containing sex factor (C) or both sex factor and pLE12I plasmid (D). Bands corresponding to donor plasmid (D), recipient plasmid (R), and mobility product (M) are highlighted. Ll.LtrB mobility efficiency, % Mob; +, presence of sex factor (Sex F); –, absence of sex factor.

**Figure 3.4**

Plasmid- and sex factor-based Ll.LtrB dissemination between *L. lactis* strains. A conjugation assay between the donor NZ9800/pLE12I and recipient MMS372/pMNHS strains can generate three types of transconjugant cells. In the first step (1), the recipient cell acquires pLE12I plasmid (1a), sex factor (1b), or both mobilizable elements (1c) by conjugation. The second step (2) represents mobility of the Ll.LtrB intron from its carrying element(s) (pLE12I, sex factor) to the homing site on the recipient plasmid (pMNHS). The percentage and nature of Ll.LtrB mobility products (pMNHS+) depends on which Ll.LtrB-harbouring element(s) is present within the recipient cell. Ll.LtrB group II intron, black; Ll.LtrB exons, E1 and E2; *td* intron, I; *td* exons, grey.

## 3.10 Tables

**Table 3.1** Efficiency of LI.LtrB mobility following its transfer by conjugation[a]

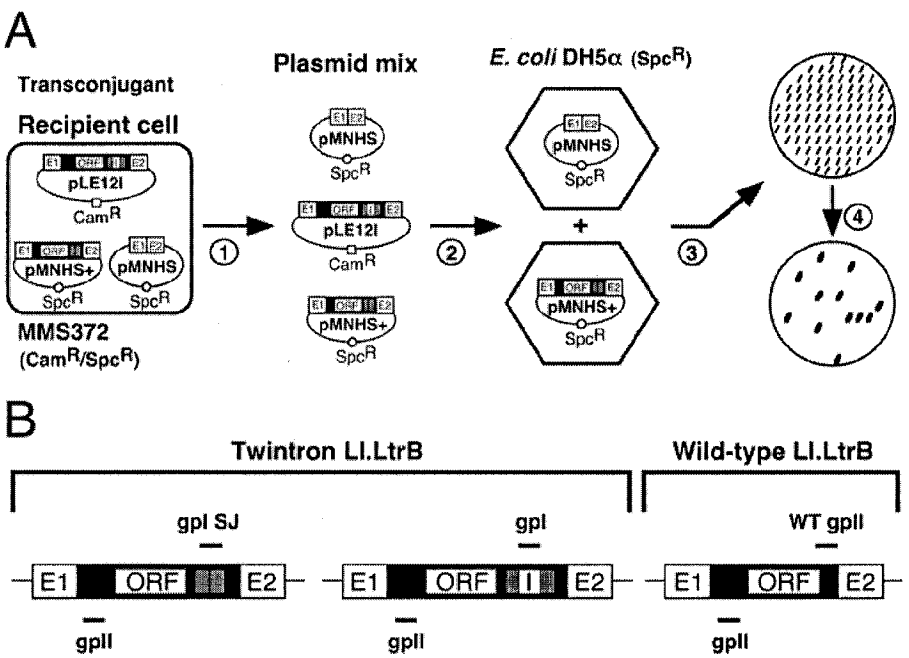
| Isolate no. | Mobility efficiency (%) after conjugation: | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NZ/pLE12I × MMS/pMNHS-CR | | | | | NZΔ*ltrB*/pLE12I × MMS/pMNHS-CR[b] | | | | NZ/pLE12I × MMS/pMNHS | | | NZ/pLE12I × MMS/pMNHS[c] | | | |
| | gpII | gpI SJ | gpI SJ/gpII | gpI | WT gpII | gpII | gpI SJ | gpI SJ/gpII | gp I | gpII | WT gpII | gpII-WT gpII | gpII | gpI SJ | gpI SJ/gpII | WT gpII |
| 1 | *61* | 25 | 41 | 0 | 36 | 13 | 13 | 100 | 0 | 95 | 91 | 4 | *63* | 16 | 25 | 47 |
| 2 | *51* | 17 | 33 | 0 | 34 | 12 | 12 | 100 | 0 | 88 | 79 | 9 | 34 | 34 | 100 | 0 |
| 3 | 21 | 21 | 100 | 0 | 0 | 12 | 12 | 100 | 0 | 98 | 98 | 0 | | | | |
| 4 | *70* | 9 | 13 | 0 | 61 | 19 | 19 | 100 | 0 | 71 | 50 | 21 | | | | |
| 5 | 25 | 25 | 100 | 0 | 0 | 24 | 24 | 100 | 0 | 98 | 59 | 39 | | | | |
| 6 | 13 | 13 | 100 | 0 | 0 | 13 | 13 | 100 | 0 | 92 | 87 | 5 | | | | |
| 7 | 51 | 51 | 100 | 0 | 0 | 14 | 14 | 100 | 0 | 93 | 85 | 8 | | | | |
| 8 | 13 | 13 | 100 | 0 | 0 | 11 | 10 | 91 | 1 | 43 | 19 | 24 | | | | |
| 9 | 43 | 43 | 100 | 0 | 0 | 6 | 6 | 100 | 0 | 98 | 98 | 0 | *86* | 29 | 34 | 57 |
| 10 | 19 | 18 | 95 | 1 | 0 | 16 | 16 | 100 | 0 | 98 | 96 | 2 | | | | |
| Avg ± SEM | 26.4 ± 5.6 | 23.5 ± 4.3 | | | | 14.0 ± 1.5 | 13.9 ± 1.6 | | | 87.4 ± 5.6 | 76.2 ± 8.2 | 11.2 ± 4.1 | 26.3 ± 5.4 | | | |

[a] The mobility efficiencies (%) of the LI.LtrB intron were calculated by colony patch hybridization assays (Materials and Methods and Fig. 1A) analyzing 100 colonies for each of the 10 independent transconjugants (1 to 10). Unexpectedly high LI.LtrB mobility efficiencies are italicized and not included in the means. NZ9800 (NZ) and MMS372 (MMS) are *L. lactis* strains; NZΔ*ltrB* is NZ9800Δ*ltrB*::Tet$^r$ (11). gpII, number of colonies revealed with the group II intron probe (5′ end, Fig. 1B); gpI SJ, number of colonies revealed with the *td* group I intron splice junction probe (ligated exons, Fig. 1B); gpI SJ/gpII, % of LI.LtrB mobility events missing the *td* group I intron; gpI, number of colonies revealed with the *td* group I intron probe (Fig. 1B); WT gpII, number of colonies revealed with the wild-type group II intron probe (3′ end, Fig. 1B). The recipient strain in our previous study (3) contained the pMNHS plasmid.

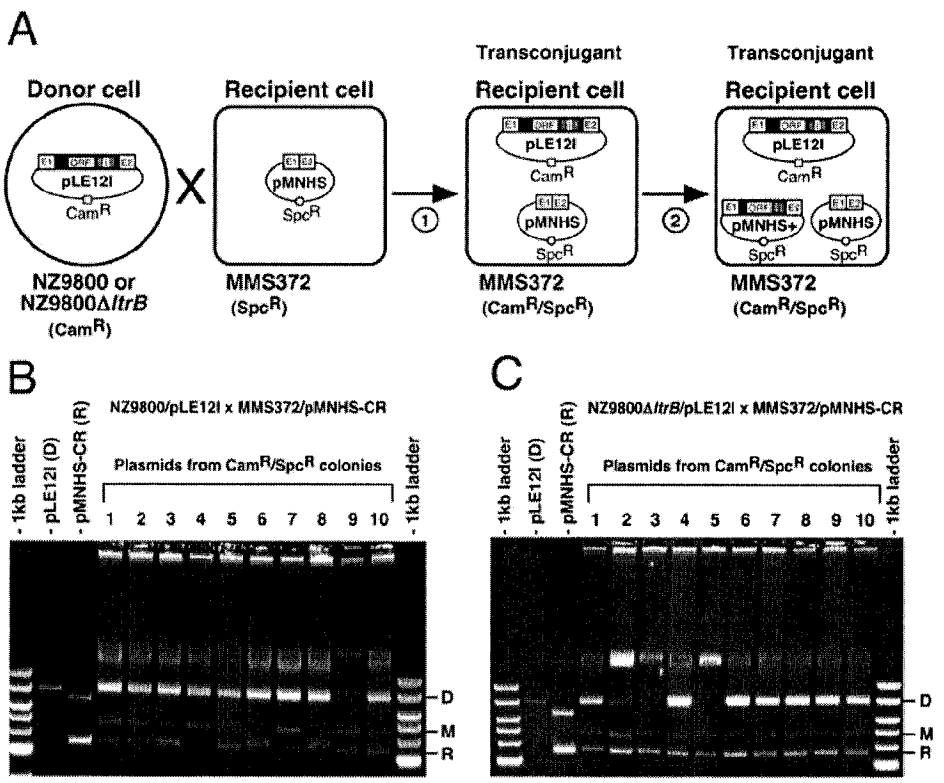
[b] The mobility efficiency detected with the WT gpII probe was zero for all transconjugants.

[c] From reference 3. The mobility efficiency detected with the gpI probe was zero for all transconjugants.

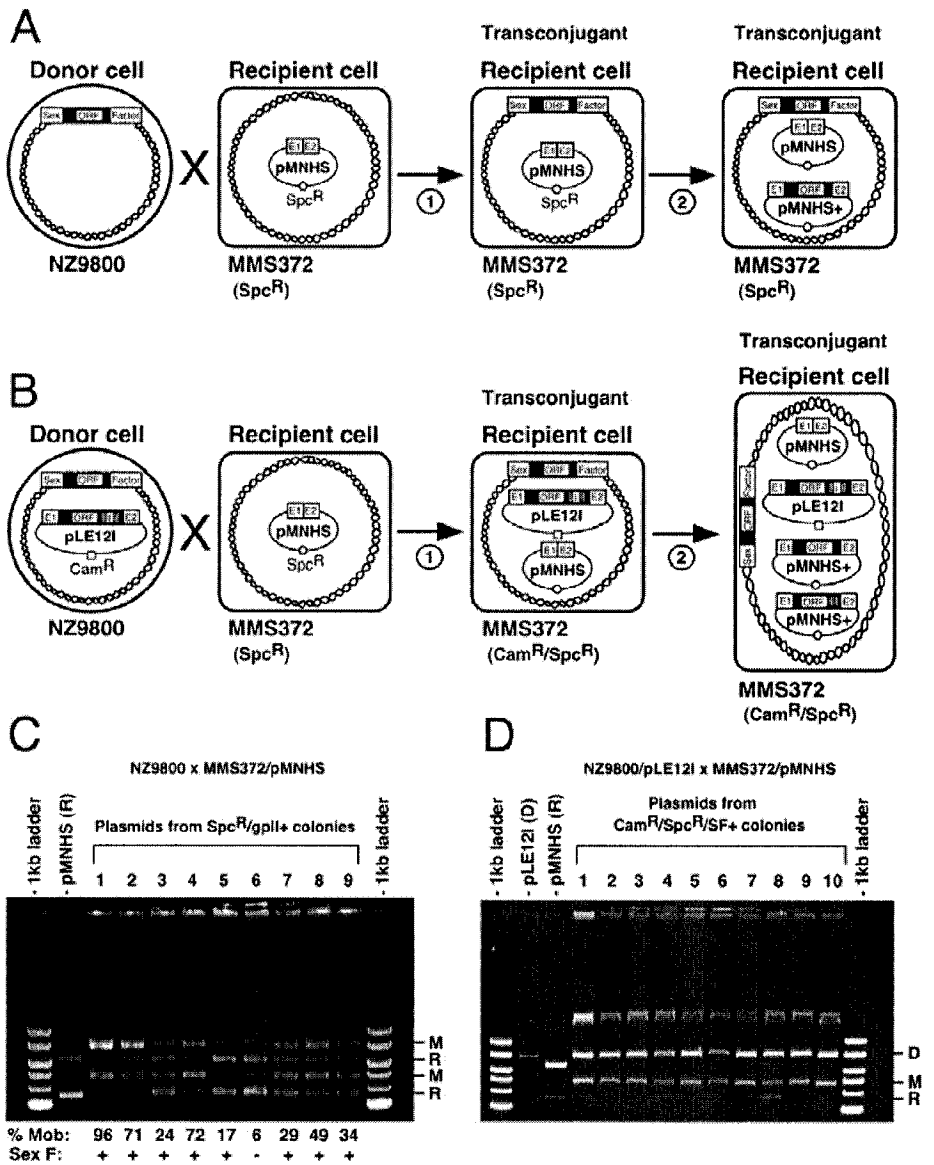
# Figure 3.1


A
Transconjugant
Recipient cell
pLE12I
CamR
pMNHS+
pMNHS
SpcR
SpcR
MMS372
(CamR/SpcR)
Plasmid mix
pMNHS
SpcR
pLE12I
CamR
pMNHS+
SpcR
E. coli DH5α (SpcR)
pMNHS
SpcR
+
pMNHS+
SpcR
1
2
3
4
B
Twintron LI.LtrB
Wild-type LI.LtrB
gpI SJ
gpI
WT gpII
E1
ORF
E2
gpII
gpII
gpII

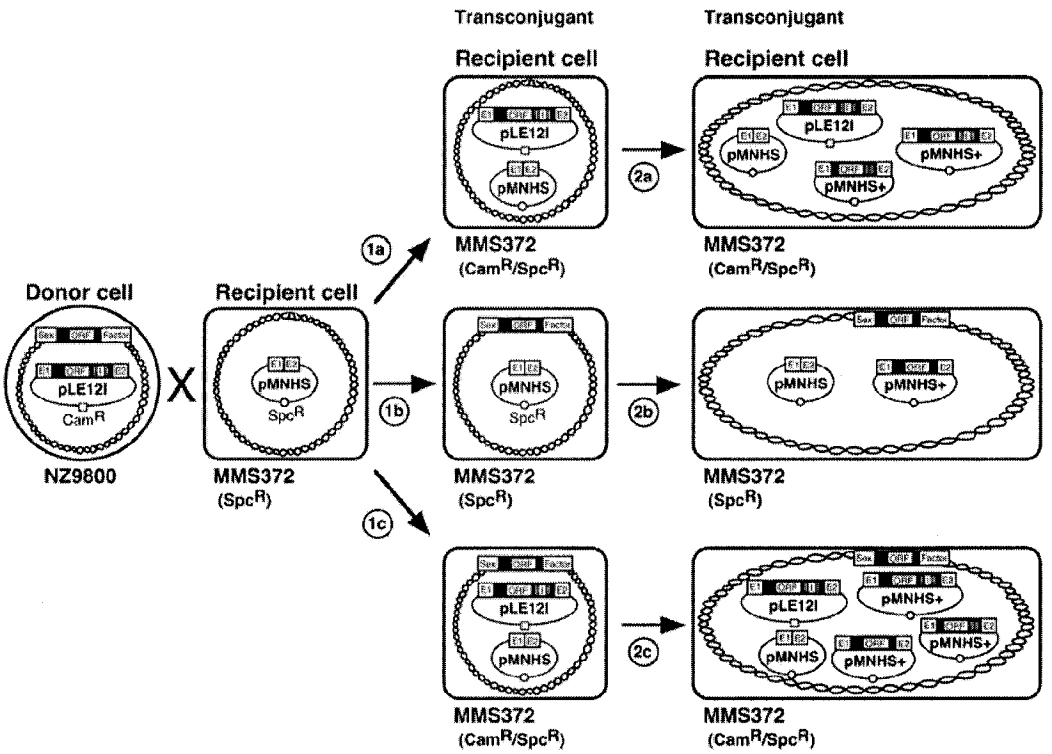
# Figure 3.2


A
Donor cell
Recipient cell
Transconjugant
Recipient cell
Transconjugant
Recipient cell
pLE12I
CamR
pMNHS
SpcR
pMNHS+
NZ9800 or
NZ9800ΔltrB
(CamR)
MMS372
(SpcR)
MMS372
(CamR/SpcR)
MMS372
(CamR/SpcR)
B
C
NZ9800/pLE12I x MMS372/pMNHS-CR
NZ9800ΔltrB/pLE12I x MMS372/pMNHS-CR
Plasmids from CamR/SpcR colonies
- 1kb ladder
- pLE12I (D)
- pMNHS-CR (R)
1 2 3 4 5 6 7 8 9 10
D
M
R

# Figure 3.3


A
Donor cell
Recipient cell
Transconjugant
Recipient cell
Transconjugant
Recipient cell
pMNHS
SpcR
pMNHS+
NZ9800
MMS372
(SpcR)
MMS372
(SpcR)
MMS372
(SpcR)
B
Donor cell
Recipient cell
Transconjugant
Recipient cell
Transconjugant
Recipient cell
pLE12I
CamR
NZ9800
MMS372
(SpcR)
MMS372
(CamR/SpcR)
MMS372
(CamR/SpcR)
C
NZ9800 x MMS372/pMNHS
1kb ladder
pMNHS (R)
Plasmids from SpcR/gpII+ colonies
1 2 3 4 5 6 7 8 9
1kb ladder
M
R
M
R
% Mob: 96 71 24 72 17 6 29 49 34
Sex F: + + + + + - + + +
D
NZ9800/pLE12I x MMS372/pMNHS
1kb ladder
pLE12I (D)
pMNHS (R)
Plasmids from
CamR/SpcR/SF+ colonies
1 2 3 4 5 6 7 8 9 10
1kb ladder
D
M
R

# Figure 3.4


Transconjugant
Recipient cell
Transconjugant
Recipient cell
pLE12I
pMNHS
pMNHS+
MMS372
(CamR/SpcR)
MMS372
(CamR/SpcR)
Donor cell
Recipient cell
1a
2a
pLE12I
CamR
pMNHS
SpcR
1b
2b
NZ9800
MMS372
(SpcR)
MMS372
(SpcR)
MMS372
(SpcR)
1c
2c
MMS372
(CamR/SpcR)
MMS372
(CamR/SpcR)

# CHAPTER FOUR

## Conjugative Transfer of the *Lactococcus lactis* sex factor and pRS01 plasmid to *Enterococcus faecalis*

### 4.1 Preface

In Chapter Two, we reported that the pLE12 mobilizable plasmid can be transferred from *Lactococcus lactis* to *Enterococcus faecalis*. pLE12 harbours the origin of transfer (*oriT*) of pRS01/sex factor. Therefore, it is specifically recognized and transferred by the conjugative machinery of the *L. lactis* chromosomal sex factor. This finding implies that the sex factor machinery is able to promote conjugative transfer to other bacterial species. However, it was previously stated that the *L. lactis* sex factor could not be transferred to other species (Godon et al., 1995).

We demonstrated in Chapter Three that transfer of the sex factor between *L. lactis* strains promotes dissemination of Ll.LtrB. We thus asked whether the sex factor and the closely-related pRS01 plasmid can be transferred to *E. faecalis* by conjugation and promote dissemination of their carried intron in the new host by retrohoming.

## Conjugative transfer of the *Lactococcus lactis* sex factor and pRS01 plasmid to *Enterococcus faecalis*

Kamila Belhocine, Victoria Mandilaras, Bonnie Yeung & Benoit Cousineau

Department of Microbiology and Immunology, McGill University, Montréal, QC, Canada

In: ***FEMS Microbiology Letters*** (2007) **269**:289-294


### 4.2 Abstract

The low G+C Gram-positive bacterium *Lactococcus lactis* harbours two highly similar conjugative elements: an integrative and conjugative element called sex factor and the pRS01 plasmid. Originally, it was believed that the host range of the sex factor was limited to *L. lactis* sub-species. Here, we report that pTRK28 co-integrates of a spectinomycin-marked *L. lactis* sex factor and of the pRS01 conjugative plasmid can be transferred from *L. lactis* to *Enterococcus faecalis*. These results demonstrate the conjugative transfer of these elements to other bacterial species. Furthermore, we report that Ll.LtrB, a mobile group II intron carried by both elements, can invade its recognition site upon pRS01 conjugative transfer to *E. faecalis*.

### 4.3 Introduction

The low G+C Gram-positive bacterium *Lactococcus lactis* harbours two conjugative elements that are nearly identical: an integrative and conjugative element called the sex factor (Gasson *et al.*, 1995), and the pRS01 plasmid (Mills

*et al.*, 1994). These two elements are efficiently transferred between lactococcal strains through conjugation ($10^{-2}$-$10^{-3}$ range) (Gasson *et al.*, 1995; Mills *et al.*, 1994).

The sex factor is a chromosomally embedded element found in *L. lactis* 712. It lacks a functional origin of replication and needs to be integrated in the chromosome of the host to be replicated (Gasson *et al.*, 1995). Before conjugation, the sex factor excises and circularizes (Godon *et al.*, 1995). Once transferred, the sex factor reintegrates the chromosome of its new host via site-specific recombination at its integration site, *attB*, a conserved sequence of 24 nucleotides (Gasson *et al.*, 1995). It has been proposed that the host range of the sex factor is limited to *L. lactis* sub-species and that this element cannot be transferred to other related Gram-positive bacteria, including *Enterococcus faecalis* (Gasson *et al.*, 1995).

The pRS01 plasmid was isolated from *L. lactis* strain ML3. Four transfer regions have been identified in pRS01. The Tra1-2 region bears the origin of transfer (*oriT*), the relaxase gene, as well as other genes essential for conjugative transfer (Mills *et al.*, 1994). The Tra3 and Tra4 regions of pRS01 are believed to code for proteins that mediate contact between the donor and recipient cells during conjugation (Mills *et al.*, 1994).

The Ll.LtrB group II intron interrupts the putative relaxase gene, *ltrB*, in both the sex factor and pRS01 (Mills *et al.*, 1996; Shearman *et al.*, 1996). Self-splicing group II introns are found in bacteria, archaea and organelles of lower eukaryotes and higher plants (Lambowitz and Zimmerly, 2004). Some group II introns, like Ll.LtrB, are also mobile retroelements. In bacteria, group II introns

are often associated with other mobile elements such as transposons and conjugative plasmids. It is believed that group II introns were widely disseminated between species by lateral (horizontal) transfer throughout evolution (Lambowitz and Zimmerly, 2004). These introns are thought to spread following the transfer of their host mobile elements (Belhocine *et al.*, 2004, 2005).

Ll.LtrB was the first bacterial group II intron shown to self-splice and integrate into new DNA sites *in vivo* (Mills *et al.*, 1996, 1997). It harbours an open reading frame, *ltrA*. The LtrA protein (599 amino acids) has three enzymatic activities: reverse transcriptase, maturase and endonuclease. Each of these activities is essential for mobility of the intron into an intronless allele, also called the homing site (HS) (Matsuura *et al.*, 1997; Cousineau *et al.*, 1998).

We showed that Ll.LtrB can be transferred by conjugation of a pRS01-based mobilizable plasmid (pLE12) between *L. lactis* sub-species and from *L. lactis* to the closely related Gram-positive bacterium *E. faecalis* (Belhocine *et al.*, 2004). We also showed that conjugative transfer of the sex factor promotes dissemination of Ll.LtrB between *L. lactis* sub-species (Belhocine *et al.*, 2005). Here we show that co-integrates of both pRS01 and the chromosomal sex factor can be transferred from *L. lactis* to *E. faecalis*, albeit at low efficiencies. Moreover, we show that pRS01 transfer from *L. lactis* to *E. faecalis* supports the dissemination of Ll.LtrB.

## 4.4 Materials and Methods

### 4.4.1 Strains, plasmids and media

The strains and plasmids used in this study are described in Table 4.1. *L.*

*lactis* strains were grown in GM17 at 30°C without shaking. *E. faecalis* JH2-2 was grown in BHI at 37°C without shaking. *E. coli* strains were grown in LB at 37°C with shaking. Selective media contained the following concentration (μg/ml) of antibiotics: erythromycin (Erm), 150; fusidic acid (Fus), 25; chloramphenicol (Cam), 10; tetracycline (Tet), 3; spectinomycin (Spc), 300.

### 4.4.2 Generation of a marked sex factor

The kanamycin resistance gene within Ll.LtrB in the pLEIItd+KR" plasmid (Cousineau *et al.*, 1998) was replaced by the *aad*9 spectinomycin resistance gene (*Sal*I). This marked version of Ll.LtrB was then subcloned into pBlueScript and used for gene replacement to create the NZISp strain.

### 4.4.3 Conjugation assays

Mating was performed on milk plates (intraspecies, *L. lactis*) or on filters (interspecies) as previously described (Belhocine *et al.*, 2004). Conjugation efficiency is the ratio of the number of transconjugant cells to donor cells (three independent assays). To monitor intron mobility following conjugation, the recipient strain contained a recipient plasmid, either pMNHS or pMNHSTet (Cousineau *et al.*, 1998; Belhocine *et al.*, 2004, 2005). Colony-patch hybridization was used to score intron mobility (Belhocine *et al.*, 2004, 2005).

### 4.4.4 Southern blot

Total DNA from transconjugant cells (10 μg) was digested with *Pst*I and

run on a 1% agarose gel. DNA was transferred onto a membrane (Hybond-N, Amersham) and hybridized with an *ltrB*-specific 5'-$^{32}$P end-labeled DNA probe (5'-CATTTGAGGTTCATCAAGCAGC-3'). The membrane was revealed with a phosphor-imager (Bio-Rad FX).

## 4.5 Results and Discussion

### 4.5.1 Generating a spectinomycin-marked sex factor

It was previously proposed that conjugative transfer of the *L. lactis* sex factor to other bacterial species is impossible (Gasson *et al.*, 1995). However, our laboratory has since shown that a mobilizable plasmid, pLE12, containing the Tral-2 region from pRS01 encompassing the origin of transfer (*oriT*), can be transferred to the Gram-positive bacterium *E. faecalis* ($1.6 \times 10^{-7}$). In these experiments, pLE12 transfer relied on the chromosomal sex factor (Belhocine *et al.*, 2004), hinting that its conjugation machinery can sustain interspecies transfer. Therefore, we investigated whether the sex factor or pRS01 can be transferred by conjugation outside of lactococci.

The sex factor is transferred between *L. lactis* sub-species through conjugation at a rate of $7.3 \times 10^{-3}$ (Table 4.2; Belhocine *et al.*, 2005). Furthermore, the Ll.LtrB intron harboured on the sex factor invades its homing site after conjugation at a highly variable rate (6% to 96%) (Table 4.2; Belhocine *et al.*, 2005). To monitor interspecies transfer of the sex factor and Ll.LtrB mobility in the new host, the chromosomal sex factor of *L. lactis* NZ9800 was marked with a spectinomycin resistance gene (Spc) as well as the *td* group I intron (NZISp) (Figure 4.1A). The *td* group I intron allows the detection of Ll.LtrB mobility via

an RNA intermediate (Cousineau *et al.*, 1998).

To verify that the marked sex factor (SFISp) is still proficient in conjugation, mating was performed between NZISp and *L. lactis* LM0230 containing the pMNHSTet recipient plasmid, which harbours the Ll.LtrB homing site. The marked sex factor is transferred less efficiently between *L. lactis* sub-species than the wild-type conjugative element ($10^{-4}$, Table 4.2). This slight drop in conjugation efficiency ($10^{-4}$ vs. $10^{-3}$) was expected since insertion of additional sequences within Ll.LtrB typically hinders splicing efficiency (Plante and Cousineau, 2006). This may reduce the amount of mature relaxase transcript and relaxase enzyme produced, leading to reduced conjugation ability. The presence of foreign sequences within Ll.LtrB also decreases its mobility efficiency in comparison to the wild type intron (<1% - 47% (ten independent transconjugants) vs. 6% - 96%) (Cousineau *et al.*, 1998; Plante and Cousineau, 2006).

Interspecies conjugations of SFISp from *L. lactis* NZISp to *E. faecalis* JH2-2 were unsuccessful. We hypothesized that this non-replicative conjugative element cannot maintain itself in the new host species if transferred. Indeed, only native bacterial hosts of the sex factor would be expected to harbour its attachment site, *attB*, needed for chromosomal insertion and maintenance.

### 4.5.2 Conjugative transfer of the *L. lactis* sex factor to *E. faecalis*

We used the integrative plasmid pTRK28 to supply an origin of replication to the marked sex factor. The pTRK28 plasmid contains an insertion sequence, IS946, which allows it to randomly integrate into DNA (Mills *et al.*, 1994). Integration of pTRK28 within excised sex factor allows this non-replicative

conjugative element to be maintained independently from the chromosome (Figure 4.1B, step 1). To select for conjugation-competent co-integrates (pTRK28::SFISp), NZISp transformed with pTRK28 was conjugated with *L. lactis* LM0231 (Figure 4.1B, step 2). Each transconjugant received a different co-integrate where pTRK28 was inserted into different positions within the marked sex factor. The same procedure was used to generate co-integrates of the wild type sex factor.

Transconjugants from the pTRK28::SFISp pool were selected based on their aggregation phenotype in liquid culture as an indicator for conjugation competence. Three transconjugants of twenty-four had a strong aggregation phenotype. These were used as donor strains for intraspecies conjugation assays (LM0231/pTRK28::SFISp x LM0230/pLEHS). Transfer of pTRK28::SFISp between *L. lactis* strains occurred at rates similar to SFISp ($2.24 \times 10^{-4}$ vs. $8.51 \times 10^{-4}$, Table 4.2). The transconjugant found to have the best conjugation efficiency ($3.50 \times 10^{-4}$) was used as the donor strain for conjugation with *E. faecalis*

We were able to detect the conjugative transfer of pTRK28::SFISp from *L. lactis* LM0231 to *E. faecalis* JH2-2, harbouring or not the pMNHSTet recipient plasmid (Figure 4.1C, step 1; $7.29 \times 10^{-9}$, Table 4.2). Presence of the pTRK28::SFISp co-integrate in *E. faecalis* transconjugants was confirmed by the acquisition of spectinomycin resistance, and ascertained by Southern blot analysis (Figure 4.1D). The identity of transconjugant cells was confirmed by their resistance to rifampicin and by their genomic DNA restriction pattern.

These data constitute the first observation of interspecies transfer of the *L. lactis* sex factor. Therefore, the observed limitation of the sex factor host range is

likely not due to its conjugation machinery, but rather to its instability in a new cellular environment that does not harbour the attachment site, *attB*.

Notably, no mobility of the Spc-marked Ll.LtrB intron was detected in transconjugant cells following the acquisition of the marked sex factor (nine transconjugants analyzed, Figure 4.1C, step 2). This can be attributed to the presence of markers within the intron that affect splicing and/or mobility in *E. faecalis* (Plante and Cousineau, 2006). In fact, we previously observed a significant drop in Ll.LtrB mobility in *E. faecalis* transconjugant cells when the intron carries markers (pLE12IK, 1.2%) in comparison to the wild-type intron (pLE12, 6%) (Belhocine *et al.*, 2004).

Conjugations were also performed using six independent pTRK28 co-integrates with the wild-type sex factor. Intraspecies *L. lactis* conjugations occurred at low frequencies ($5.33 \times 10^{-7}$, Table 4.2), and mating with *E. faecalis* was unsuccessful. The reason for the lowered transfer efficiency of the wild type sex factor co-integrates in comparison to the marked sex factor co-integrates remains unclear.

### 4.5.3 Interspecies conjugation of a pTRK28::pRS01 co-integrate

A pTRK28::pRS01 co-integrate, pM2036, was found to conjugate at high levels with other *L. lactis* species ($5.7 \times 10^{-2}$) (Mills *et al.*, 1994). To determine if pRS01 can be transferred to *E. faecalis*, we attempted to conjugate pM2036 from *L. lactis* MMS370 (strain DM2036; Mills *et al.*, 1994) to *E. faecalis* JH2-2 containing or not the pMNHS plasmid (Figure 4.2A, step 1). Transconjugants were obtained at an efficiency of $6.91 \times 10^{-9}$. Southern blot analysis confirmed the

presence of the pTRK28::pRS01 co-integrate within *E. faecalis* transconjugants (Figure 4.2B). The strain identity of transconjugant cells was confirmed by their resistance to both rifampicin and fusidic acid, and by their genomic DNA restriction pattern.

These data demonstrate that similar to the sex factor, the conjugation machinery of the pRS01 plasmid can mediate interspecies transfer of this conjugative element, although at low efficiency.

### 4.5.4 Mobility of Ll.LtrB following conjugation of the pTRK28::pRS01 co-integrate to *E. faecalis*

Ll.LtrB was already shown to be mobile in *E. faecalis* (Belhocine *et al.*, 2004; Staddon *et al.*, 2004). However, to determine whether Ll.LtrB endogenous to pRS01 mobilized in its new bacterial host (Figure 4.2A, step 2) following conjugation, plasmid mixes from eight *E. faecalis* transconjugants were analyzed. A recipient plasmid harbouring the Ll.LtrB homing site was provided in the recipient strain to track the ability of Ll.LtrB to retrohome in the new host. We first assessed the plasmid mixes by agarose gel electrophoresis (Figure 4.2C). The presence of a characteristic band larger than pMNHS indicates the invasion of a proportion of homing sites by Ll.LtrB (Cousineau *et al.*, 1998; Belhocine *et al.*, 2004, 2005). We scored the precise mobility efficiency of Ll.LtrB from pRS01 in *E. faecalis* using a patch hybridization assay (Belhocine *et al.*, 2004; 2005). We observed variable rates of Ll.LtrB mobility ranging from <1% to 86% (Figure 4.2C, Table 4.2). Interestingly, similar variability in mobility efficiency of the intron after conjugation was observed following transfer of the wild type sex

factor between *L. lactis* sub-species, (6% to 96%; Belhocine *et al.*, 2005). Sequencing of the intron/exon junctions of ten mobility products confirmed that the intron invaded its homing site.

The mobility efficiency of Ll.LtrB supplied by pRS01 in *E. faecalis* (<1% to 86%) is higher than when supplied by the pLE12 mobilizable plasmid (6%; Belhocine *et al.*, 2004). This observation parallels the drop in mobility detected in *L. lactis* when comparing Ll.LtrB retrohoming from the sex factor to retrohoming from pLE12 (6% to 96% vs. 10.5%; Belhocine *et al.*, 2005). Taken together, these results demonstrate that Ll.LtrB mobility is variable and more efficient from the complete conjugative elements (sex factor, pRS01) than from the smaller mobilizable plasmid pLE12, which harbours only parts of the conjugation machinery.

This study demonstrates that the sex factor and pRS01, both conjugative elements native to *L. lactis*, can be transferred outside of lactococci to *E. faecalis* through conjugation. Furthermore, Ll.LtrB originating from pRS01 can be disseminated in this novel host species following conjugative transfer. These results are biologically relevant as pRS01 is a natural host element for Ll.LtrB. Therefore, this work supports the theory suggesting that group II introns are laterally transferred in nature following the relocation of their carrying mobile elements (Belhocine *et al.*, 2004, 2005).

## 4.6 Acknowledgments

We thank G. Dunny for providing the DM2036 strain and pTRK28, and K.K. Yam for critical reading of the manuscript. This work was supported by a

grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to B.C.. B.C. was a CIHR New Investigator Scholar and is an FRSQ Chercheur-Boursier Junior 2 and a McGill University William Dawson Scholar. K.B. holds a Post-Graduate Studies Fellowship from NSERC and V.M. was the recipient of an Undergraduate Summer Research Award (USRA) from NSERC.

## 4.7 References

**Belhocine K., Plante I. and Cousineau B.** (2004) Conjugation mediates transfer of the Ll.LtrB group II intron between different bacterial species. *Mol. Biol.* **51:** 1459-1569.

**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.

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

**Gasson, M.J., Godon, J.-J., Pillidge, C.J., Eaton, T.J., Jury, K. and Shearman, C.A.** (1995) Characterization and exploitation of conjugation in *Lactococcus lactis*. *International Dairy Journal* **5:** 757-762.

**Godon, J.-J., Pillidge, C.J., Jury, K., Shearman, C.A. and Gasson, M.J.** (1995) Molecular Analysis of the *Lactococcus lactis* Sex Factor. *Dev. Biol. Stand.* **85:** 423-430.

**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.

**McKay, L.L. and Baldwin, K.A.** (1975) Plasmid distribution and evidence for a proteinase plasmid in *Streptococcus lactis* C2-1. *Appl. Microbiol.* **29:** 546-548.

**Mills, D.A., Manias, D.A., McKay, L.L. and Dunny, G.M.** (1997) Homing of a group II intron from *Lactococcus lactis* subsp. *lactis* ML3. *J. Bacteriol.* **179:** 6107-6111.

**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.

**Mills, D.A., Choi, C.K., Dunny, G.M. and McKay, L.L.** (1994) Genetic Analysis of Regions of the *Lactococcus lactis* subsp. *lactis* Plasmid pRS01 Involved in Conjugative Transfer. *Appl. Environ. Microbiol.* **60:** 4413-4420.

**Plante, I. and Cousineau, B.** (2006) Restriction for gene insertion within the *Lactococcus lactis* Ll.LtrB group II intron. *RNA* **12:** 1980-92.

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

**Staddon, J.H., Bryan, E.M., Manias, D.A. and Dunny, G.M.** (2004) Conserved target for group II intron insertion in relaxase genes of conjugative elements of gram-positive bacteria. *J. Bacteriol.* **186:** 2393-2401.

## 4.8 Figure legends

### Figure 4.1

Conjugation of the sex factor. **A.** Schematic of the marked sex factor (SFISp). Insertion of markers within Ll.LtrB avoids disruption of genes necessary for conjugation. The group I intron (I) is used to assess Ll.LtrB mobility using an RNA intermediate. The spectinomycin resistance gene (Spc) allows tracking of sex factor transfer and Ll.LtrB mobility. **B.** Generation of a bank of SFISp co-integrates. The integrative plasmid pTRK28 is transformed into *L. lactis* NZISp. SFISp is illustrated in its excised circular form. pTRK28 randomly integrates into circularized SFISp providing the conjugative element with an origin of replication (*oriC*) and a selection marker ($Erm^R$) (step 1). The random integrations generate a bank of SFISp co-integrates (SFISp::pTRK28). To isolate conjugation competent co-integrates, NZISp harbouring pTRK28 is conjugated with another *L. lactis* strain, LM0231 (step 2). A recipient cell that received a co-integrate becomes a pTRK28::SFISp donor cell for interspecies transfer. **C.** Interspecies transfer of SFISp::pTRK28 co-integrates and Ll.LtrB mobility. LM0231 harbouring a SFISp::pTRK28 co-integrate is mated with *E. faecalis* JH2-2. Step 1 shows the transfer of SFISp::pTRK28 from *L. lactis* to *E. faecalis*. The mobility of Ll.LtrB into a recipient plasmid in *E. faecalis* (step 2) can be monitored. **D.** Southern blot analysis of *E. faecalis* SFISp::pTRK28 transconjugants. The *ltrB* specific probe confirms the presence of the sex factor in all representative transconjugants (TC1-3) as well as in the donor strain (LM0231/ SFISp::pTRK28, D), and its absence from the recipient strain (JH2-2/pMNHSTet, R).

**Figure 4.2**

Interspecies transfer of a pRS01 co-integrate and subsequent Ll.LtrB mobilization. **A.** Conjugation of the pTRK28::pRS01 co-integrate (pM2036) to *E. faecalis* and Ll.LtrB mobility. Step 1 shows the conjugative transfer of pTRK28::pRS01 to *E. faecalis* JH2-2 harbouring a Ll.LtrB recipient plasmid. Step 2 shows mobilization of Ll.LtrB into its homing site. **B.** Southern blot analysis of *E. faecalis* pTRK28::pRS01 transconjugants. The *ltrB* specific probe confirms the presence of pRS01 in all representative transconjugants (TC1-3), in the donor strain (DM2036, D), and its absence from the recipient strain (JH2-2/pMNHS, R). **C.** Agarose gel (0.5%) containing plasmid mixes isolated from 8 representative pTRK28::pRS01 transconjugants. Ll.LtrB mobility efficiency is indicated under each lane. R, uninterrupted recipient plasmid; M, Ll.LtrB mobility product; MWM, molecular weight marker.

## 4.9 Tables

**Table 4.1** Relevant characteristics of strains and plasmids

| Strain or plasmid | Relevant characteristics[a] | Source, description and/or reference |
|---|---|---|
| ***L. lactis* strains** | | |
| DM2036 | pTRK28::pRS01, $Erm^R$ | Donor for inter-species conjugation[b] |
| LM0230 | Plasmid free | Recipient for conjugation[b] |
| LM0231 | Plasmid free, $Fus^R$ | Recipient for conjugation with sex factor co-integrate, donor for conjugation of co-integrate[c] |
| MG1363 | Plasmid free, sex factor | Used for generation of sex-factor co-integrates[d] |
| NZ9800 | Plasmid free, sex factor | Derivative of MG1363[e] |
| NZISp | Plasmid free, $Spc^R$ gene and *td* group I intron in Ll.LtrB of sex factor | Used for generation of $Spc^R$ and group I *td* intron marked sex factor co-integrates (this study) |
| ***E. faecalis* strain** | | |
| JH2-2 | Plasmid free, $Rif^R$, $Fus^R$ | Recipient strains for conjugation with *L. lactis* strains[e] |
| **Plasmids** | | |
| pLEHS | $Cam^R$, 9 kb | pLE1 shuttle plasmid with Ll.LtrB homing site (271 bp) (this study) |
| pMNHS | $Spc^R$, 6.9 kb | Shuttle plasmid pDL278 containing Ll.LtrB homing site (271 bp in *Hind*III site)[e,f] |
| pMNTet | $Tet^R$, 9.2 kb | Shuttle plasmid pDL278 containing Ll.LtrB homing site (271 bp in *Hind*III site) with $Tet^R$ gene cloned into $Spc^R$ (*Kpn*I) (this study) |
| pTRK28 | $Erm^R$, 11 kb | pSA3 non-conjugative cloning vector with IS946 insertion sequence[b] |
| pTRK28::pRS01 | $Tra^+$ $Clu^{-/+}$ $Erm^R$, 59.4 kb | Co-integrate found in DM2036[b] |
| pTRK28::SFISp | $Tra^+$ $Clu^{-/+}$ $Erm^R$ $Spc^R$ approx. 70 kb | Conjugative co-integrate formed by pTRK28 integration in NZISp marked sex factor (this study) |

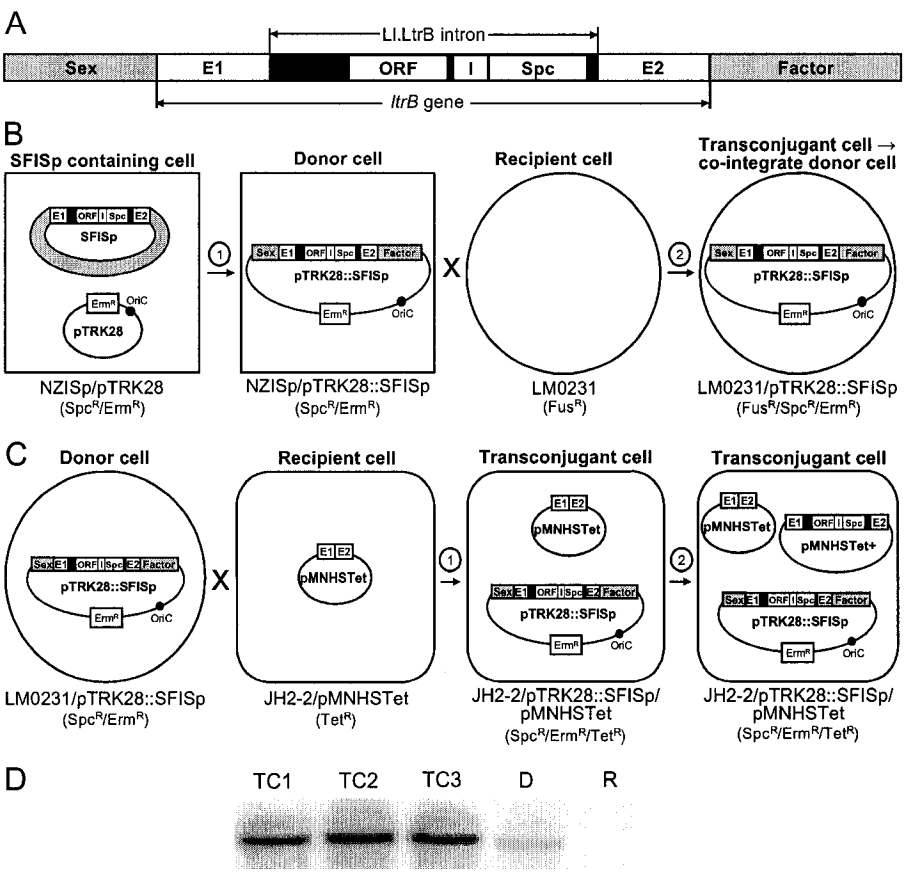
[a] $Cam^R$, chloramphenicol resistant; $Erm^R$, erythromycin resistant; $Fus^R$, fusidic acid resistant, $Rif^R$; rifampicin resistant; $Spc^R$, spectinomycin resistant; $Tet^R$, tetracycline resistant; Tra, conjugative proficiency; Clu, cell aggregation; -, negative; +, positive.
[b] (Mills *et al.*, 1994)
[c] (McKay and Baldwin, 1975)
[d] (Gasson *et al.*, 1995)
[e] (Belhocine *et al.*, 2004)
[f] (Mills *et al.*, 1997; Cousineau *et al.*, 1998)

**Table 4.2** Conjugation efficiencies of the sex factor and pRS01 co-integrates, and Ll.LtrB mobility efficiency in *E. faecalis*

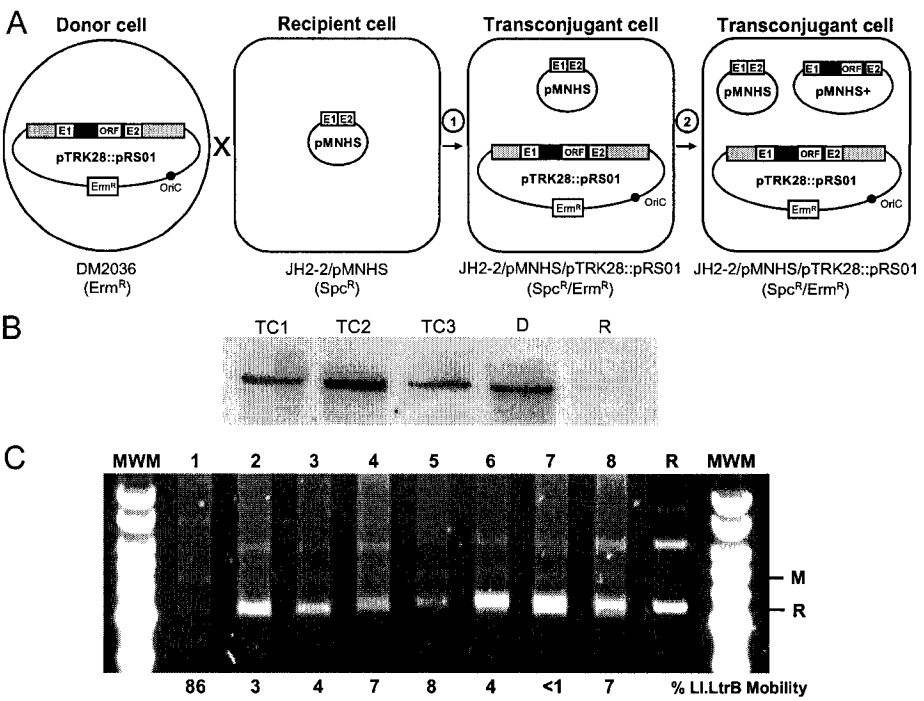
| Conjugative Element | Conjugation Efficiency | | Ll.LtrB Mobility Efficiency in *E. faecalis* |
|---|---|---|---|
| | *L. lactis* x *L. lactis* | *L. lactis* x *E. faecalis* | |
| Wild type sex factor | $7.33 \times 10^{-3}$ [a] | N.D. | N/A |
| SFISp | $8.51 \times 10^{-4}$ | N.D. | N/A |
| pTRK28::SFISp | $3.50 \times 10^{-4}$ | $7.29 \times 10^{-9}$ | < 1% |
| pTRK28::SF | $5.33 \times 10^{-7}$ | N.D. | N/A |
| pTRK28::pRS01 | $5.70 \times 10^{-2}$ [b] | $6.94 \times 10^{-9}$ | From <1% to 86% |

[a] (Belhocine *et al.*, 2005)
[b] (Mills *et al.*, 1994)
N.D. not detected
N/A not applicable

# Figure 4.1


A
LI.LtrB intron
Sex
E1
ORF
I
Spc
E2
Factor
ltrB gene
B
SFISp containing cell
Donor cell
Recipient cell
Transconjugant cell → co-integrate donor cell
SFISp
pTRK28
pTRK28::SFISp
NZISp/pTRK28
(SpcR/ErmR)
NZISp/pTRK28::SFISp
(SpcR/ErmR)
LM0231
(FusR)
LM0231/pTRK28::SFISp
(FusR/SpcR/ErmR)
C
Donor cell
Recipient cell
Transconjugant cell
Transconjugant cell
pMNHSTet
pMNHSTet+
LM0231/pTRK28::SFISp
(SpcR/ErmR)
JH2-2/pMNHSTet
(TetR)
JH2-2/pTRK28::SFISp/ pMNHSTet
(SpcR/ErmR/TetR)
JH2-2/pTRK28::SFISp/ pMNHSTet
(SpcR/ErmR/TetR)
D
TC1
TC2
TC3
D
R

# Figure 4.2


A
Donor cell
Recipient cell
Transconjugant cell
Transconjugant cell
E1
ORF
E2
pTRK28::pRS01
Erm^R
OriC
X
pMNHS
pMNHS+
1
2
DM2036
(Erm^R)
JH2-2/pMNHS
(Spc^R)
JH2-2/pMNHS/pTRK28::pRS01
(Spc^R/Erm^R)
JH2-2/pMNHS/pTRK28::pRS01
(Spc^R/Erm^R)
B
TC1
TC2
TC3
D
R
C
MWM
1
2
3
4
5
6
7
8
R
MWM
M
R
86
3
4
7
8
4
<1
7
% LI.LtrB Mobility

# CHAPTER FIVE

## *Trans*-splicing of the Ll.LtrB group II intron in *Lactococcus lactis*

### 5.1 Preface

The objective of the present thesis is to address evolutionary questions concerning group II introns with experimental tools. After group II intron dissemination, we addressed the evolutionary theory proposing that nuclear introns and the five small nuclear RNAs (snRNAs) part of the spliceosome evolved from fragmentation of an ancestral group II intron (Sharp, 1985; Cech, 1986). This theory is supported by structural and functional resemblances between portions of group II introns and snRNAs, as well as similarity of nuclear and group II intron boundaries and splicing pathways (Cech, 1986; Jacquier, 1990). According to this theory, an ancestral group II intron was fragmented to give rise to the five snRNAs (Jacquier, 1990; Sharp, 1991). The discovery of fragmented introns that undergo *trans*-splicing provided substantial support to this theory (Sharp, 1991). However, *trans*-splicing introns were only found fragmented in domains I, III and IV, which does not fully rationalize fragmentation into "five easy pieces" (Sharp, *Science* 1991). We thus decided to explore the ability of group II introns to tolerate fragmentation at different sites and splice *in trans*, with

the eventual aim to fragment a group II intron consistently with the origin of the five snRNAs.

To achieve our goal, we constructed a model system for *trans*-splicing of group II introns based on Ll.LtrB. *Trans*-splicing has never been reported in bacteria and no fragmented variants of Ll.LtrB were reported in nature. Therefore, we engineered artificial fragmentation sites within Ll.LtrB to assess its *trans*-splicing potential. To validate our system, the fragmentation sites that we engineered within Ll.LtrB mimicked fragmentation of natural group II introns (domain I, III and multiple locations within domain IV). To fully assess the *trans*-splicing potential of Ll.LtrB, we built a sensitive splicing/conjugation assay that exploits the dependence of conjugative transfer of the sex factor on Ll.LtrB splicing and relaxase production.

*Published online 27 March 2007* *Nucleic Acids Research, 2007, Vol. 35, No. 7 2257–2268*
*doi:10.1093/nar/gkl1146*

# Trans-splicing of the Ll.LtrB group II intron in *Lactococcus lactis*

**Kamila Belhocine, Anthony B. Mak and Benoit Cousineau***

Department of Microbiology and Immunology, McGill University, Montréal, Québec, Canada H3A 2B4

In: ***Nucleic Acids Research*** (2007) **35**:2257-2268


## 5.2 Abstract

The Ll.LtrB intron from the Gram-positive bacterium *Lactococcus lactis* is one of the most studied bacterial group II introns. Ll.LtrB interrupts the relaxase gene of three *L. lactis* conjugative elements. The relaxase enzyme recognizes the origin of transfer (*oriT*) and initiates the intercellular transfer of its conjugative element. The splicing efficiency of Ll.LtrB from the relaxase transcript thus controls the conjugation level of its host element. Here, we used the level of sex factor conjugation as a read-out for Ll.LtrB splicing efficiency. Using this highly sensitive splicing/conjugation assay ($10^7$-fold detection range), we demonstrate that Ll.LtrB can *trans*-splice in *L. lactis* when fragmented at various positions such as: three different locations within domain IV, within domain I, and within domain III. We also demonstrate that the intron-encoded protein, LtrA, is absolutely required for Ll.LtrB *trans*-splicing. Characteristic Y-branched *trans*-spliced introns and ligated exons are detected by RT-PCR from total RNA extracts of cells harbouring fragmented Ll.LtrB. The splicing/conjugation assay we developed constitutes the first model system

to study group II intron *trans*-splicing *in vivo*. Although only previously observed in bacterial-derived organelles, we demonstrate that assembly and *trans*-splicing of a fragmented group II intron can take place efficiently in bacterial cells.

## 5.3 Introduction

Group II introns are large ribozymes that catalyze their own excision from RNA transcripts through a process called splicing (Michel and Ferat, 1995; Lehmann and Schmidt, 2003; Lambowitz and Zimmerly, 2004). Some group II introns that harbour an open reading frame (ORF) are also mobile retroelements, which are capable of invading new DNA sites using an RNA intermediate (Lambowitz and Zimmerly, 2004). Group II introns are found in bacteria, archaea, and bacterial-derived organelles of lower eukaryotes and higher plants (Michel and Ferat, 1995).

Splicing of group II introns is achieved by two consecutive trans-esterification reactions (Figure 5.1A) (Lehmann and Schmidt, 2003; Lambowitz and Zimmerly, 2004). The 2'-OH of a bulged adenosine (A) found near the 3' end of the intron initiates the first nucleophilic attack on the exon 1-intron junction (Figure 5.1A, step 1). This results in a 2'-5' branching of the intron on the bulged adenosine and the release of exon 1. The liberated 3'-OH of exon 1 initiates the second nucleophilic attack on the intron-exon 2 junction (Figure 5.1A, step 2), thereby releasing the intron in the form of a lariat and ligating the two exons. This splicing mechanism is identical to the removal of nuclear eukaryotic introns by

the spliceosome, therefore suggesting a common origin for these two classes of introns (Sharp, 1991; Michel and Ferat, 1995; Lambowitz and Zimmerly, 2004).

Group II introns share very little sequence similarity. However, their secondary structure is universally conserved and is composed of six domains radiating from a central hub (e.g. Figure 5.1B) (Michel and Ferat, 1995; Lehmann and Schmidt, 2003; Lambowitz and Zimmerly, 2004). Domain I is transcribed first, and when folded, it provides the scaffold to dock the other folded intron domains (Su et al., 2005). Domain IV plays little if any role in catalysis (Lehmann and Schmidt, 2003; Lambowitz and Zimmerly, 2004) and was shown to be dispensable for splicing *in vitro* (Jarrell et al., 1988). In many cases, the loop region of domain IV contains an ORF expressing an intron-encoded protein (IEP), which provides essential functions for both intron splicing and mobility (Mohr et al., 1993; Michel and Ferat, 1995; Lehmann and Schmidt, 2003; Lambowitz and Zimmerly, 2004). Domain V is the catalytic domain of these ribozymes and its sequence is the most conserved among group II introns (Michel and Ferat, 1995; Qin and Pyle, 1998; Lehmann and Schmidt, 2003; Lambowitz and Zimmerly, 2004). Domain VI carries the branch point nucleotide, which is most often an adenosine. The structure of this domain is believed to place the branch point nucleotide near the 5' splice site in a position that allows initiation of the first nucleophilic attack (Figure 5.1A, step 1) (Qin and Pyle, 1998).

In addition to local base pairing, a series of long-range interactions between the various folded domains were characterized (e.g., Figure 5.1B, Greek symbols) (Michel and Ferat, 1995; Qin and Pyle, 1998; Lehmann and Schmidt, 2003; Lambowitz and Zimmerly, 2004). For instance, recognition of the exon-

intron boundaries is achieved via long-range interactions between domain I and both exons. In general, domain I harbours the exon binding sites 1 and 2 (e.g., Figure 5.1B, EBS1 and EBS2), which are involved in 5' splice site recognition through base pairing interactions with exon 1. The sequences of these exon binding sites are complementary to contiguous intron binding sites 1 and 2 located at the 3' end of exon 1 (e.g., Figure 5.1B, IBS1 and IBS2) (Lehmann and Schmidt, 2003). Similarly, the intron-exon 2 splice junction is recognized by either the δ-δ' or EBS3-IBS3 interaction between domain I and exon 2 (Michel and Ferat, 1995; Lehmann and Schmidt, 2003). Another important long-range interaction, ζ-ζ', links the GNRA tetraloop of domain V to a phylogenetically conserved stem-loop structure within domain I (e.g., Figure 5.1B). This tertiary interaction is thought to form the intron catalytic core (Bonen, 1993; Qin and Pyle, 1998). It was previously demonstrated that tertiary interactions allow the assembly and splicing of an intron fragmented within domain IV *in vitro*, even if no base-pairing can take place between the two intron fragments (Jarrell et al., 1988).

Remarkably, the fragments of some group II introns that became fragmented by genome rearrangements in organelles have the capacity to re-assemble and splice accurately *in vivo* (Bonen, 1993; Michel and Ferat, 1995). This phenomenon, known as *trans*-splicing, ligates exons expressed on two separate transcripts. Most fragmented group II introns are found interrupted within domain IV (Bonen, 1993; Michel and Ferat, 1995). When these introns harbour an IEP within domain IV, the fragmentation site is located either

upstream or downstream from the ORF. There are two examples of group II introns fragmented within domain III; they are found in the chloroplastic genomes of the liverwort *Marchantia polymorpha* and the tobacco plant *Nicotiana tabacum* (Kohchi et al., 1988). In addition, two tripartite group II introns, i.e. introns fragmented into three pieces, have been reported to functionally *trans*-splice in their respective hosts. The first example ligates exons 1 and 2 of the *psaA* chloroplast gene of the green alga *Chlamydomonas reinhardtii* (Goldschmidt-Clermont et al., 1991). The second example joins exons c and d of the *nad5* subunit gene in mitochondria of the angiosperm plant *Oenothera berteriana* (Knoop et al., 1997). These tripartite introns are both fragmented in domain I upstream of the EBS regions and in domain IV, which, in both cases, do not appear to contain an ORF.

The discovery of group II introns in bacteria greatly facilitated the study of their splicing and mobility pathways. One of the best characterized bacterial group II introns is Ll.LtrB, a 2.5-kb intron found in the low G+C Gram-positive bacterium *Lactococcus lactis*. It was the first bacterial group II intron shown to splice (Mills et al., 1996) and mobilize (Mills et al., 1997) *in vivo*. The splicing (Zhou et al., 2000; Matsuura et al., 2001; Klein et al., 2004; Chen et al., 2005) and mobility pathways (Cousineau et al., 1998, 2000; Ichiyanagi et al., 2002; Smith et al., 2005) of this intron have been extensively studied *in vivo*.

Ll.LtrB harbours an ORF in domain IV coding for a 599 amino acid protein, LtrA, with three enzymatic activities: reverse transcriptase, maturase and endonuclease (Matsuura et al., 1997). The maturase activity of LtrA is essential to promote intron splicing by assisting the folding of this large ribozyme into its

catalytically active conformation (Matsuura et al., 1997; Cousineau et al., 1998). LtrA binds the intron RNA primarily at a distinctive structure found in domain IVa (Figure 5.1B, DIVa) (Matsuura et al., 2001). By binding to this structural motif, the IEP occludes its own Shine-Dalgarno sequence and autoregulates its translation (Singh et al., 2002).

The Ll.LtrB group II intron is found in three conjugative elements from *L. lactis*: the pRS01 (Mills et al., 1996) and pAH90 (O'Sullivan et al., 2001) plasmids, and the *L. lactis* 712 sex factor, an integrative and conjugative element closely related to pRS01 (Shearman et al., 1996). Interestingly, in all three cases, the intron interrupts the gene coding for a putative relaxase at the same position. The relaxase enzyme is essential for conjugation; it initiates DNA transfer by inflicting a site- and strand-specific nick at the origin of transfer (*oriT*) of the conjugative element. Then, the enzyme carries one DNA strand to the recipient cell by rolling-circle replication via the mating pore, and re-ligates the ends of the conjugated element (Byrd and Matson, 1997).

Since the relaxase enzyme is essential for conjugation, splicing of Ll.LtrB is a prerequisite for the intercellular transfer of its host elements (Mills et al., 1996; Shearman et al., 1996). Klein and colleagues showed that the intimate relationship between splicing and conjugation can be exploited to monitor splicing of Ll.LtrB. They confirmed by real-time RT-PCR analyses that the conjugation efficiency of a mobilizable plasmid is directly proportional to splicing of Ll.LtrB from the relaxase transcript (Klein et al., 2004).

In this study, we use the Ll.LtrB intron from *L. lactis* as a model system to study group II intron *trans*-splicing in bacteria. We developed a splicing/

conjugation assay that measures Ll.LtrB splicing efficiency by monitoring sex factor conjugation levels. We used this splicing/conjugation assay to investigate *trans*-splicing of the lactococcal intron. When fragmented at analogous locations to various naturally occurring fragmentation sites, Ll.LtrB consistently shows the ability to splice *in trans*, sometimes at very high efficiencies. Furthermore, we demonstrate that the LtrA protein is essential for *trans*-splicing. These findings constitute the first demonstration and characterization of *trans*-splicing of a group II intron in bacteria.

## 5.4 Experimental procedures

### 5.4.1 Strains and plasmids

*L. lactis* strains NZ9800Δ*ltrB*::*tet* (Ichiyanagi et al., 2002) and LM0231 (Mills et al., 1994) were grown at 30°C without shaking in M17 media supplemented with 0.5% glucose (GM17). *E. coli* strain DH10β was used for cloning and was grown at 37°C with shaking in LB broth. Antibiotics were added at the following concentrations: 3 μg/ml for tetracycline (Tet), 300 μg/ml for spectinomycin (Spc), 25 μg/ml for fusidic acid (Fus) and 10 μg/ml for chloramphenicol (Cam).

Two 43 bp-long $P_{23}$ constitutive promoters ($P_{23}$ right and $P_{23}$ left) (van der Vossen et al., 1987) were obtained by annealing two complementary oligonucleotides (SuppTable 5.1, $P_{23}$ right top/bottom, $P_{23}$ left top/bottom). The pDL-$P_{23}^2$ plasmid was constructed by cloning, in opposite directions, the two $P_{23}$ constitutive promoters within the pDL278 shuttle plasmid (LeBlanc et al., 1992).

The $P_{23}$ left promoter is oriented towards the *rrnB* terminator (PvuI) while the $P_{23}$ right promoter is directed towards the λ terminator (BamHI). The two promoters are 135 bp apart and initiate transcription in opposite orientations (Figure 5.2A, Figure 5.3B). When the oligonucleotides for $P_{23}$ left were designed, a BssHII cloning site was introduced 6 nucleotides downstream of the transcription initiation site (SuppTable 5.1, $P_{23}$ left top/bottom). Similarly, a NotI cloning site was introduced at the +6 position of $P_{23}$ right (SuppTable 5.1, $P_{23}$ right top/bottom). These two unique restriction sites were subsequently used to clone the various relaxase fragments.

The pDL-$P_{23}^2$-*ltrB* plasmid was obtained by cloning the promoter-less *ltrB* gene in pDL-$P_{23}^2$ under the control of $P_{23}$ right. The *ltrB* gene was PCR-amplified from genomic DNA extracted from the sex factor-containing *L. lactis* strain NZ9800 (Belhocine et al., 2004) using the 5'*ltrB* and 3'*ltrB* primers (SuppTable 5.1). To generate the intron-mutated versions of this plasmid, the BsrGI/HpaI fragment spanning most of the Ll.LtrB intron and a portion of the *ltrB* exon 2 was exchanged for the corresponding fragment from pLE12ΔD5, pLE12ΔORF and pLE12Mat$^-$ respectively (Matsuura et al., 1997; Belhocine et al., 2004).

The pLE-$P_{23}^2$ plasmid was generated by cloning the 1.8 kb SapI fragment from pDL-$P_{23}^2$ containing the divergent $P_{23}$ transcriptional units, into the unique SmaI site of the pLE1 plasmid (Mills et al., 1994). The resulting plasmid was used for expression of two of the three intron pieces when assessing tripartite variants of Ll.LtrB, and for overexpression of LtrA in ORF-complementation assays. For

the latter experiments, the *ltrA* gene was PCR-amplified using primers 5'*ltrA* and 3'*ltrA* (SuppTable 5.1), and cloned in the NotI restriction site of pLE-$P_{23}$[2].

To generate fragmented variants of Ll.LtrB, the sites of fragmentation, Sn (Figure 5.1B, S1 to S5), were chosen according to the occurrence of fragmented group II introns in nature. Two primers were designed for each Sn fragmentation site. The Sn-P1 primer was always used in conjunction with the 5'*ltrB* primer to generate the 5' fragment of the relaxase by PCR amplification (Figure 5.3A). Similarly, the Sn-P2 primer was always used with the 3'*ltrB* primer to generate the 3' fragment of the relaxase (Figure 5.3A). The 5' and 3' fragments were cloned downstream from the $P_{23}$ right and $P_{23}$ left promoters respectively. All relaxase fragments were PCR-amplified from genomic DNA of the sex factor-containing *L. lactis* strain NZ9800 (Belhocine et al., 2004). To generate the maturase-mutated and ΔORF variants of fragmented Ll.LtrB constructs, the ORF-containing intron fragment was PCR-amplified from the pLE12Mat$^-$ and pLE12ΔORF plasmids respectively. The sequence integrity of all relaxase fragments generated by PCR amplification was confirmed by sequence analyses.

### 5.4.2 Conjugation assays

Mating was performed on milk plates made of 5% non-fat dried milk (Carnation$^{TM}$), 1% glucose and 1.5% agar (Belhocine et al., 2004). Donor and recipient *L. lactis* strains were inoculated from saturated overnight cultures (0.4 ml into 10 ml), and grown for 7 hrs. Cells were then collected by centrifugation, mixed and spread on milk plates. After a 16-hour incubation at 30°C, cells were

recovered in 1 ml 1X PBS, and appropriate dilutions were plated on selective medium for donors, recipients and transconjugants. Conjugation efficiencies were calculated as the ratio of transconjugant cells ($Fus^R/Tet^R$) to donor cells ($Spc^R$), for three independent assays. The *L. lactis* strains NZ9800Δ*ltrB*::*tet* and LM0231 ($Fus^R$) were used as the donor and recipient strains, respectively.

### 5.4.3 RNA isolation, Northern blot hybridizations and RT-PCR

Total RNA was isolated from *L. lactis* cells inoculated from saturated overnight cultures (0.4 ml into 10 ml), and grown for 7 hrs. Cell pellets were mixed with 500 ml TRIzol (Invitrogen life technologies) and 250 µg of acid-washed glass beads (Sigma). The mixture was vortexed for 3 min and incubated at 55°C for 5 min; this treatment was repeated three times. The rest of the extraction was performed according to the TRIzol manufacturer's protocol.

For Northern blot hybridization, 10 µg of total RNA was resolved on 1% agarose gel containing 5% formaldehyde and transferred by capillarity to nylon membrane (Hybond-N; Amersham Biosciences). The membranes were hybridized with the appropriate 5'-$^{32}$P labelled oligonucleotide probe (see figure legends and SuppTable 5.1). Radiolabelling of the probes was performed in a final volume of 10 µl containing 10 pmol of the oligonucleotide, 6.4 pmol of [γ-$^{32}$P]ATP (6,000 Ci/mmol; Amersham Biosciences) and 5 U of T4 polynucleotide kinase (New England Biolabs) at 37°C for 1 h. The labelling mix was then purified on Sephadex G-50 columns. The membranes were hybridized for 2 hrs at 42°C, washed and exposed on a phosphor screen, which was visualized with the

Molecular Imager FX software from Bio-Rad. As an RNA loading control, the membranes were stripped and re-probed with an oligonucleotide specific for the 23S rRNA of *L. lactis* (Figure 5.2B).

For reverse transcriptase (RT)-PCR analyses, cDNAs were synthesized using the SuperScript II cDNA synthesis kit from Invitrogen according to the manufacturer's instructions, using 5 μg of total RNA and 2 pmol of the appropriate primer. The BP cDNA primer was used to synthesize cDNA across the Ll.LtrB branch point, while the E2 cDNA primer was used to synthesize cDNA of the *ltrB* ligated exons (SuppTable 5.1). PCR amplification was performed using 10% of the cDNA obtained and 2 U of Taq DNA polymerase (New England Biolabs).

### 5.4.4 Protein isolation and Western blot

Total proteins were extracted from *L. lactis* cells grown to mid-log phase. Cells were pelleted and resuspended in 100 μl TES buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 25% sucrose) containing 1 mg/ml lysozyme (Roche), 50 U/ml mutanolysin (Sigma®) and 0.1 mg/ml RNaseA (USB). This mixture was incubated for 1 hr at 37°C. Cells were then lysed by mixing with 100 μl 1X TE buffer containing 4% SDS and boiling for 5 min. Samples were loaded on SDS-PAGE (10%) by mixing 10 μl of the total protein extract with 10 μl of SDS loading buffer containing 5% β-mercaptoethanol.

For Western blot analysis, the gel was transferred electrophoretically onto PVDF membrane (Immobilon-P, Millipore). The membrane was blocked O/N in

1X PBS - 0.1% Tween 20 (Fisher) (PBS-T) containing 10% non-fat dried milk (Carnation™). Anti-LtrB antibodies (Chen et al., 2005) were subjected to a membrane containing protein extracts from *L. lactis* cells not expressing LtrB to reduce the amounts of antibodies reactive to non-LtrB proteins, like previously described (Zhou et al., 2000). Following blocking, the membrane containing the samples for analysis was incubated 1 hr in PBS-T + 5% milk containing 1:5000 depleted anti-LtrB antibodies. After three 5 min washes in PBS-T, the membrane was incubated 1 hr in PBS-T containing horseradish peroxidase (HRP)-conjugated anti-rabbit antibody (1:30,000; Bio-Rad), then washed four times for 5 min in PBS-T. The membrane was detected with Immobilon™ Western Chemi-luminescent HRP Substrate (Millipore) and visualized with the Molecular Imager VersaDoc from Bio-Rad and on X-ray film. The software Quantity One® (Bio-Rad) was used for quantitative analysis of LtrB amounts.

## 5.5 Results

### 5.5.1 A conjugation-based system to monitor Ll.LtrB splicing

We developed a genetic assay to monitor Ll.LtrB splicing by measuring the conjugation efficiency of the chromosomal sex factor between *L. lactis* strains (Figure 5.2A). This splicing/conjugation assay exploits the dependence of conjugative transfer of the sex factor on Ll.LtrB splicing from the relaxase transcript (Shearman et al., 1996; Klein et al., 2004). The sex factor donor strain used in this assay is NZ9800Δ*ltrB*::*tet*, in which the sex factor carries a deletion of the Ll.LtrB intron and portions of its flanking exons destroying the relaxase gene

(Ichiyanaga et al., 2002). In the absence of relaxase, the sex factor transfer efficiency drops from 7.33 x $10^{-3}$ (Belhocine et al., 2005) to 1.88 x $10^{-9}$ (Table 5.1, cf. SF wild-type and SFΔ*ltrB* / pDL-$P_{23}^2$).

A complementation system was built where the intron-interrupted *ltrB* gene was cloned in the pDL-$P_{23}^2$ plasmid and expressed under the control of the $P_{23}$ constitutive promoter (Figure 5.2A, pDL-$P_{23}^2$-*ltrB*). When the relaxase-deleted sex factor is thereby complemented, transfer efficiency is restored from 1.88 x $10^{-9}$ to 5.80 x $10^{-2}$ (Table 5.1, SFΔ*ltrB* / pDL-$P_{23}^2$-*ltrB*). The difference in transfer efficiency between the complemented (5.80 x $10^{-2}$) and the wild-type sex factor (7.33 x $10^{-3}$) (~ 8-fold) is likely due to overexpression of the relaxase gene from a constitutive promoter on a multiple copy plasmid. Notably, the relaxase transcript produced from the wild-type chromosomal sex factor is undetectable by Northern blot. On the other hand, expression from the $P_{23}$ promoter produces high levels of both the full-length relaxase transcript and mature mRNA (Figure 5.2B, compare lanes NZ9800 and *ltrB*-WT in both blots). Accordingly, significantly less ligated exons and no intron lariats are detected by RT-PCR in NZ9800, where the chromosomal sex factor expresses a functional relaxase, compared to expression of the relaxase gene from the $P_{23}$ promoter (see below, Fig 5.3B and C, cf. NZ9800 and *ltrB*-WT). Western blot analysis confirms the significant increase in LtrB production from the plasmid when compared to its expression from the chromosomal sex factor (Figure 5.2C, cf. NZ9800 and *ltrB*-WT). Since no signal for LtrB is detected from NZ9800, the increase in LtrB expression from the $P_{23}$ promoter was estimated to be of at least 112-fold. Therefore, the slight increase in

conjugation efficiency between the wild-type and complemented sex factor (~ 8-fold) is not comparable to the great difference in levels of relaxase expression observed (at least 112-fold). The complementation system is thus most likely saturated in relaxase enzyme, and the maximal rate of sex factor conjugation efficiency is probably reached ($10^{-2}$ range). This suggests that the conjugation efficiency observed for the complemented sex factor underestimates the real splicing efficiency of Ll.LtrB. Therefore, the conjugation efficiencies observed throughout this study will be interpreted as increases from the background level rather than decreases from the saturated maximum level.

To verify the dependence of sex factor conjugation on Ll.LtrB splicing, we tested an Ll.LtrB variant where the catalytic domain, i.e. domain V, was deleted (pDL-$P_{23}{}^2$-*ltrB*-ΔD5) (Matsuura et al., 1997). Complementation with a relaxase gene harbouring the catalytically inactive intron failed to sustain sex factor transfer; the conjugation efficiency observed was only 3.79 x $10^{-9}$ (Table 5.1, SFΔ*ltrB* / pDL-$P_{23}{}^2$-*ltrB*-ΔD5). This conjugation efficiency is similar to the background level observed for the *ltrB*-deficient sex factor (Table 5.1, SFΔ*ltrB* / pDL-$P_{23}{}^2$). Northern blot analyses of total RNA extracted from donor cells harbouring the pDL-$P_{23}{}^2$-*ltrB*-ΔD5 plasmid showed accumulation of the pre-mRNA precursor while no traces of ligated exons was detected (Figure 5.2B, lane *ltrB*-ΔD5 in both blots). Accordingly, neither ligated exons nor excised intron lariats were detected by RT-PCR analyses (see below, Figure 5.4C and 3B respectively). Taken together, these results demonstrate that transfer of the *L. lactis* sex factor is dependent on Ll.LtrB splicing from the relaxase transcript. We have

therefore built a highly sensitive splicing/conjugation assay ($10^7$-fold detection range) to monitor Ll.LtrB splicing in *L. lactis*.

### 5.5.2 *Trans*-splicing of Ll.LtrB fragmented within domain IV

We used the above splicing/conjugation assay to assess if Ll.LtrB is capable of splicing when severed into two fragments at various positions. Several locations were chosen to mimic fragmentation sites found in naturally occurring *trans*-splicing group II introns (Figure 5.1B, S1 to S5). To introduce a fragmentation site within Ll.LtrB (Sn, n varies from 1 to 5), the complete relaxase gene was PCR-amplified as two non-overlapping fragments (Figure 5.3A). To ensure equimolar production of each relaxase fragment, these were expressed from two constitutive $P_{23}$ promoters located on the same plasmid, pDL- $P_{23}^2$ (Figure 5.3B). The two promoters were cloned in opposite orientation to ensure that two independent RNA transcripts are generated. Proper expression of relaxase fragments was assessed by Northern blot analyses and all the transcripts displayed the expected size. However, the fragments expressed from the $P_{23}$-right promoter, consisting of exon 1 and the first portion of the Ll.LtrB intron, invariably showed slightly higher expression level than those expressed from the $P_{23}$-left promoter, corresponding to the second portion of Ll.LtrB and exon 2 (data not shown). The lower expression of the 3' relaxase fragments is likely due to their instability since transcription of these fragments is initiated at different positions within Ll.LtrB. This probably leads to a slight underestimation of the Ll.LtrB *trans*-splicing efficiency.

Most *trans*-splicing group II introns are found fragmented within domain IV. When these introns harbour an ORF, the fragmentation site can be found either upstream or downstream of the ORF. Therefore, we investigated the effect of fragmenting Ll.LtrB in domain IV, either upstream or downstream of the *ltrA* ORF. Two sites were chosen for fragmentation upstream of the ORF. The first site lies at the bottom of the DIV stem (position 524, Figure 5.1B, S1). Fragmentation of Ll.LtrB at this location does not allow base-pairing interactions between the two intron RNA pieces, but leaves only a series of long-range tertiary interactions for intron fragment recognition and assembly (Figure 5.1B, Greek letters). The second site lies within the DIVa stem at position 543 (Figure 5.1B, S2). According to the potential tertiary contacts between the two Ll.LtrB fragments, this fragmentation site is equivalent to S1. However, S2 also allows a potential 16 base-pair interaction between the two intron pieces. When the intron is fragmented at the S1 site, the relaxase can promote sex factor conjugation to an efficiency of $1.66 \times 10^{-6}$, ~ 890-fold over the background (Figure 5.5, S1). This shows that tertiary interactions are sufficient to promote intron fragment assembly and splicing. On the other hand, when the relaxase is fragmented at the S2 site, sex factor conjugation is ~ 87 times more efficient than S1, ~ 76,000-fold over the background ($1.44 \times 10^{-4}$) (Figure 5.5, S2). These results suggest that base-pairing interactions between intron fragments further stimulate their assembly and splicing *in vivo*.

Fragmentation of Ll.LtrB in domain IV downstream of the ORF was also tested. The intron was fragmented downstream from the *ltrA* stop codon at position 2375, which allows a potential 15 base-pair interaction between the two

intron fragments (Figure 5.1B, S3). Even though the potential tertiary interactions between the two intron pieces are the same as with S1 and S2, this Ll.LtrB *trans*-splicing variant harbours *ltrA* and the high affinity LtrA binding site (Figure 5.1B, DIVa) on the 5' instead of the 3' intron fragment. The relaxase gene fragmented at the S3 site promoted sex factor conjugation to very high levels ($1.16 \times 10^{-2}$) (Figure 5.5, S3, two fragments), ~ 6.17 million-fold over the background (Table 5.1, SF$\Delta$*ltrB* / pDL-$P_{23}^2$). Moreover, when the two relaxase pieces were expressed independently, sex factor conjugation remained near basal levels ($10^{-8}$), showing a slight increase compared to the background (Figure 5.5, S3, 5' fragment and 3' fragment). Therefore, expression of both relaxase fragments is absolutely required to promote sex factor conjugation. Taken together, these results demonstrate that when the Ll.LtrB intron is fragmented independently at three different positions within domain IV, a significant proportion of the two intron fragments can fold correctly, assemble, and *trans*-splice to ligate their flanking exons.

### 5.5.3 *Trans*-splicing of Ll.LtrB fragmented at other naturally occurring group II intron fragmentation sites

Fragmentation sites outside domain IV can be found in domains I or III in some group II introns. The next Ll.LtrB *trans*-splicing variant that we analyzed mimics the fragmentation site in domain I of the tripartite intron from the *psaA* chloroplast gene of the green algae *C. reinhardtii* (Goldschmidt-Clermont et al., 1991). The fragmentation site was introduced at position 172 upstream of the EBS1 and 2 sequences. When Ll.LtrB is fragmented at this position, the possible

tertiary interactions contributing to intron fragment assembly are different from those of the previous fragmentation sites (S1 to S3). Moreover, a potential 19 base-pair interaction can take place between the two fragments of the intron (Figure 5.1B, S4). When the two relaxase fragments are expressed independently, no significant increase in conjugation efficiency is observed ($10^{-8}$ range) (Figure 5.5, S4, 5' fragment and 3' fragment). However, when both fragments are co-expressed, the rate of sex factor conjugation increases to $4.52 \times 10^{-4}$ (Figure 5.5, S4, two fragments). The substantial increase observed (~ 240,000-fold) in sex factor transfer efficiency from basal level when both relaxase fragments are produced demonstrates that Ll.LtrB can also *trans*-splice when fragmented in domain I at this position.

The last fragmentation site engineered in Ll.LtrB lies in the loop region of domain III at position 499 (Figure 5.1B, S5) and mimics the fragmentation site of the *rps12* intron in *M. polymorpha* and *N. tabacum* chloroplasts (Kohchi et al., 1988). Fragmentation of Ll.LtrB at this location allows the same tertiary contacts between the two intron fragments as the S1, S2 and S3 sites. Moreover, 17 base pairs could be formed between the two intron fragments. When the *ltrB* gene is fragmented at this position, the amount of relaxase produced supports sex factor transfer to a level of $6.03 \times 10^{-5}$ (Figure 5.5, S5), which represents an increase of approximately 32,000-fold over the background level.

These results demonstrate that when the intron is fragmented within domains I or III, a significant proportion of the two independently expressed Ll.LtrB fragments can fold correctly, assemble, splice *in trans*, and ligate their flanking exons.

### 5.5.4 Detection of *trans*-spliced intron molecules and ligated exons by RT-PCR

Excised group II intron lariats can be detected from total RNA extracts by RT-PCR across the branch point (Figure 5.4A, top, black arrows) (Li-Pook-Than and Bonen, 2006). Using a similar approach, we designed an RT-PCR assay to detect the *trans*-spliced Y-shaped intron molecules (Figure 5.4A, bottom). This assay was designed to yield a 201-bp fragment when a lariat or a Y-shaped intron molecule is produced. The two RT-PCR primers lie outside of the region where the fragmentation sites were engineered, and could thus be used with all fragmented relaxase variants (Figure 5.4A, bottom, black arrows). The assay was performed on total RNA extracted from donor cells (NZ9800Δ*ltrB*) containing various Ll.LtrB constructs. As shown in figure 5.3B, donor cells harbouring the full length relaxase gene yield a unique RT-PCR product of the expected size, representing the Ll.LtrB lariat (Figure 5.4B, *ltrB*-WT). On the other hand, Y-branched intron molecules are only detected with the most efficient *trans*-splicing construct, where the intron is fragmented downstream of the ORF (Figure 5.5, S3, two fragments). Failure to detect Y-branched introns from the other fragmented variants of Ll.LtrB is most likely due to the low level of these molecules as well as their instability. Indeed, Y-shaped molecules have three free ends and are likely more sensitive to degradation than intron lariats (Figure 5.4A, bottom).

The presence of ligated exons of the relaxase gene was also assessed by RT-PCR. This assay was designed to produce a 521-bp fragment when the exons

are properly ligated (Figure 5.4A, open arrows). As can be seen in the agarose gel presented in figure 5.3C, ligated exons are detected in total RNA extracted from all donor cells containing Ll.LtrB constructs that showed conjugation efficiencies above background. Specifically, ligated exons are detected in the case of the full-length *ltrB* gene but not from the catalytically inactive *ltrB*-ΔD5 variant (Figure 5.4C, *ltrB* WT and *ltrB*-ΔD5). Ligated exons are also detected in all five cases of Ll.LtrB *trans*-splicing (Figure 5.4C, S1 to S5). Sequence analyses of all RT-PCR products confirmed that the exons were ligated precisely at the intron splice junctions. These RT-PCR assays confirm that in all cases where sex factor conjugation increased over the background, properly ligated relaxase exons were readily detectable in donor cells. Furthermore, Y-branched intron molecules could be detected only for the fragmented relaxase construct that sustains the highest conjugation efficiency (S3).

### 5.5.5 The LtrA protein is essential for *trans*-splicing of Ll.LtrB

Splicing of group II introns *in vivo* requires the assistance of splicing co-factors that function as maturases to stabilize the catalytically active intron structure (Lambowitz and Zimmerly, 2004). Some group II introns, like Ll.LtrB, encode their own maturase. The role of IEPs in *trans*-splicing has never been investigated.

To assess the role of the LtrA protein in Ll.LtrB *trans*-splicing, we studied two different protein mutants. Firstly, the Ll.LtrB-Mat$^-$ mutant carries point mutations that only affect the maturase activity of LtrA (Matsuura et al., 1997).

Secondly, the Ll.LtrB-ΔORF intron carries a deletion of 88% of the *ltrA* coding sequence, which completely removes the maturase domain (Matsuura et al., 1997).

Cis-splicing of these Ll.LtrB mutants was first analysed in our splicing/conjugation assay. As expected, the Ll.LtrB-ΔORF mutant appears completely inactive. When the relaxase gene harbours this intron mutant, sex factor conjugation efficiency is only 10-fold higher than the background (Table 5.1, cf. SFΔ*ltrB* / pDL-$P_{23}^{2}$-*ltrB* ΔORF and SFΔ*ltrB* / pDL-$P_{23}^{2}$). Surprisingly, the maturase mutant efficiently induced sex factor transfer to a level ~ 152,000-fold over the background (Table 5.1, cf. SFΔ*ltrB* / pDL-$P_{23}^{2}$-*ltrB* Mat$^{-}$ and SFΔ*ltrB* / pDL-$P_{23}^{2}$). Such a result suggests that this LtrA mutant can still bind Ll.LtrB and promote splicing.

Next, the two LtrA mutants were analyzed in the context of two fragmented introns: the S4 variant fragmented within domain I, and the S3 variant fragmented downstream of the ORF within domain IV. These two introns were chosen because they *trans*-splice efficiently and are fragmented in different domains (DI vs. DIV). The intron fragment that is bound by the LtrA protein also differs in these two variants. Indeed, when the intron is fragmented within domain I (S4), the LtrA primary binding site (stem DIVa, Figure 5.1B) lies in the 3' fragment of the intron. On the other hand, when the fragmentation site is located downstream of the ORF (S3), the LtrA primary binding site lies in the 5' fragment of the intron (Figure 5.1B). As shown in Figure 5.5, the introduction of point mutations, that only abolish the maturase activity of LtrA to both fragmented

introns, caused a consistent decrease of ~ 35,000-fold in sex factor conjugation efficiency. (Figure 5.5, S4 and S3, two fragments/Mat$^-$). Moreover, removal of the ORF from Ll.LtrB completely abolishes *trans*-splicing as the sex factor conjugation efficiency drops to background levels with both variants tested ($10^{-9}$) (Figure 5.5, S3 and S4, two fragments/ΔORF). Interestingly, if a second plasmid overexpressing LtrA (pLE-$P_{23}^2$-*ltrA*) is introduced into cells harbouring the ΔORF variants of fragmented Ll.LtrB, conjugation efficiency is almost completely restored (Figure 5.5, S3 and S4, two fragments/ΔORF + ORF). This increase in sex factor conjugation efficiency is not a direct effect of LtrA on the sex factor since the LtrA-expressing plasmid alone does not sustain conjugation of the sex factor ($1.31 \pm 0.85 \times 10^{-9}$). These results demonstrate that the LtrA protein, primarily through its maturase function, is an essential co-factor for Ll.LtrB *trans*-splicing in *L. lactis*.

## 5.6 Discussion

In this study, we describe a highly sensitive conjugation-based system to monitor splicing of the Ll.LtrB group II intron in *L. lactis*. One of the important features of this experimental system lies in its extremely broad detection range. This simple genetic assay allows the detection and study of Ll.LtrB splicing quantitatively on a $10^7$-fold detection range. As we demonstrate in this study, the Ll.LtrB splicing/conjugation assay will be particularly useful to study group II intron *trans*-splicing, an aspect of group II intron splicing that is not very well studied or understood.

Although this genetic assay was developed to study Ll.LtrB *trans*-splicing, some interesting facts regarding sex factor conjugation and Ll.LtrB cis-splicing were discovered. First, we observe that expressing higher amounts of the relaxase enzyme increased sex factor conjugation efficiency from $7.33 \times 10^{-3}$ (Belhocine et al., 2005) to $5.80 \times 10^{-2}$ (Table 5.1). Our data show that the sex factor conjugation machinery can sustain even higher transfer rates than what is observed with the wild-type element. This demonstrates that the presence of Ll.LtrB within the relaxase gene not only controls conjugation of the *L. lactis* sex factor, but also that the normal LtrB expression level in NZ9800 is sub-optimal for the sex factor conjugation machinery. This suggests that any mutation within Ll.LtrB that hinders splicing would also directly, and proportionally, reduce LtrB expression and sex factor conjugation efficiency.

The catalytic domain (DV) of Ll.LtrB was previously deleted to generate a splicing deficient mutant (Matsuura et al., 1997). When the relaxase gene contains this Ll.LtrB-ΔD5 variant, the sex factor conjugation efficiency is similar to the

background observed when no relaxase is provided *in trans* (Table 5.1, 3.79 x $10^{-9}$ vs. 1.88 x $10^{-9}$). Moreover, no RT-PCR signals for excised intron lariats or ligated exons were obtained for this Ll.LtrB mutant (Figure 5.4B and C). Taken together, these results confirm that deletion of domain V from Ll.LtrB prevents any residual splicing in *L. lactis*.

When the relaxase gene provided *in trans* is interrupted by the Ll.LtrB maturase-deficient mutant (Ll.LtrB-Mat$^-$), sex factor conjugation efficiency increases by ~ 152,000-fold over the background (Table 5.1, 2.86 x $10^{-4}$ vs. 1.88 x $10^{-9}$). This shows that this LtrA maturase mutant can bind the intron and promote splicing to significant levels. When Ll.LtrB is missing most of the LtrA coding region (Ll.LtrB-ΔORF), we observe only a minor increase in sex factor conjugation efficiency (Table 5.1). In fact, this increase is comparable to the background conjugation rate obtained upon expression of individual fragment of the relaxase variants fragmented at position S3 and S4 (Figure 5.5, S3 / 5' fragment and 3' fragment, S4 / 5' fragment and 3' fragment). Conjugation efficiencies in the $10^{-8}$ range correspond to only a few transconjugant colonies per assay; very close to the limit of detection of our system.

Using our splicing/conjugation assay, we report the first experimental evidence that group II introns can *trans*-splice in bacteria. All the naturally occurring group II intron fragmentation sites were independently engineered within Ll.LtrB: three sites located in domain IV, two upstream and one downstream of the ORF, and one site located in both domains I and III. The *trans*-splicing efficiency of Ll.LtrB was found to vary greatly with the location of the

fragmentation site (Figure 5.5). Considering the limited number of characterized *trans*-splicing introns and the lack of similarity between these introns and Ll.LtrB, it was surprising to detect *trans*-splicing with all fragmented variants of the lactococcal intron. The fragmentation site we engineered in domain I was actually never found as a unique fragmentation site. It was only observed in nature in conjunction with a second fragmentation within domain IV in two tripartite group II introns (Goldschmidt-Clermont et al., 1991, Knoop et al., 1997). Nevertheless, when fragmented at this location within domain I, Ll.LtrB *trans*-splices efficiently since the sex factor conjugation is ~ 240,000-fold higher than the background.

In domain IV, two different intron fragmentation sites upstream of the *ltrA* coding region were studied. When the intron is fragmented at position S1 (Figure 5.1B), the *trans*-splicing efficiency is the lowest observed ($1.66 \times 10^{-6}$; Figure 5.5, S1). Due to the position of the fragmentation site, at the bottom of the DIV stem, the two intron fragments cannot interact by base-pairing but only through long-range tertiary interactions. However, when the intron is fragmented at a position only 19 nt downstream (Figure 5.1B, S2), the two intron fragments potentially have an additional 16 nt base-pair interaction. In this case, sex factor conjugation increases by ~ 100-fold ($1.44 \times 10^{-4}$; Figure 5.5, S2). This finding suggests that base-pairing between intron fragments plays an important role in intron assembly and splicing *in vivo*, contrary to what was previously proposed from *in vitro* data (Jarrell et al., 1988). Notably, the intron fragments of the vast majority of fragmented group II introns reported kept the potential to interact by a series of base pairs. Even though long range tertiary interactions are expected to be key in the assembly of folded group II intron fragments, it is conceivable that base-

pairing plays also a role in the assembly of fragmented group II introns at physiological conditions (Michel and Ferat, 1995).

The *trans*-splicing efficiency of Ll.LtrB is greater when the fragmentation in domain IV occurs downstream rather than upstream from the ORF (Figure 5.5, cf. S3/two fragments to both S1 and S2). A reason for this difference may be that interrupting Ll.LtrB upstream of its ORF can disrupt the DIVa stem (Figure 5.1B), which is the primary binding site for LtrA (Singh et al., 2002; Watanabe and Lambowitz, 2004). Since we showed that the LtrA protein is absolutely necessary for the *trans*-splicing reaction, it is likely that disruption of its primary interaction site within the intron RNA will decrease *trans*-splicing efficiency. Therefore, it would be difficult to generalize that group II introns *trans*-splice more efficiently when fragmented downstream rather than upstream from their carried ORFs within domain IV. Interestingly, a thorough comparative analysis by Qiu and Palmer showed that a mitochondrial group II intron in angiosperms, *nad1i*728, underwent fifteen independent incidences of fragmentation within domain IV throughout differentiation of these species (Qiu and Palmer, 2004). Among these cases, ten independent fragmentation events occurred upstream of the intron-encoded *matR* gene while only five occurred downstream. However, the authors did not conclude that fragmentation of a group II intron upstream rather than downstream of its carried ORF enabled better *trans*-splicing.

Since Ll.LtrB showed satisfactory *trans*-splicing efficiencies when fragmented in domain I and IV independently, we assessed the *trans*-splicing potential of two tripartite Ll.LtrB introns. We combined one fragmentation within domain IV, upstream or downstream of the ORF, with the fragmentation in

domain I (Figure 5.1B, S2&S4, S3&S4). These two tripartite variants of Ll.LtrB failed to *trans*-splice since the sex factor conjugation efficiency observed was in the $10^{-8}$ range, which is comparable to overexpression of independent fragments of the relaxase fragmented at sites S3 and S4 (Figure 5.5). Notably, in chlamydomonas chloroplast, it was shown that at least 9 cellular factors are needed for *trans*-splicing of the *psaA* tripartite intron (Perron et al., 2004). One may imagine that following the fragmentation of a group II intron, the recruitment of splicing factors becomes important in maintaining reasonable splicing efficiency and gene expression.

Using the intron variants fragmented in domain I (S4) and domain IV downstream of the ORF (S3), we show that the LtrA protein is essential for Ll.LtrB *trans*-splicing *in vivo* (Figure 5.5). Indeed, *trans*-splicing of these introns is abolished in the absence of LtrA, and is almost completely restored when LtrA is produced *in trans* from a second plasmid. Interestingly, LtrA remains important for the assembly of Ll.LtrB fragments whether its primary binding site (DIVa) rests in the 5' or the 3' intron fragment. These results suggest that for both of these Ll.LtrB fragmented variants, as well as for the ORF-encoding fragmented introns (S1 to S5), the LtrA binding structure is correctly folded in a significant proportion of the intron fragments to allow LtrA binding and Ll.LtrB splicing.

It is well established that the maturase activity of the LtrA protein is essential for Ll.LtrB splicing *in vivo* (Matsuura et al., 1997; Cousineau et al., 1998). In this study, we demonstrate that the maturase function of LtrA is also extremely important for Ll.LtrB *trans*-splicing. Using a maturase mutant for the introns fragmented at sites S3 and S4, we observe a significant and consistent

drop of ~ 35,000-fold in sex factor conjugation efficiency for both constructs (Figure 5.5, c.f. S3/two fragments and S4/two fragments and their respective Mat$^-$ counterparts). In our system, it is impossible to use conjugation efficiencies to quantify the difference observed between the cis-splicing full length intron and its maturase-deficient counterpart since conjugation is saturated when the full-length *ltrB* gene is overexpressed. Yet, one may expect that the assistance of the maturase function of LtrA is more important for bimolecular splicing reactions than for cis-splicing.

We demonstrated that a bacterial group II intron can *trans*-splice efficiently *in vivo* when fragmented at different positions. Although several splicing-competent fragmented introns were identified in organellar genomes, no *trans*-splicing group II intron was yet discovered in bacteria. Fragmentation of group II introns through genome rearrangements may occur more readily in organelles than in bacteria. Indeed, numerous *trans*-splicing group II introns were found in the mitochondrial genomes of higher plants, which are more plastic and subjected to high rearrangement rates (Fauron et al., 1997; Qiu and Palmer, 2004). Nevertheless, this study suggests that *trans*-splicing group II introns may exist and eventually be uncovered in prokaryotes. Group II introns can splice *in trans* in bacterial cells and therefore could be maintained if ever generated.

## 5.7 Acknowledgements

We thank Karen K. Yam and Greg T. Marczynski for providing comments on the manuscript, and Victoria Mandilaras for technical assistance. We are also thankful to Gary M. Dunny for kindly providing the anti-LtrB antibody and B.

Franz Lang for helpful discussions on mitochondrial genome plasticity. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institutes of Health Research (CIHR) to B.C.. B.C. was a CIHR New Investigator Scholar and is currently an FRSQ Chercheur-Boursier Junior 2 and a McGill University William Dawson Scholar. K.B. holds a Postgraduate Studies Fellowship from NSERC.

## 5.8 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., 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.

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

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

**Cech, T.R.** (1986) The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* **44:** 207-210. Added in thesis

**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.

**Cousineau, B., Lawrence, S., Smith, D. and Belfort, M.** (2000) Retrotransposition of a bacterial group II intron. *Nature* **404:** 1018-1021. Erratum in (2001) *Nature* **414:** 84.

**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.

**Fauron, C, Casper, M., Gao, Y. and Moore, B.** (1995) The maize mitochondrial genome: dynamic, yet functional. *Trends Genet.* **11:** 228-235.

**Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M. and Rochaix, J.D.** (1991) A small chloroplast RNA may be required for *trans*-splicing in *Chlamydomonas reinhardtii*. *Cell* **65:** 135-143.

**Ichiyanagi, 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.

**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.

**Jacquier, A.** (1990) Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *Trends Biochem. Sci.* **15:** 351-354. Added in thesis

**Klein, J.R., Chen, Y., Manias, D.A., Zhuo, J., Zhou, L., Peebles, C.L. and Dunny, G.M.** (2004) A conjugation-based system for genetic analysis of group II intron splicing in *Lactococcus lactis*. *J. Bacteriol.* **186:** 1991-1998.

**Knoop, V., Altwasser, M. and Brennicke, A.** (1997) A tripartite group II intron in mitochondria of an angiosperm plant. *Mol. Gen. Genet.* **255:** 269-276.

**Kohchi, T., Umesono, K., Ogura, Y., Komine, Y., Nakahigashi, K., Komano, T., Yamada, Y., Ozeki, H. and Ohyama, K.** (1988) A nicked group II intron and *trans*-splicing in liverwort, *Marchantia polymorpha*, chloroplasts. *Nucleic Acids Res.* **16:** 10025-10036.

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

**LeBlanc, D.J., Lee, L.N. and Abu-Al-Jaibat, A.** (1992) Molecular, genetic, and functional analysis of the basic replicon of pVA380-1, a plasmid of oral streptococcal origin. *Plasmid* **28:** 130-145.

**Lehmann, K. and Schmidt, U.** (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit Rev. Biochem. Mol. Biol.* **38:** 249-303.

**Li-Pook-Than, J. and Bonen, L.** (2006) Multiple physical forms of excised group II intron RNAs in wheat mitochondria. *Nucleic Acids Res.* **34:** 2782-2790.

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

**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.

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

**Mills, D.A., Choi, C.K., Dunny, G.M. and McKay, L.L.** (1994) Genetic analysis of regions of the *Lactococcus lactis* subsp. *lactis* plasmid pRS01 involved in conjugative transfer. *Appl. Environ. Microbiol.* **60:** 4413-4420.

**Mills, D.A., Manias, D.A., McKay, L.L. and Dunny, G.M.** (1997) Homing of a group II intron from *Lactococcus lactis* subsp. *lactis* ML3. *J. Bacteriol.* **179:** 6107-6111.

**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.

**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.

**O'Sullivan, D., Ross, R.P., Twomey, D.P., Fitzgerald, G.F., Hill, C. and Coffey, A.** (2001) Naturally occurring lactococcal plasmid pAH90 links bacteriophage resistance and mobility functions to a food-grade selectable marker. *Appl. Environ. Microbiol.* **67:** 929-937.

**Perron, K., Goldschmidt-Clermont, M. and Rochaix, J.D.** (2004) A multiprotein complex involved in chloroplast group II intron splicing. *RNA* **10:** 704-711.

**Qin, P.Z. and Pyle, A.M.** (1998) The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol.* **8:** 301-308.

**Qiu, Y.L. and Palmer, J.D.** (2004) Many independent origins of *trans*-splicing of a plant mitochondrial group II intron. *J. Mol. Evol.* **59:** 80-89.

**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.

**Singh, R.N., Saldanha, R.J., D'Souza, L.M. and Lambowitz, A.M.** (2002) Binding of a group II intron-encoded reverse transcriptase/maturase to its high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation. *J. Mol. Biol.* **318:** 287-303.

**Smith, D., Zhong, J., Matsuura, M., Lambowitz, A.M. and Belfort, M.** (2005) Recruitment of host functions suggests a repair pathway for late steps in group II intron retrohoming. *Genes Dev.* **19:** 2477-2487.

**Su, L.J., Waldsich, C. and Pyle, A.M.** (2005) An obligate intermediate along the slow folding pathway of a group II intron ribozyme. *Nucleic Acids Res.* **33:** 6674-6687.

**van der Vossen, J.M., van der Lelie, D. and Venema, G.** (1987) Isolation and characterization of *Streptococcus cremoris* Wg2-specific promoters. *Appl. Environ. Microbiol.* **53:** 2452-2457.

**Watanabe, K. and Lambowitz, A.M.** (2004) High-affinity binding site for a group II intron-encoded reverse transcriptase/maturase within a stem-loop structure in the intron RNA. *RNA.* **10:** 1433-1443.

**Zhou, L., Manias, D.A. and Dunny, G.M.** (2000) Regulation of intron function: efficient splicing in vivo of a bacterial group II intron requires a functional promoter within the intron. *Mol. Microbiol.* **37:** 639-651.

## 5.9 Figure legends

### Figure 5.1

Group II intron splicing pathway and Ll.LtrB secondary structure. **A.** Splicing pathway of group II introns. The 2'OH of the bulged adenosine present in domain VI of the intron (circled A) performs the initial nucleophilic attack on

the exon 1-intron junction (step 1), generating a 2'-5' linkage, also known as a branch point. Then, the 3'-OH of the released exon 1 performs the second nucleophilic attack on the intron-exon 2 junction (step 2), releasing the intron lariat and ligating the flanking exons. Group II intron, black line; exon 1 and 2, E1 and E2; branch point, circled A. **B.** Schematic of the Ll.LtrB secondary structure. The six domains of Ll.LtrB are indicated (I to VI) and the detailed secondary structure of a portion of domain IV is shown (top right). The LtrA start and stop codons are boxed and the Shine-Dalgarno sequence is underlined. Exon 1 and 2 on both sides of the intron are also boxed. Pairs of Greek letters, linked by dashed lines, correspond to long-range interactions between different portions of the intron. The five chosen Ll.LtrB fragmentation sites are mapped with black arrowheads (S1 to S5). EBS1 and EBS2, exon binding site 1 and 2; IBS1 and IBS2, intron binding site 1 and 2; branch point, circled A.

**Figure 5.2**

Ll.LtrB splicing/conjugation assay and relaxase expression levels. **A.** Schematic of the Ll.LtrB splicing/conjugation assay. The *L. lactis* strain used as the donor strain harbours a sex factor that has a defective relaxase (NZ9800Δ*ltrB*::*tet*) while the recipient strain lacks the sex factor (LM0231). The Ll.LtrB intron, along with portions of its exons, was replaced within the chromosomal sex factor by a tetracycline resistance marker, which prevents relaxase expression. The *ltrB* relaxase is constitutively produced from the pDL-$P_{23}$[2]-*ltrB* complementation plasmid. When *ltrB* is interrupted by a splicing

proficient variant of Ll.LtrB, relaxase is produced and mediates transfer of the sex factor from the donor to the recipient cell. The level of sex factor conjugation observed is proportional to the Ll.LtrB splicing efficiency. **B.** Relaxase expression and Ll.LtrB splicing. Northern blots were performed using total RNA from NZ9800 and NZ9800Δ*ltrB* harbouring different pDL-$P_{23}$[2]-based constructs. The amounts of total relaxase RNA (left) and mature relaxase transcript (right) produced were assessed (two independent blots, RNA was loaded in duplicate on the same gel). The exon 2 probe (left panel) (SuppTable 5.1, RT E2) and the splice-junction probe (right panel) (SuppTable 5.1, SJ) are depicted as grey bars. As an RNA loading control, the two membranes were stripped and probed with a 23S rRNA specific probe (SuppTable 5.1, 23S rRNA). **C.** Western blot on total protein extracts using LtrB-specific antibodies. Ll.LtrB group II intron, black line; exon 1 and 2, E1 and E2; branch point, circled A; sex factor, grey; tetracycline resistance marker, Tet; spectinomycin resistance marker, Spc; $P_{23}$ promoter, $P_{23}$; transcription terminator, schematic stem-loop; *L. lactis* chromosome, scribble.

**Figure 5.3**

RT-PCR analyses of excised Ll.LtrB introns and ligated exons. **A.** Schematic of Ll.LtrB splicing (top) and *trans*-splicing (bottom). Primers used for RT-PCR amplifications of released intron lariats or Y-branched molecules are shown as black arrows (BP1 and 2). Ligated exons were amplified by RT-PCR with the RTE1 and 2 primers (open arrows). Agarose gels showing the RT-PCR amplification of Ll.LtrB across the branch point (**B**) and across the *ltrB* ligated

exons (**C**) are presented. RT-PCR analyses were performed on total RNA extracts from NZ9800 or NZ9800Δ*ltrB* harbouring different pDL-$P_{23}$[2] plasmids expressing *ltrB* in one (WT, ΔD5) or two pieces (S1 to S5). S3ΔORF and S4ΔORF represent the ΔORF variant of the S3 and S4 constructs without LtrA complementation. Ll.LtrB group II intron, black line; exon 1 and 2, E1 and E2; branch point, circled A.

**Figure 5.4**

Ll.LtrB *trans*-splicing/conjugation assay. **A.** Strategy to generate a variant of Ll.LtrB fragmented at the Sn site (S1 to S5). The relaxase gene is amplified as two non-overlapping fragments using the indicated primers (thick arrows, SuppTable 5.1). Each fragment is cloned under the control of a $P_{23}$ promoter on pDL-$P_{23}$[2]. **B.** Ll.LtrB *trans*-splicing/conjugation assay. The pDL-$P_{23}$[2]-Sn plasmid is transformed in NZ9800Δ*ltrB*. If the two fragments of the intron can correctly fold, assemble, and *trans*-splice, the *ltrB* exons are ligated allowing the production of relaxase, which is required for sex factor transfer. Ll.LtrB splicing efficiency is monitored by the level of sex factor conjugation from the donor (NZ9800Δ*ltrB*) to the recipient *L. lactis* strain (LM0231). Ll.LtrB group II intron, thick black line; exon 1 and 2, E1 and E2; branch point, circled A; sex factor, grey; tetracycline resistance marker, Tet; spectinomycin resistance marker, Spc; $P_{23}$ promoters, $P_{23}$; transcription terminators, schematic stem-loops; *L. lactis* chromosome, scribble.

**Figure 5.5**

Sex factor conjugation efficiencies observed with different fragmented variants of Ll.LtrB. The fragmentation sites within Ll.LtrB correspond to those shown in Figure 5.1B (S1 to S5). The Ll.LtrB variant column indicates which fragment(s) was/were overexpressed from the pDL-$P_{23}{}^2$ plasmid, and which mutated version of the ORF was used when applicable. The ΔORF + ORF variants correspond to the assays where LtrA was overexpressed from a second plasmid to complement the deleted version within Ll.LtrB. The schematics show linear representations of Ll.LtrB and its six domains (I-VI), and the location of the fragmentation sites. Sex factor conjugation efficiency is the average of three independent assays and the error is the standard deviation of the three obtained values. Ll.LtrB group II intron, black line; exon 1 and 2, empty boxes; branch point, circled A

## 5.10 Tables

**Table 5.1** Sex factor conjugation efficiencies observed with different relaxase variants.

| Donor Strain | SF variant / complementation plasmid | SF conjugation efficiency |
|---|---|---|
| NZ9800 | SF wild-type | [a](7.33 ± 0.58) x $10^{-3}$ |
| NZ9800Δ*ltrB* | SFΔ*ltrB* | (1.88 ± 1.96) x $10^{-9}$ |
| NZ9800Δ*ltrB* | SFΔ*ltrB* / pDL-$P_{23}^{2}$-*ltrB* | (5.80 ± 2.89) x $10^{-2}$ |
| NZ9800Δ*ltrB* | SFΔ*ltrB* / pDL-$P_{23}^{2}$-*ltrB* ΔD5 | (3.79 ± 1.75) x $10^{-9}$ |
| NZ9800Δ*ltrB* | SFΔ*ltrB* / pDL-$P_{23}^{2}$-*ltrB* ΔORF | (1.93 ± 0.96) x $10^{-8}$ |
| NZ9800Δ*ltrB* | SFΔ*ltrB* / pDL-$P_{23}^{2}$-*ltrB* Mat$^{-}$ | (2.86 ± 0.86) x $10^{-4}$ |

SF; sex factor

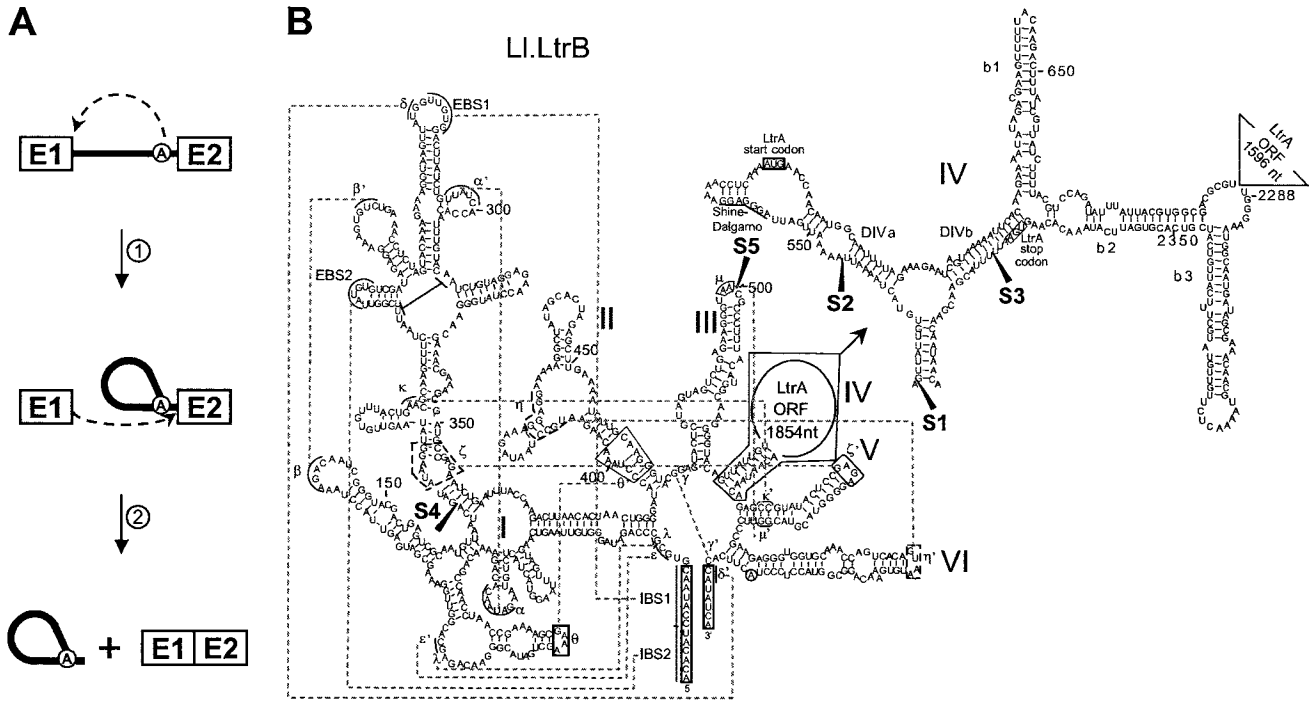
[a] Belhocine *et al.*, 2005

**SuppTable 5.1** Sequences of oligonucleotides used in this study

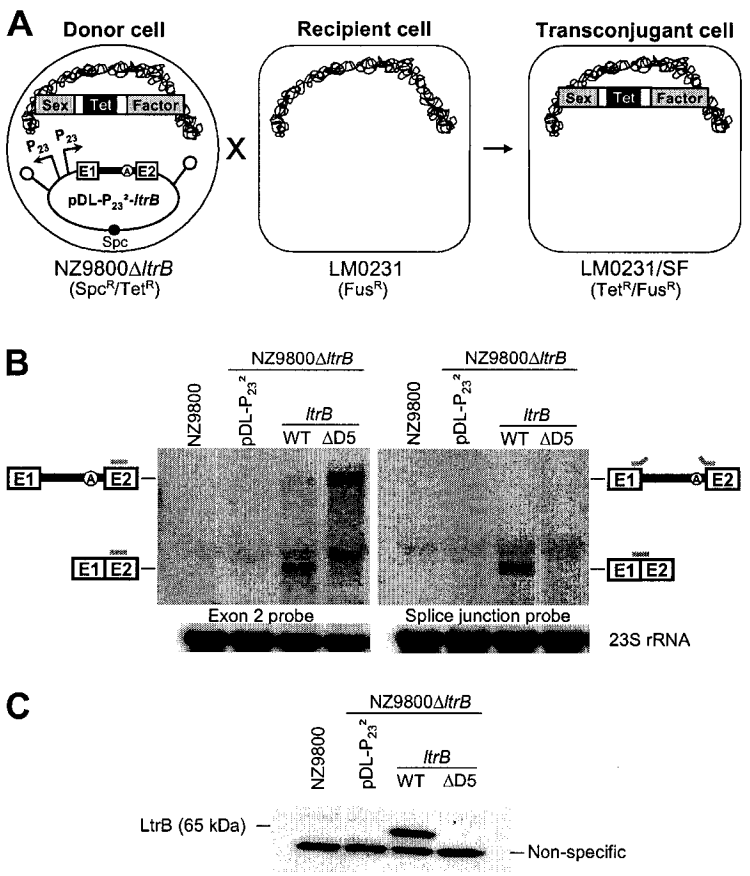
| Name | Sequence |
|---|---|
| $P_{23}$ right top | 5'-GATCCTGATGACAAAAAGAGCAAATTTTGATAAAATAGTATTAGAATTGCGGCCGCG -3' |
| $P_{23}$ right bottom | 5'-GATCCGCGGCCGCAATTCTAATACTATTTTATCAAAATTTGCTCTTTTTGTCATCAG -3' |
| $P_{23}$ left top | 5'-CGTGATGACAAAAAGAGCAAATTTTGATAAAATAGTATTAGAATTGCGCGCCGAT -3' |
| $P_{23}$ left bottom | 5'-CGGCGCGCAATTCTAATACTATTTTATCAAAATTTGCTCTTTTTGTCATCACGAT -3' |
| 5' *ltrB* | 5'- AAAGCGGCCGCCAGAACGATTTAAAGAAGAATTGAA -3' |
| 3' *ltrB* | 5'- AAAAGCGCGCACTACATCCGTTCATAAACTATAC -3' |
| SJ | 5'- AAAATGATATGGTTATGGATGT -3' |
| E2 cDNA | 5'- GAGCCGTTCAATAATAGATTCCA -3' |
| RT E1 | 5'- TTGGTCATCACCTCATCCAATC -3' |
| RT E2 | 5'- CATTTGAGGTTCATCAAGCAGC -3' |
| BP cDNA | 5'- CGATTGTCTTTAGGTAACTCAT -3' |
| BP1 | 5'- CCGTGCTCTGTTCCCGTATCAGC -3' |
| BP2 | 5'- CAGAGCCGTATACTCCGAG -3' |
| S1-P1 | 5'- AAAGCGGCCGCTGTACCCCTTTGCCATGTAAAG -3' |
| S1-P2 | 5'- AAAAGCGCGCGTTATTGTGTACTAAAATTAAAAATTG -3' |
| S2-P1 | 5'- AAAGCGGCCGCAATTTTAGTACACAATAACTGTACCCC -3' |
| S2-P2 | 5'- AAAAGCGCGCAAAAATTGATTAGGGAGGAAAACC -3' |
| S3-P1 | 5' AAAGCGGCCGCAATTCACTTGTGTTTATGAATCAC -3' |
| S3-P2 | 5'- AAAAGCGCGCTTTACGAACGAACAATAACAGAG -3' |
| S4-P1 | 5'- AAAGCGGCCGCTGATTAACATTGCGACTCAGTC -3' |
| S4-P2 | 5'- AAAAGCGCGCGATATAAGGTATAAGTTGTGTTTAC- 3' |
| S5-P1 | 5'- AAAGCGGCCGCTTACCCTTCTCAGACTACTACG -3' |
| S5-P2 | 5'- AAAAGCGCGCCGCCCTTTACATGGCAAAGG -3' |
| 5' *ltrA* | 5'- CGGGGTACCTTGATTAGGGAGGAAAACCTCAA -3' |
| 3' *ltrA* | 5'- CGGGGTACCTCACTTGTGTTTATGAATCACGT -3' |
| 23S rRNA | 5'- ACCCGACAAGGAATTTCGCTACC -3' |

GCGGCCGC, NotI; GCGCGC, BssHII

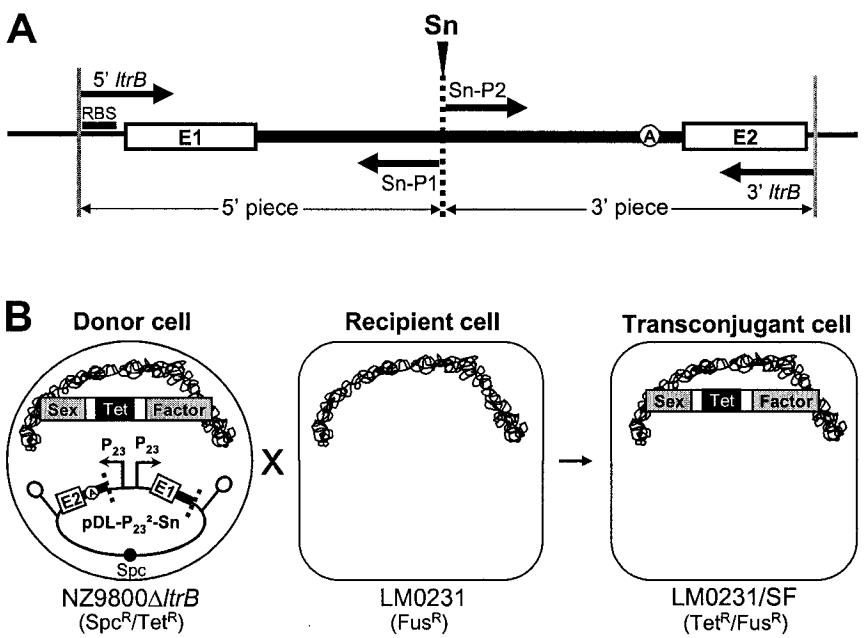
# Figure 5.1


A
E1
E2
①
②
+
B
LI.LtrB
EBS1
EBS2
IBS1
IBS2
LtrA start codon
Shine-Dalgarno
LtrA stop codon
LtrA ORF 1854nt
LtrA ORF 1596 nt
DIVa
DIVb
S1
S2
S3
S4
S5
I
II
III
IV
V
VI
b1
b2
b3

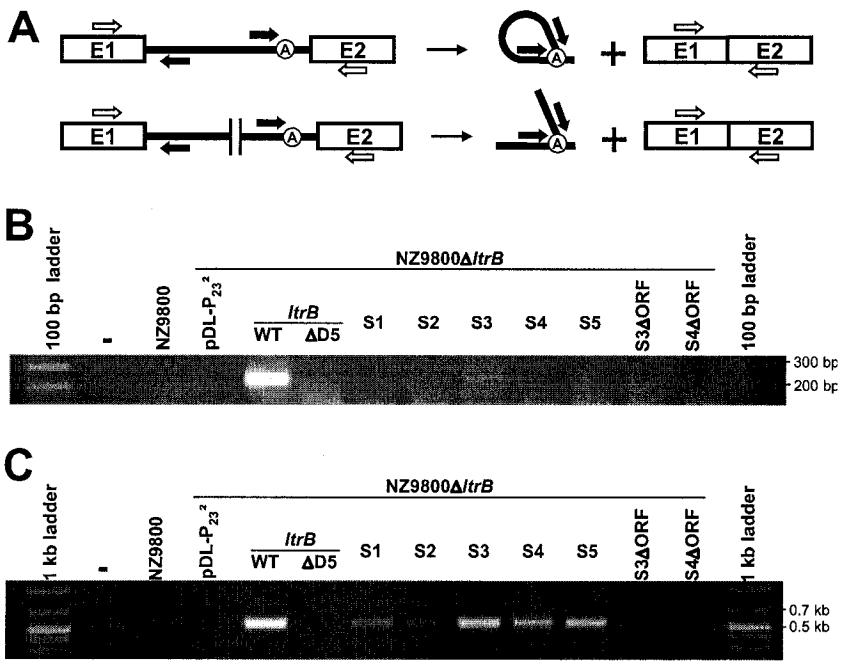
# Figure 5.2


A
Donor cell
Recipient cell
Transconjugant cell
Sex
Tet
Factor
P23
P23
E1
A
E2
pDL-P23²-ltrB
Spc
X
NZ9800ΔltrB
(SpcR/TetR)
LM0231
(FusR)
LM0231/SF
(TetR/FusR)
B
NZ9800ΔltrB
NZ9800
pDL-P23²
ltrB
WT ΔD5
E1
E1 E2
Exon 2 probe
Splice junction probe
23S rRNA
C
NZ9800ΔltrB
NZ9800
pDL-P23²
ltrB
WT ΔD5
LtrB (65 kDa)
Non-specific

# Figure 5.3


A
Sn
5' ltrB
RBS
E1
Sn-P2
A
E2
Sn-P1
3' ltrB
5' piece
3' piece
B
Donor cell
Recipient cell
Transconjugant cell
Sex
Tet
Factor
P23 P23
E2
A
E1
pDL-P23²-Sn
Spc
X
NZ9800ΔltrB
(SpcR/TetR)
LM0231
(FusR)
LM0231/SF
(TetR/FusR)

# Figure 5.4


A
E1
E2
A
+
E1
E2
B
100 bp ladder
-
NZ9800
NZ9800ΔltrB
pDL-P23
ltrB
WT
ΔD5
S1
S2
S3
S4
S5
S3ΔORF
S4ΔORF
100 bp ladder
300 bp
200 bp
C
1 kb ladder
-
NZ9800
NZ9800ΔltrB
pDL-P23
ltrB
WT
ΔD5
S1
S2
S3
S4
S5
S3ΔORF
S4ΔORF
1 kb ladder
0.7 kb
0.5 kb

# Figure 5.5

| Fragmentation Site | LLLtrB variants | Schematic of LLLtrB domains | Sex factor conjugation efficiency |
|---|---|---|---|
| | | I II III IV V VI | |
| No relaxase | Empty vector | | $(1.88 \pm 1.96) \times 10^{-9}$ |
| None | Full length | ORF | $(5.80 \pm 2.89) \times 10^{-2}$ |
| S1 | Two fragments | ORF | $(1.66 \pm 0.88) \times 10^{-6}$ |
| S2 | Two fragments | ORF | $(1.44 \pm 0.37) \times 10^{-4}$ |
| S3 | 5' fragment | ORF | $(0.79 \pm 0.74) \times 10^{-8}$ |
| | 3' fragment | | $(1.57 \pm 0.51) \times 10^{-8}$ |
| | Two fragments | ORF | $(1.16 \pm 0.64) \times 10^{-2}$ |
| | Two fragments / Mat$^-$ | Mat$^-$ | $(3.31 \pm 1.38) \times 10^{-7}$ |
| | Two fragments / ΔORF | | $(4.91 \pm 3.00) \times 10^{-9}$ |
| | Two fragments / ΔORF + ORF | | $(5.00 \pm 1.37) \times 10^{-3}$ |
| S4 | 5' fragment | | $(2.23 \pm 0.44) \times 10^{-8}$ |
| | 3' fragment | ORF | $(2.80 \pm 0.07) \times 10^{-8}$ |
| | Two fragments | ORF | $(4.52 \pm 1.60) \times 10^{-4}$ |
| | Two fragments / Mat$^-$ | Mat$^-$ | $(1.35 \pm 0.49) \times 10^{-8}$ |
| | Two fragments / ΔORF | | $(3.90 \pm 1.56) \times 10^{-9}$ |
| | Two fragments / ΔORF + ORF | | $(2.38 \pm 0.24) \times 10^{-4}$ |
| S5 | Two fragments | ORF | $(6.03 \pm 1.64) \times 10^{-5}$ |

# CHAPTER SIX

## Group II intron *trans*-splicing versatility: experimental support for the origin of snRNAs

### 6.1 Preface

Our original goal was to address the evolutionary hypothesis of group II intron fragmentation towards snRNAs. In this perspective, we aimed at investigating the potential of fragmenting a group II intron consistently with the origin of the five snRNAs, and monitor its *trans*-splicing efficiency. In the previous chapter, we validated the use of Ll.LtrB as a model system to study *trans*-splicing of group II introns. We reported that Ll.LtrB tolerates fragmentation at domains I, III and different locations in domain IV. These fragmentation sites were engineered to mimick natural group II introns' fragmentations supporting *trans*-splicing.

To assess if group II introns can be fragmented at locations distinct from those observed in nature, we undertook a comprehensive approach to test all possible fragmentations of Ll.LtrB that are permissive for *trans*-splicing. We engineered a Tn5 transposon to harbour a transcriptional terminator followed by a promoter, and used this transposon to introduce random fragmentations within the Ll.LtrB intron. We exploited the dependence of conjugation of the pLE12

mobilizable plasmid on Ll.LtrB splicing. This unique feature of Ll.LtrB was used to select for splicing-proficient fragmented variants of Ll.LtrB, thus selecting for all possible fragmentations supporting *trans*-splicing.

# Group II intron *trans*-splicing versatility: experimental support for the origin of snRNAs

**Kamila Belhocine, Anthony B. Mak, and Benoit Cousineau**
Manuscript in preparation

## 6.2 Introductory paragraph

Self-splicing group II introns are large ribozymes found in organelles, bacteria and archaea (Michel and Ferat, 1995; Lambowitz and Zimmerly, 2004). They share a conserved secondary structure composed of six domains (I-VI) radiating from a central wheel (Michel and Ferat, 1995; Lambowitz and Zimmerly, 2004). Many harbour, within domain IV, an open reading frame (ORF) with reverse transcriptase, maturase and for the most part endonuclease activities (Lambowitz and Zimmerly, 2004). Group II introns require intron-encoded or free-standing maturases to excise from primary RNA transcripts *in vivo* (Michel and Ferat, 1995; Lambowitz and Zimmerly, 2004). Some ORF-containing group II introns were shown to be mobile retroelements invading DNA by retrohoming or retrotransposition (Lambowitz and Zimmerly, 2004). Group II introns are also hypothesized to be the ancestors of the spliceosome-dependent nuclear introns and the five small nuclear RNAs (snRNAs – U1, U2, U4, U5 and U6) that are part of the spliceosome (Sharp, 1985,1991; Cech, 1986; Jacquier, 1990; Cavalier-Smith, 1991; Bonen, 1993; Roger and Doolittle, 1993; Michel and Ferat, 1995; Lambowitz and Zimmerly, 2004). The ability of some fragmented organellar

group II introns to undergo splicing *in trans* supports the theory that the five snRNAs evolved from portions of group II introns (Sharp, 1991; Bonen, 1993). 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*. We report that numerous fragmentation sites located throughout Ll.LtrB support splicing *in trans* showing that this intron is remarkably tolerant to fragmentation. Moreover, the observed fragmentation sites within Ll.LtrB were found clustered between regions functionally and/or structurally analogous to the five snRNAs. This study provides *in vivo* experimental support to long-standing 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, 1985, 1991; Cech, 1986; Jacquier, 1990; Cavalier-Smith, 1991; Bonen, 1993; Roger and Doolittle, 1993; Michel and Ferat, 1995; Lambowitz and Zimmerly, 2004).

## 6.3 Results and discussion

The Ll.LtrB group II intron was found in two nearly identical conjugative elements of *L. lactis*: the pRS01 plasmid (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, an enzyme that initiates conjugation by nicking the conjugative element at the origin of transfer (*oriT*) (Byrd and Matson, 1997). Splicing of Ll.LtrB from its pre-mRNA transcript is thus essential for LtrB

production and subsequent DNA transfer (Mills *et al.*, 1996; Shearman *et al.*, 1996).

We developed a sensitive splicing/conjugation assay in *L. lactis* ($10^5$-fold detection range) where Ll.LtrB splicing is monitored by the conjugation efficiency of its carrier, the pLE12 mobilizable plasmid (Figure 6.1). pLE12 contains a portion of pRS01 that includes the *oriT* and the *ltrB* relaxase gene (Mills *et al.*, 1994; Belhocine *et al.*, 2004). The *L. lactis* strain used as the donor in the splicing/conjugation assay harbours a relaxase-deficient chromosomal sex factor (NZ9800Δ*ltrB*::*tet*) (Ichiyanagi *et al.*, 2002). When pLE12 bears a catalytic mutant of Ll.LtrB (pLE12ΔDV, Belhocine *et al.*, 2004) or when the intron-encoded maturase LtrA located within domain IV is deleted (pLE12ΔORF, Belhocine *et al.*, 2004), the transfer efficiency drops significantly (~$10^4$-fold, Table 6.1). This demonstrates that conjugation of pLE12 is dependent on Ll.LtrB splicing.

Ll.LtrB splices *in trans* when fragmented independently within domains I, III or IV, either upstream or downstream from *ltrA* (Belhocine *et al.*, 2007). Interestingly, all these fragmented Ll.LtrB variants, which were engineered to mimic natural *trans*-splicing introns, were able to *trans*-splice (Belhocine *et al.*, 2007). We therefore developed a Tn5-based genetic screen to investigate the full *trans*-splicing potential of Ll.LtrB. We engineered a Tn5 transposon (Goryshin and Reznikoff, 1998) to carry a transcriptional terminator followed by a constitutive *L. lactis* promoter (Figure 6.1). Insertion of this transposon into a target gene fragments the RNA transcript; the 5' fragment ending at the

transcriptional terminator and the 3' fragment extending from the inserted promoter.

We generated a bank of pLE12 plasmids (pLE12::Tn) where the entire intron and portions of exon 1 and 2 were saturated with unique Tn5 insertions. The bank was used to transform NZ9800Δ*ltrB* and the transformants were pooled for conjugation (Figure 6.1). pLE12::Tn variants can only be transferred to recipient cells if the Tn5-fragmented Ll.LtrB intron *trans*-splices and ligates its flanking exons, resulting in relaxase expression. Plasmids were recovered from 66 independent transconjugant cells and the Tn5 insertion sites were identified. The great majority of insertions were found within domain I between intron positions 37 and 181 (63.6%) and within domain IV (21.2%) either upstream of *ltrA* or downstream from the maturase domain (Figure 6.2a, black arrowheads). Interestingly, these three locations are comparable to fragmentations observed in naturally occurring organellar *trans*-splicing group II introns (Bonen, 1993). The remaining Tn5 insertions within domains I and II do not correspond to characterised fragmentations of group II introns. No introns fragmented within domains III, V and VI were isolated. The presence of multiple Tn5 insertions within *ltrA*, downstream of the maturase domain, suggests that these *trans*-splicing introns express active truncated maturase variants of LtrA. Accordingly, when the conjugation screen was repeated in the presence of LtrA expressed *in trans* from pDE*ltrA*, the pressure to produce an active maturase from the intron was relieved and consequently, numerous Tn5 insertions were found randomly distributed within *ltrA* (Figure 6.2a, open arrowheads). These results confirm the

importance of the maturase function of LtrA for Ll.LtrB *trans*-splicing (Belhocine *et al.*, 2007) and demonstrate that a portion of domain I and domain IV are the regions within Ll.LtrB that best tolerate fragmentation.

Selected plasmids carrying different Tn5-fragmented introns (Figure 6.2a, F1-F6) were further analyzed and demonstrated significantly higher conjugation efficiencies ($10^{-6}$-$10^{-5}$) than the pLE12 variants harbouring splicing-deficient introns ($10^{-9}$) (Table 6.1). Deletion of the promoter from two selected variants significantly decreased their conjugation efficiency (Table 6.1, F1P$^-$ and F3P$^-$). These results demonstrate that production of the second fragment is required for Ll.LtrB *trans*-splicing and that the relaxase gene is indeed expressed as two independent transcripts. Detection of ligated exons by RT-PCR confirmed that, when the intron is fragmented at these locations, both Ll.LtrB fragments are able to fold, assemble, and *trans*-splice in order to correctly ligate their flanking exons (Figure 6.3, F1-F6). Finally, when LtrA was provided *in trans*, all fragmented introns *trans*-spliced more efficiently (Table 6.1, 1.1- to 14.2-fold), comparable to Ll.LtrB *cis*-splicing from pLE12 (Table 6.1, 6.5-fold). These results suggest that LtrA production is sub-optimal when expressed from its natural promoter and that the *trans*-splicing ability of Ll.LtrB is only slightly affected by the increased expression of LtrA.

Next, to identify fragmentation sites in other locations within Ll.LtrB, introns harbouring Tn5 insertions limited to sub-domains of the intron were subjected to the screen (Figure 6.2b; see brackets, I-III, II-III, V-VI). As found in the original full-length screens, the majority of Tn5 insertions isolated from the I-III screen were located at the beginning of domain I, some at identical positions

(cf. Figure 6.2a and 2b). The II-III screen, which excluded the beginning of domain I and most of domain IV, allowed for the identification of other regions supporting Ll.LtrB fragmentation within domains I, II and III. The V-VI screen revealed that Ll.LtrB *trans*-splices when fragmented at certain positions within domain VI. Even though four Tn5 insertions were found within domain V, they induced fragmentations structurally similar to those found at the 3' end of domain IV. Upon insertion, Tn5 creates a 9-nucleotide direct repeat (Goryshin and Reznikoff, 1998). Therefore, although the first fragments end at the indicated arrowheads (Figure 6.2), the second *ltrB* fragments start 9 nucleotides upstream of the arrowheads. The second fragment of the four Tn5 insertions in domain V actually harbours an almost complete domain V. Selected samples from the partial screens (Figure 6.2b, F7-F11) were found to have similar conjugation efficiencies as the Ll.LtrB *trans*-splicing variants analyzed from the initial screens (Table 6.1, cf. F7-F11 with F1-F6).

To further evaluate our screens, we identified Tn5 insertion sites in 43 independent plasmids from the original pLE12::Tn bank. The insertions had a random distribution both within the intron and flanking *ltrB* exons and occurred in both orientations. Conversely, the 247 plasmids selected by conjugation displayed a biased distribution of Tn5 insertions (Figure 6.2). All Tn5 insertions were found in the same orientation as *ltrB* such that the terminator and promoter created bipartite introns. In addition, all fragmented introns harboured either full-length *ltrA* or a variant truncated downstream from the maturase domain. However, when LtrA was provided *in trans* during selection, introns containing Tn5

insertions throughout *ltrA* were selected. These results confirm the reliability of our screens.

The identified fragmentation sites were located in specific regions of the Ll.LtrB secondary structure (Figure 6.4b, grey areas). Interestingly, fragmentation sites were clustered between intron sub-domains which are functionally and/or structurally equivalent to the spliceosomal snRNAs (Figure 6.4, cf. a and b). Indeed, U1 is functionally similar to the 5' region of group II introns (ε' and λ) in that they both recognize the 5' end of introns through base pairing. In fact, the ε-ε' interaction within Ll.LtrB (Figure 6.2) may be extended to the last nucleotide of exon I similarly to the interaction of U1 with the 5' splice site of nuclear introns (Jacquier, 1990; Nilsen, 1994; Patel and Steitz, 2003). The portion of domain I that recognizes both flanking exons by base pairing (EBS1, EBS2, δ) is functionally analogous to U5 (Newman and Norman, 1992; Nilsen, 1994; Hetzer *et al.*, 1997; Patel and Steitz, 2003;). The catalytic domain of group II introns (domain V) was shown to be structurally (Seetharaman *et al.*, 2006) and functionally (Shukla and Padget, 2002) equivalent to U6. Finally, the base pair interaction between U2 and the 3' end of nuclear introns bulges the branch point nucleotide (A) similarly to the domain VI stem of group II introns (Jacquier, 1990; Nilsen, 1994; Patel and Steitz, 2003).

The first two sets of fragmentations identified within Ll.LtrB were located in domain I upstream from the region analogous to U1 (Figure 6.4b, A). Therefore, they likely separated the 5' end of the intron from the domains analogous to the five snRNAs (Figure 6.4c, model A). Similarly, fragmentations within domain VI

separated the domains analogous to the five snRNAs from the 3' end of the intron (Figure 6.4, model E). Interestingly, although numerous efficiently *trans*-splicing Ll.LtrB variants fragmented at these sites were selected, no fragmentations in those areas are observed in natural *trans*-splicing group II introns.

The next sets of Ll.LtrB fragmentations (Figure 6.4b, B) clustered between regions equivalent to U1 and U5 (Figure 6.4c, model B). These sites correspond to one of the two fragmentations observed in the tripartite group II intron of the *nad5* mitochondrial gene from *Oenothera berteriana* (Knoop *et al.*, 1997). Moreover, the 5' fragment of these Tn5-interrupted introns correspond to the spliced leader RNAs (SL RNAs) that are incorporated *in trans* at the 5' end of transcripts in nematodes, trypanosomes and dinoflagellates (Bonen, 1993; Zhang *et al.*, 2007). The intron segment removed from the *trans*-spliced exon was shown to be functionally equivalent to U1 (Bruzik and Steitz, 1990; Bonen, 1993).

Fragmentations within domain I downstream of the EBS1-2 and δ sequences (Figure 6.4b, model C) resulted in the association of the U1- and U5-analogous portions of Ll.LtrB with the 5' intron fragment while U4-, U6- and U2-equivalent portions would be contained within the 3' intron fragment (Figure 6.4c, model C). On the other hand, fragmentations at the end of domain IV downstream of *ltrA* (Figure 6.4b, D) support the model in which U1-, U5- and U4-analogous portions of Ll.LtrB remain associated with the 5' intron fragment, and U6- and U2-equivalent segments with the 3' intron fragment (Figure 6.4c, model D). While no group II introns have been found fragmented in domain I downstream of the EBS1-2 and δ sequences, numerous *trans*-splicing organellar introns are

fragmented within domain IV either upstream or downstream from the ORF (Bonen, 1993; Qiu and Palmer, 2004). These include tripartite introns from the *nad5* mitochondrial gene of *Oenothera berteriana* (Knoop *et al.*, 1997) and the *psaA* chloroplast gene of *Chlamydomonas reinhardtii* (Goldschmidt-Clermont *et al.*, 1991).

Fragmentations within domains II and III and within domain IV upstream or within *ltrA* (Figure 6.4b, C/D) left the intron domains corresponding to U1 and U5 associated with the 5' intron fragment, and those corresponding to U6 and U2 with the 3' intron fragment. Nevertheless, we could not predict which of these Tn5-induced fragmentations support the separation between the U5/U4 and U4/U6 analogous regions (Figure 6.4c, models C/D). The function of U4 is accessory in the spliceosome (Nilsen, 1994; Patel and Steitz, 2003) and it is therefore difficult to delineate an analogous region for this snRNA within group II introns. Finally, we did not select introns fragmented between domains V (U6) and VI (U2). This might be due to the fact that these domains are the most conserved, very important for splicing, and adjacent (Michel and Ferat, 1995; Lambowitz and Zimmerly, 2004).

Since Tn5-induced fragmentations are broadly distributed throughout Ll.LtrB and these fragmented introns have similar *trans*-splicing efficiencies regardless of the position of their fragmentation sites, we show that group II introns are more versatile in their ability to splice *in trans* than expected from previously characterized *trans*-splicing introns. Moreover, the selected fragmentation sites within Ll.LtrB lie at positions supporting the division between regions analogous to the five snRNAs. Taken together, the results presented in

this study provide *in vivo* experimental support for long-standing 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, 1985, 1991; Cech, 1986; Jacquier, 1990; Cavalier-Smith, 1991; Bonen, 1993; Roger and Doolittle, 1993; Michel and Ferat, 1995; Lambowitz and Zimmerly, 2004)..

## 6.4 Methods

### 6.4.1 Plasmids

pBS*ltrB* was constructed by cloning the XbaI/SwaI fragment from pLE12 (Mills *et al.*, 1996) into pBlueScript SK- (SmaI). This plasmid was used as the target for the Tn5 *in vitro* transposition assay. This insertion bank (pBS*ltrB*::Tn, 22 insertions/position) 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 domain IV, was used to create the pBS*ltrB*AII::Tn bank (1.9 insertions/position). The AatII site was engineered by site-directed mutagenesis, *changing* nucleotides 2384-2387 from 5'-ACGA-3' to 5'-CGTC-3'. The various pLE12::Tn banks were engineered by sub-cloning 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 was removed.

pMOD$P_{23}$ was constructed by cloning $P_{23}$ (complementary oligos, SuppTable 6.1) between the two Tn5 mosaic ends (BamHI) within the pMOD™ - 2<MCS> Transposon Construction Vector (Epicentre). The *pepN* terminator was

PCR-amplified from pLEII*td*+KR" (SuppTable 6.1) and cloned upstream of $P_{23}$ (HindIII). Finally, the *aad9* gene (Spc) was PCR-amplified from pDL278 (SuppTable 6.1) and cloned between the terminator and $P_{23}$ (PstI).

pDE*ltrA* was constructed by cloning the PCR-amplified *ltrA* gene (SuppTable 6.1) in the pDL-$P_{23}$[2] plasmid (Belhocine *et al.*, 2007) under the control of $P_{23}$ (NotI). The *ermAM* gene (Erm) was PCR-amplified from pMSP3545 (SuppTable 6.1) and cloned in the resulting plasmid (SmaI). Finally, the *aad9* gene (Spc) was deleted (PacI/SwaI).

### 6.4.2 *in vitro* transposition assay

The Tn5 transposon was excised from pMOD (PvuII) and incubated with the pBS*ltrB* or pBS*ltrB*AII target plasmids (14:1 ratio) at 37°C for 2 hours in the presence of the EZ::Tn5™ transposase recombinant protein (Epicentre). Conjugations and RT-PCR assays were performed as previously described (Belhocine *et al.*, 2007).

## 6.5 Acknowledgements

We thank L. Bonen, L. Kronstad, F. Lang, R. Lease, A. Stern and 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. was a CIHR New Investigator Scholar and is currently an FRSQ Chercheur-Boursier Junior 2 and a McGill University William Dawson Scholar. K.B. holds a Postgraduate Studies Fellowship from NSERC.

## 6.6 References

**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.

**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.

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

**Bruzik, J.P. and Steitz, J.A.** (1990) Spliced leader RNA sequences can substitute for the essential 5' end of U1 RNA during splicing in a mammalian in vitro system. *Cell* **62:** 889-899.

**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.

**Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M., and Rochaix, J.D.** (1991). A small chloroplast RNA may be required for *trans*-splicing in *Chlamydomonas reinhardtii*. *Cell* **65:** 135-143.

**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.

**Ichiyanagi, 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.

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

**Knoop, V., Altwasser, M. and Brennicke, A.** (1997) A tripartite group II intron in mitochondria of an angiosperm plant. *Mol. Gen. Genet.* **255:** 269-276.

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

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

**Mills, D.A., Choi, C.K., Dunny, G.M. and McKay, L.L.** (1994) Genetic analysis of regions of the *Lactococcus lactis* subsp. *lactis* plasmid pRS01 involved in conjugative transfer. *Appl. Environ. Microbiol.* **60:** 4413-4420.

**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: unravelling 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.

**Qiu, Y.L. and Palmer, J.D.** (2004) Many independent origins of *trans*-splicing of a plant mitochondrial group II intron. *J. Mol. Evol.* **59:** 80-89.

**Roger, A.J. and Doolittle, W.F.** (1993) Molecular evolution. Why introns-in-pieces? *Nature* **364:** 289-290.

**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.

**Zhang, H., Hou, Y., Miranda, L., Campbell D.A., Sturm N.R., Gaasterland T. and Lin, S.** (2007) Spliced leader RNA *trans*-splicing in dinoflagellates. *Proc. Natl. Acad. Sci. USA* **104:** 4618-4623.

## 6.7 Figure Legends

### Figure 6.1

Tn5 transposition assay and *trans*-splicing/conjugation assay. **A.** Schematic of the Pst I fragment of pRS01 sub-cloned into the pLE12 plasmid. Genes are represented as boxes and the Ll.LtrB intron is shown as a black thick line. The origin of transfer is depicted as a black circle. The restriction sites used to sub-clone the various bank portions are indicated. **B.** Tn5 transposition assay and generation of the various pLE12::Tn banks. The Tn5 transposon is mixed with the pBS*ltrB* target plasmid and a pBS*ltrB*::Tn insertion bank is thereby generated, harbouring Tn5 insertions at all possible locations of the pBS*ltrB* plasmid (step 1). To generate a pLE12::Tn bank bearing Tn insertions in a specific portion of the

*ltrB* gene, this portion is excised from the pBS*ltrB*::Tn bank and cloned into the pLE12 plasmid (step 2). **C.** *Trans*-splicing/conjugation assay. The pLE12::Tn banks are screened by conjugation between *L. lactis* strains. This assay selects for plasmids that harbour proficient Ll.LtrB *trans*-splicing variants that produce relaxase and initiate their conjugative transfer. The conjugation machinery to support plasmid transfer is provided from the relaxase-deficient chromosomal sex factor. Ll.LtrB, thick black line; exon 1 and 2, E1 and E2; Origin of conjugative transfer (*oriT*), black circle; chloramphenicol resistance gene, Cam; Tn5 transposon, grey box; Tn5 mosaic ends, arrowheads; *pepN* transcriptional terminator, schematic stem-loop; $P_{23}$ constitutive promoter, broken arrow; chromosome, scribble.

**Figure 6.2**

Selected Tn5 insertion sites supporting Ll.LtrB *trans*-splicing. The Tn5 insertion sites are depicted on the Ll.LtrB secondary structure model. **a.** Ll.LtrB full-length screen. The screen was performed with (open arrowheads) or without (black arrowheads) the expression of LtrA *in trans* (pDE*ltrA*). **b.** Screens of Ll.LtrB sub-domains. The various sections of Ll.LtrB that were independently screened are delineated by brackets (I-III, open arrowheads; II-III, black arrowheads; V-VI, grey arrowheads). Isolates found multiple times (numbers) and the ones further analyzed are indicated (F1 to F11). The Ll.LtrB intron is 2492 nucleotides long with the *ltrA* gene comprising 1797 nucleotides (599 aa). Exon binding sites 1 and 2; EBS1 and 2; intron binding sites 1 and 2, IBS1 and 2; Exons, open boxes; *ltrA*, grey shading; ribosome binding site, RBS; LtrA reverse

transcriptase domain, RT; LtrA maturase domain, Mat; LtrA endonuclease domain, Endo; tertiary interactions, pairs of Greek letters.

**Figure 6.3**

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 *ltrB* ligated exons junction. Amplifications were performed on total RNA extracts of NZ9800Δ*ltrB* containing plasmids harbouring either continuous (pLE12, pLE12ΔD5) or fragmented introns (F1-F6). Variants of the F1 and F3 constructs where the $P_{23}$ promoter was deleted were also analyzed ($F1P^-$ and $F3P^-$).

**Figure 6.4**

Model for group II intron fragmentation between regions analogous to snRNAs. **a.** Key interactions between snRNAs and nuclear introns (adapted from Patel and Steitz, 2003). **b.** Distribution of the selected fragmentation sites within the schematic Ll.LtrB secondary structure. Ll.LtrB regions functionally and/or structurally analogous to the snRNAs (a, color coded) are compared to the position of the selected Tn5 insertion sites (grey areas). **c.** Linear representations of the selected fragmentation sites within Ll.LtrB. Represented are the different groups of Ll.LtrB *trans*-splicers (models A-E) and the Ll.LtrB regions functionally and/or structurally analogous to the five snRNAs.

## 6.8 Tables

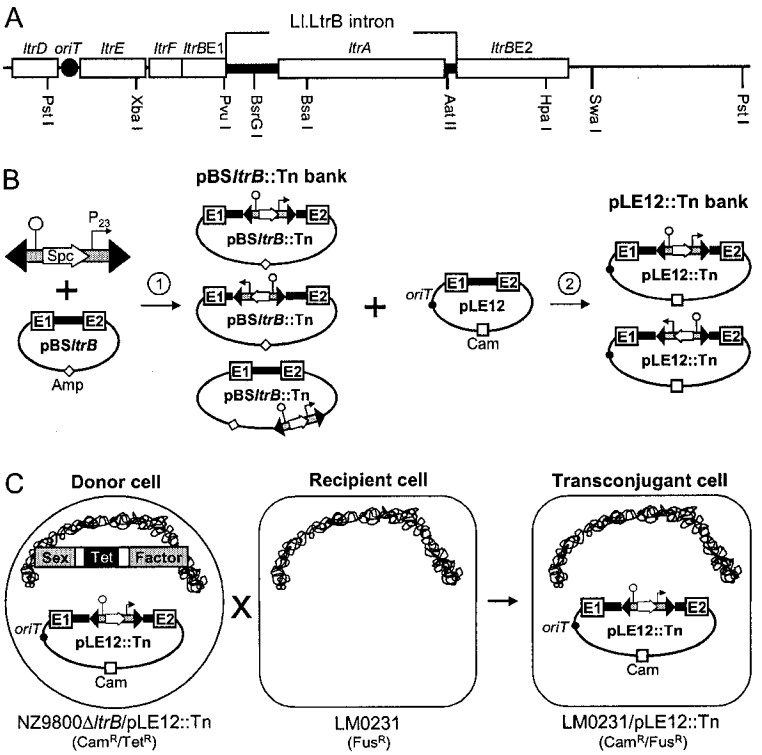
**Table 6.1:** Conjugation efficiencies of pLE12 variants

| 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$^-$ | $(3.17 \pm 0.36) \times 10^{-7}$ | $(5.75 \pm 4.20) \times 10^{-7}$ |

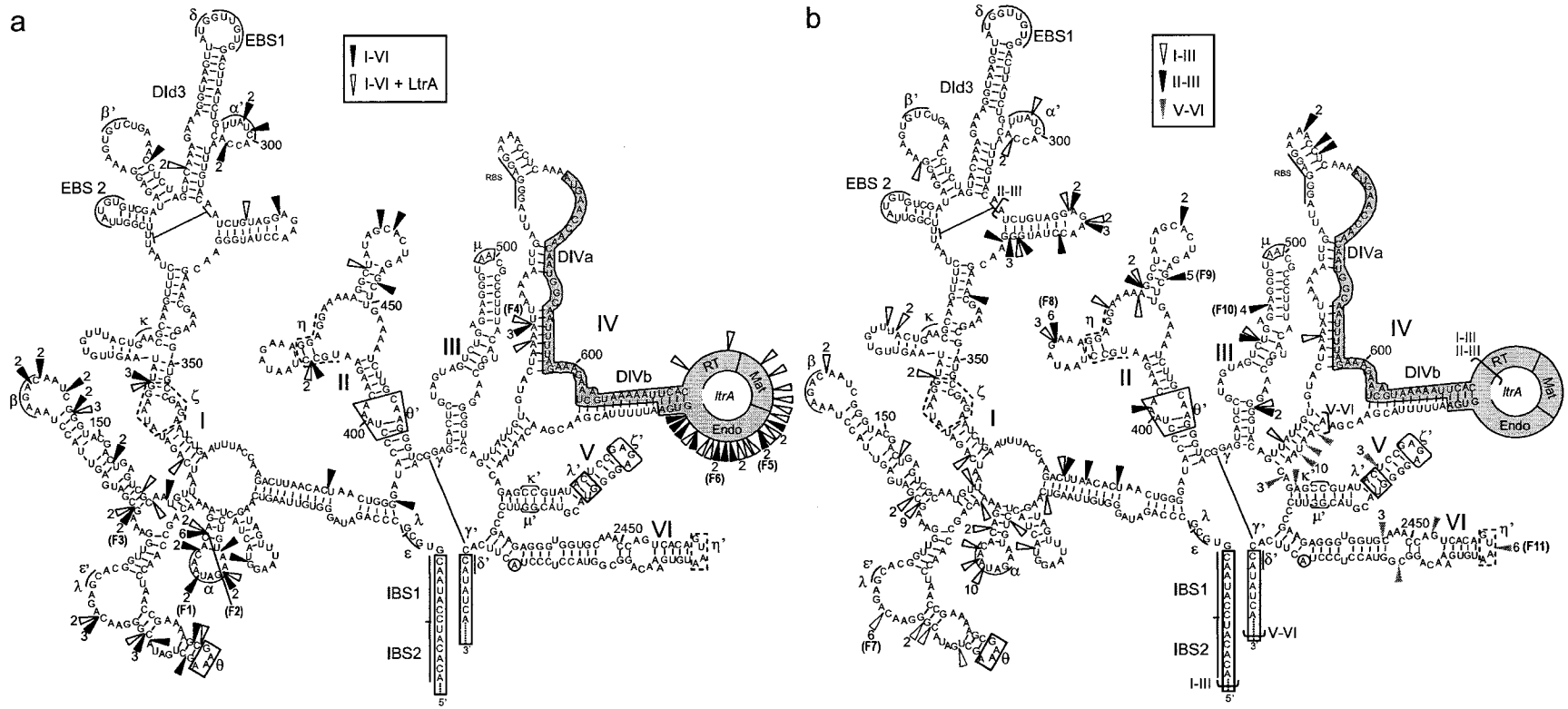
**SuppTable 6.1** Primers

| 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' |

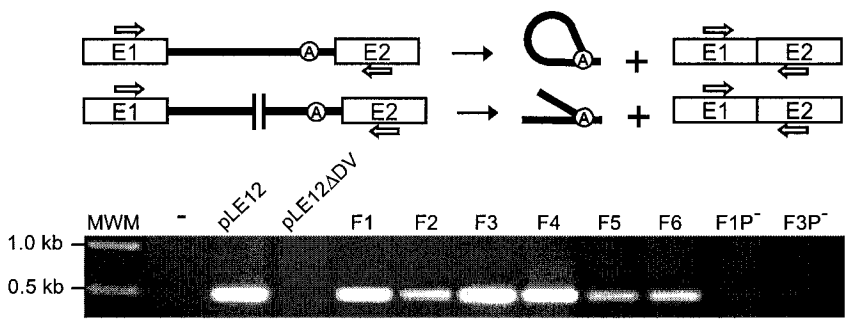
# Figure 6.1


A
LI.LtrB intron
ltrD oriT ltrE ltrF ltrBE1 ltrA ltrBE2
Pst I
Xba I
Pvu I
BsrG I
Bsa I
Aat II
Hpa I
Swa I
Pst I
B
pBSltrB::Tn bank
pLE12::Tn bank
P23
Spc
E1
E2
pBSltrB
Amp
pBSltrB::Tn
pLE12
Cam
pLE12::Tn
oriT
1
2
C
Donor cell
Recipient cell
Transconjugant cell
Sex
Tet
Factor
X
NZ9800ΔltrB/pLE12::Tn
(CamR/TetR)
LM0231
(FusR)
LM0231/pLE12::Tn
(CamR/FusR)

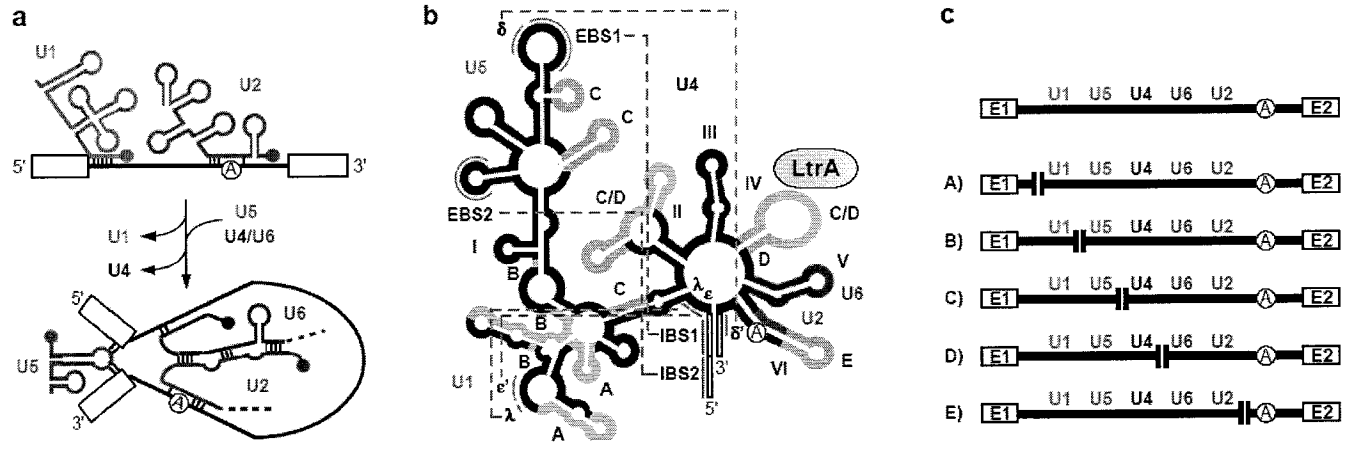
# Figure 6.2


a
b
EBS1
Dld3
EBS 2
I-VI
I-VI + LtrA
I-III
II-III
V-VI
RBS
DIVa
DIVb
IBS1
IBS2
RT
Mat
ItrA
Endo
I
II
III
IV
V
VI

# Figure 6.3


E1
E2
A
+
E1
E2
MWM
-
pLE12
pLE12ΔDV
F1
F2
F3
F4
F5
F6
F1P-
F3P-
1.0 kb
0.5 kb

# Figure 6.4


a
U1
U2
5'
3'
A
U5
U1
U4/U6
U4
5'
U6
U5
U2
3'
A
b
δ
EBS1
U5
C
C
U4
III
LtrA
IV
C/D
EBS2
C/D
II
I
B
C
λ
ε
D
V
U6
B
IBS1
δ'
A
U2
U1
B
IBS2
3'
VI
E
ε'
A
λ'
5'
c
U1 U5 U4 U6 U2
E1
A
E2
A)
B)
C)
D)
E)

# CHAPTER SEVEN

## Conclusions and perspectives

The goal of our research was to address evolutionary theories about group II introns and provide them with experimental support. We used the Ll.LtrB intron from *Lactococcus lactis* as a model system because it efficiently splices and mobilizes *in vivo*. Also, Ll.LtrB is naturally associated with conjugative elements. Because it interrupts an essential gene for conjugation, Ll.LtrB controls conjugation of its host elements. Therefore, this intron offers the unique advantage of generating a quantifiable phenotype upon splicing.

### 7.1 On lateral transfer of group II introns

In Chapters Two, Three and Four, we described the various conjugation/mobility assays we built to study dissemination of the Ll.LtrB intron between various *L. lactis* strains and from *L. lactis* to *E. faecalis*. We showed that transfer of both mobilizable and non-mobilizable plasmids promotes dissemination of Ll.LtrB in *L. lactis*. Upon conjugative transfer, the intron is able to invade its homing site by retrohoming, or the host chromosome by retrotransposition (Chapter Two). We also showed that the pLE12 mobilizable plasmid is actively transferred to *E. faecalis* and promotes Ll.LtrB invasion of the new host by retrohoming (Chapter Two). Then, we showed that the *L. lactis* sex

factor can also be transferred between *L. lactis* strains, inducing high Ll.LtrB retrohoming efficiency in the new host (Chapter Three). We described co-transfer of the pLE12 plasmid and the sex factor between *L. lactis* strains, and noted competitive invasion of homing sites by the Ll.LtrB introns present on the two co-transferred elements (Chapter Three). Finally, we provided the first demonstration of conjugative transfer of the *L. lactis* sex factor and pRS01 to a different bacterial species (*E. faecalis*), and showed that pRS01 transfer promotes dissemination of Ll.LtrB in the new host by retrohoming (Chapter Four).

Taken together, these results demonstrate that conjugation can be a means of dispersal of group II introns in the bacterial kingdom. Recently, Nisa-Martinez and co-workers published a study describing retrohoming of the *Sinorhizobium meliloti* RmInt1 intron in the genome of a new *S. meliloti* host strain following its acquisition by conjugative transfer (Nisa-Martinez et al., 2007). The natural homing site of RmInt1 resides in an insertion sequence that exists in multiple copies in the genome of the *S. meliloti* strain used as a recipient (Martinez-Abarca et al., 1998; Nisa-Martinez et al., 2007). RmInt1 was shown to actively invade a fraction of its natural homing sites present in the genome of the recipient host strain following conjugative transfer. This study further illustrated the potential of group II introns to invade resident DNA of a new host following conjugative acquisition.

### 7.1.1 Relationship between conjugative transfer and Ll.LtrB mobility

We consistently observed that following conjugative transfer, retrohoming of Ll.LtrB expressed from the sex factor or from the pTRK28::pRS01 co-integrate was higher than when the intron originates from the pLE12 plasmid (Chapters Three and Four). This finding is intriguing since pLE12 consists of a portion of pRS01 cloned into pLE1. Therefore, the relaxase gene is under the control of its natural promoter and should be transcribed at the same levels as from pRS01 or the sex factor. Moreover, pLE1-based plasmids are estimated to be present in ~25 copies per cell (Ichiyanagi et al., 2003), which should result in an increased expression of Ll.LtrB RNPs. These observations may reflect that intron RNP amounts and/or mobility efficiency are tightly regulated by unknown mechanisms upon an increase in Ll.LtrB transcription.

Interestingly, we observed that an increase in sex factor conjugation efficiency results in increased Ll.LtrB mobility efficiency (Belhocine K. and Cousineau B., unpublished data). A clumping phenotype was consistently observed when the sex factor-containing strain NZ9800 was transformed with a specific construct based on the pLE1 plasmid; however, we could not identify the exact plasmid effector causing the clumping phenotype. Nevertheless, we consistently observed an increase in sex factor conjugation efficiency upon occurrence of clumping, and increased conjugation was systematically accompanied by an increase in Ll.LtrB retrohoming efficiency (Belhocine K. and Cousineau B., unpublished data). These observations suggest that the presence of Ll.LtrB in an essential gene for conjugation not only mediates its lateral transfer,

but also likely results in high expression levels of the intron in the donor cell during onset of conjugation. If this increased mobility also happens in recipient cells, it would augment chances of intron dissemination in the new host by retromobility. Taken together with the fact that Ll.LtrB specifically targets a conserved motif of relaxase genes (Staddon *et al.*, 2004), this probably constitutes an optimal dissemination opportunity for this particular group II intron.

It would be interesting to further characterize the link between increased conjugation and intron mobility. It was reported that conjugation can be triggered by specific conditions such as bacterial stress responses (Beaber et al., 2004). We conducted some exploratory experiments to assess the effect of stress on cell clumping and intron mobility as read-outs for sex factor conjugation efficiency. We tested several stress conditions including growth at elevated temperature (37°C and 42°C), addition of ethanol, sodium chloride or hydrogen peroxide at various concentrations in liquid cultures. The tested stress conditions produced no effect on either phenotype (Belhocine K. and Cousineau B., unpublished data).

An alternative would be to use an agent that triggers sex factor conjugation to conduct a systematic study of the correlation between increased sex factor conjugation and Ll.LtrB mobility. The sex factor aggregation protein, CluA, has been well characterized and was shown to increase the levels of sex factor conjugation upon overexpression in donor cells (Godon et al., 1994; Stentz et al., 2004). Therefore, CluA is a suitable candidate to induce elevated levels of sex factor conjugation and assess its effect on Ll.LtrB mobility. Both relaxase and Ll.LtrB transcript levels could be assessed by RT-PCR and Northern blot, and intron mobility efficiency could be assessed by the use of a recipient plasmid in

the cell. Intron mobility could be monitored upon induction of high conjugation levels, both within the donor and the recipient cell following transfer. This study could corroborate the conjugation-induced increase in Ll.LtrB mobility efficiency that we previously observed. Moreover, it would exemplify an advantage to the association of group II introns with other mobile elements for their dissemination.

### 7.1.2 Conjugation to distantly related species

Conjugation is a powerful means to transfer genetic information between distantly related species and even across kingdoms (Ochman et al., 2000). Conjugation can occur between Gram-negative and Gram-positive bacteria, and even from bacteria to yeast (Ochman et al., 2000). We showed that the *L. lactis* sex factor machinery supports transfer of the pLE12 mobilizable plasmid to the closely-related Gram-positive bacterium *E. faecalis* (Chapter Two). We also demonstrated the ability of sex factor and pRS01 co-integrates to promote their transfer to this host (Chapter Four). As a logical next step, we assessed the potential of these elements to be transferred to a Gram-negative bacterium. Several mating attempts were performed between *L. lactis* and *E. coli* to transfer either the pTRK28::pRS01 co-integrate or the pLE12 plasmid, and no transfer was detected (Mandilaras V., Yeung B., Belhocine K. and Cousineau B., unpublished data). Mating was attempted on milk plates and on filter, using *E. coli* strain DH10β as a recipient. Both pTRK28 and pLE12 are shuttle plasmids that harbour an *E. coli* origin of replication (Romero and Klaenhammer, 1990;

Mills et al., 1994). Therefore maintenance of the transferred plasmid in the new host was presumably not the reason for undetected transfer.

Three *E. coli* conjugative elements, R751, ColE1 and F factor, were shown to support efficient interspecies conjugation of a mobilizable plasmid to *Saccharomyces cerevisiae* (Heinemann and Sprague, Jr., 1989). We investigated whether the sex factor conjugation machinery can support conjugation from *L. lactis* to *S. cerevisiae*, with the long-term goal to test if Ll.LtrB, present on the mobilized plasmid, can invade resident DNA in the new host. For this purpose, we engineered an *E. coli – L. lactis – S. cerevisiae* shuttle plasmid by cloning a yeast replicon into the pLE12 plasmid. Several mating attempts were performed between *L. lactis* and yeast, both on milk plates and on filters, but no transfer was detected (Belhocine K. and Cousineau B., unpublished data). Taken together, these results suggest that the *L. lactis* sex factor conjugative machinery does not promote transfer to distantly related organisms within a detectable range.

### 7.1.3 Other routes of dissemination

Our work focused on dissemination of group II introns by conjugation because of the availability of a natural system linking an active group II intron to several conjugative elements, and the biological relevance that this setting offers. It would be interesting to examine group II intron dissemination by other routes of lateral transfer.

It is suggested that conjugation and biofilm formation are two synergistic events in microbial populations: not only does biofilm formation favour conjugative transfer, but it also appears that conjugation machineries promote

biofilm formation (Ghigo, 2001; Molin and Tolker-Nielsen, 2003). Interestingly, Luo and co-workers showed that *L. lactis* strains showing a strong clumping phenotype associated with pRS01 exhibited increased ability to form biofilms. They also showed that the pRS01/sex factor clumping factor CluA facilitates biofilm formation. Finally, they observed enhanced transfer efficiency of an *E. faecalis* plasmid, pAMβ1, in *L. lactis* biofilms (Luo et al., 2005). These conditions for enhanced conjugation would be suitable to study interspecies transfer of pRS01 and the sex factor, and dissemination of the Ll.LtrB intron to multiple hosts in the context of mixed biofilms of *L. lactis* and other bacterial species.

Alternatively, it would be interesting to explore other routes of lateral transfer of group II introns. The other known means of lateral transfer between bacteria include DNA uptake by transformation and DNA transfer by phage transduction. Some bacterial species are naturally competent and prone to uptake DNA from the environment (Lorenz and Wackernagel, 1994). It would be interesting to assess if a group II intron can invade the genome of naturally competent bacterial species, e.g. *Bascillus subtilis*, upon DNA uptake. Moreover, the uptake of pre-formed intron RNPs by competent species could also be assessed. Interestingly, DNA transformation is also a route of genetic exchange that is favoured in biofilms (Molin and Tolker-Nielsen, 2003). Therefore, uptake experiments could also be conducted in a biofilm context.

It would also be interesting to assess lateral transfer of Ll.LtrB by transduction. The lactococcal intron would be an appropriate model system to use

as it can readily be tagged with markers, which allows selection of mobility events (Matsuura et al., 1997; Cousineau et al., 1998, 2000). Moreover, *L. lactis* phages are very well characterized due to the industrial importance of this bacterium (Deveau et al., 2006). It would be feasible to engineer an experimental setting to analyze Ll.LtrB transfer by phage transduction. Encapsidation of relaxase DNA containing Ll.LtrB, its transduction and further intron mobilization in the new host could be assessed using a marked intron. Moreover, it would be interesting to assess if intron RNPs can be encapsidated and transported to a different host cell.

Lateral transfer of group II introns is a widely accepted model for intron evolution (Lambowitz and Belfort, 1993). It probably played, and is still playing, a significant role in shaping bacterial and organellar genomes (Dai and Zimmerly, 2002). It is also likely at the origin of intron invasion of nuclear genomes and the rise of nuclear introns. We are only beginning to understand the mechanisms of intron lateral transfer and to appreciate its consequences on bacterial and organellar genome evolution.

## 7.2 On *trans*-splicing of group II introns and the origin of spliceosomal introns and snRNAs

Evolution of nuclear introns and the five spliceosomal snRNAs from group II intron fragmentation is a well accepted hypothesis (Cech, 1986). The discovery of fragmented group II introns that undergo *trans*-splicing provided further support to this theory (Sharp, 1991). We showed that the Ll.LtrB intron is able to *trans*-splice *in vivo* and can thus be used as a model system to study *trans*-

splicing of group II introns (Chapter Five). Subsequently, we used the lactococcal intron to investigate the potential of group II introns to *trans*-splice if fragmented at any possible location. Fragmentation sites permissive for splicing clustered between regions structurally and/or functionally analogous to snRNAs (Chapter Six). These results demonstrate that it is possible to fragment a group II intron at locations analogous to fragmentation towards the five snRNAs and maintain its *trans*-splicing capability. Therefore, our work provided novel support to the evolutionary theory linking group II introns to nuclear introns and the five snRNAs.

The experimental systems that we designed could be used to address additional questions surrounding group II intron *trans*-splicing as well as the evolutionary relationship between group II and nuclear introns.

### 7.2.1 Dependence of *trans*-splicing efficiency on the fragmentation site

In Chapter Five, we showed that Ll.LtrB *trans*-splices at different efficiencies when fragmented either within domains I, III or different locations in domain IV. The Ll.LtrB variant fragmented in domain IV downstream from *ltrA* exhibited the highest *trans*-splicing efficiency (Figure 5.5). In Chapter Six, we performed a Tn5-based genetic screen for fragmentation sites that allow *trans*-splicing. Among the fragmented variants of Ll.LtrB selected by the screen, the variants fragmented in domain I showed the best catalytic activity (Table 6.1, compare F1, F2, F3 and F7 with the remaining fragmented variants). This disparity between fragmentation sites allowing the highest *trans*-splicing activity

could be due to the numerous differences existing between the experimental settings used to conduct these two studies. Indeed, in Chapter Five, fragmentation of Ll.LtrB was achieved by cloning two independent pieces, resulting in a blunt intron fragmentation whereby the second piece begins at the nucleotide at which the first piece ends (see section 5.4.1). In contrast, fragmented Ll.LtrB variants obtained in Chapter Six were generated by the insertion of a transposon carrying a terminator followed by a promoter. Transposon insertion systematically generates a nine-nucleotide duplication (Goryshin and Reznikoff, 1998), resulting in the second intron fragment starting nine nucleotides upstream from the position where the first intron fragment ends (Chapter Six).

The site of relaxase production is also different between the two systems. Relaxase produced upon Ll.LtrB *trans*-splicing acts *in trans* in the first system, since it is produced by the pDL-$P_{23}^2$ plasmids and acts on the sex factor. In the second experimental setting, relaxase is produced by the pLE12 plasmid and acts *in cis* to initiate transfer of the mobilizable plasmid itself. It has been reported that the LtrB relaxase acts better on origins of transfer located *in cis* (Chen et al., 2005).

Finally, the transcription levels of relaxase fragments differ between the two systems. In the first system, both pieces are transcribed from two identical promoters, ensuring equal production. In the second system, the first relaxase fragment is produced from the endogenous relaxase promoter located in pLE12, while the second fragment is produced from a strong constitutive promoter located on the transposon. We have yet to determine the impact of this differential expression on the observed results.

Taken together, these differences between the two systems may explain the disparity observed between fragmentations allowing best *trans*-splicing efficiencies in the two studies.

### 7.2.2 Importance of base-pairing interactions in intron fragment recognition and assembly for *trans*-splicing

In Chapter Five, we observed that the *trans*-splicing efficiency of the Ll.LtrB variants fragmented in domain IV upstream from *ltrA* increased by 100-fold when the two intron fragments had the potential to interact by base-pairing (16 base-pairs) (Figure 5.5, compare S1 and S2). This observation suggests that base-pairing interactions play a role in the recognition and assembly of intron fragments to facilitate *trans*-splicing. The exact role of base-pairing between the two intron fragments for *trans*-splicing *in vivo* could be investigated. The splicing-conjugation system that we described in Chapter Five would be suitable to conduct a systematic mutation-complementation analysis. Ll.LtrB residues potentially implicated in base-pairing interactions between the two intron fragments could be mutated to abolish base-pairing and assess the effect on *trans*-splicing. Subsequently, compensatory mutations could be introduced in the second fragment to restore base-pairing potential and assess the effect on *trans*-splicing efficiency.

### 7.2.3 *Trans*-splicing group II introns in bacteria

So far, *trans*-splicing has only been observed in organelles and not in bacteria or archaea. This is likely due to the distribution of group II introns in

these different hosts. Indeed, group II introns often interrupt genes in organellar genomes, whereas they frequently lie in untranslated regions in bacterial genomes. Therefore, fragmented organellar introns will likely have more pressure to remain efficiently splicing and ligate the exons of the gene they interrupt. Moreover, organelles are known to undergo substantial genome rearrangements (Fauron et al., 1995; Knoop, 2004), probably resulting in a high occurrence of intron fragmentation. Interestingly, many fragments of group II introns were identified in bacteria (Dai and Zimmerly, 2002; Dai et al., 2003). Whether the remaining intron portions are found within the bacterial genomes has not been addressed. Our work provided the first experimental demonstration that *trans*-splicing can take place in a bacterial host (Chapter Five). A systematic sequence search for fragmented group II introns in bacteria could be performed. The possible candidates could be assessed for their potential to *trans*-splice. *Trans*-splicing could play a yet unknown role in gene shuffling and bacterial genome reorganization.

### 7.2.4 *Trans*-splicing of an intron fragmented into three or more pieces

We demonstrated in Chapters Five and Six that Ll.LtrB can be fragmented at multiple locations and retain splicing activity as a bipartite intron. The potential of Ll.LtrB to *trans*-splice if fragmented into three or more pieces could be assessed. Two organellar tripartite group II introns were reported to *trans*-splice, and they are both fragmented in domains I and IV (Goldschmidt-Clermont et al., 1991; Knoop et al., 1997). We constructed tripartite variants of Ll.LtrB by

combining fragmentations within domains I and IV and these variants did not *trans*-splice at a detectable level (Chapter Five). The potential of Ll.LtrB to *trans*-splice when fragmented into three or more pieces could be assessed using our transposon-based approach. One of the two fragments of a bipartite Ll.LtrB variant could be further subjected to the transposon screen. This approach will assess the possibility to introduce a second fragmentation site within the intron, thereby generating a tripartite intron. Eventually, the possibility to introduce more fragmentations could be tested, with the goal to generate a group II intron fragmented into multiple pieces capable of assembly and *trans*-splicing. If this can be achieved, it would provide a major support to the theory of fragmentation of a group II intron towards nuclear introns and snRNAs.

### 7.2.5 Group II introns, nuclear introns and snRNAs

We reported that splicing-permissive fragmentations of the Ll.LtrB group II intron were clustered between regions structurally and/or functionally equivalent to snRNAs (Chapter Six). This finding constitutes a major advance in the field of the evolution of nuclear introns and the spliceosome. It will play a significant role in our understanding of evolutionary links between group II introns and nuclear introns along with their splicing machinery.

Eventually, the evolutionary relationship between group II and nuclear introns could be further addressed by several experimental initiatives. Group II and nuclear introns share the same splicing mechanism via two transesterification reactions, the first one of which is initiated by a bulged adenosine residue located near the 3' end of the intron. It would be interesting to assess the potential of a

group II intron to be recognized and processed by the spliceosome. A group II intron could be introduced in a reporter gene in yeast and splicing could be assessed via the reporter gene activity. In the eventuality of successful splicing, it would be interesting to investigate if the intron splices autocatalytically or if it is specifically recognized and processed by the spliceosome. This could be achieved using a mutant for domain V, the group II intron catalytic domain. If the spliceosome recognizes and splices a group II intron, this would constitute an additional evidence for common evolutionary origins for nuclear and group II introns.

Interestingly, it has been demonstrated that domain V, the counterpart of the U6 snRNA, can be deleted from a group II intron and supplied *in trans* to restore splicing activity *in vitro* (Jarrell et al., 1988). This study by Jarrell and co-workers constituted a major support for the theory of group II intron fragmentation towards the five snRNAs. Another interesting experimental initiative would be the generation of a "spliceosomal-like" group II intron. The catalytic machinery located at the center of a group II intron could be removed and supplied *in trans*. If the supplied catalytic core can recognize and splice the intron *in trans*, this would constitute an intermediate state between group II introns and nuclear introns with their splicing machinery. This intermediate would be equivalent to a situation where the catalytic machinery is separated from the rest of the intron; however the "five snRNAs" would be still bound together.

## 7.3 Conclusive remarks

Our work consisted in providing experimental support to strengthen important evolutionary theories in the field of group II introns. Our work shed light on the mechanisms of group II intron evolution, from their dissemination to their evolution towards nuclear introns and snRNAs. These two aspects of group II intron evolution are intimately linked since the rise of nuclear introns was likely preceded by the invasion and the spread of group II introns in the eukaryotic nucleus. The experimental systems we designed could be used to conduct further fundamental studies to explore other aspects of group II intron evolution and gain a better understanding of the origin and the fate of these unique splicing and retromobile elements.

## 7.3 References

**Beaber, J.W., Hochhut, B. and Waldor, M.K.** (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* **427:** 72-74.

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

**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.

**Cousineau, B., Lawrence, S., Smith, D., and Belfort, M.** (2000). Retrotransposition of a bacterial group II intron. *Nature* **404:** 1018-1021.

**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.

**Dai, L., Toor, N., Olson, R., Keeping, A. and Zimmerly, S.** (2003). Database for mobile group II introns. *Nucleic Acids Res.* **31:** 424-426.

**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.

**Deveau, H., Labrie, S.J., Chopin, M.C. and Moineau, S.** (2006). Biodiversity and classification of lactococcal phages. *Appl. Environ. Microbiol.* **72:** 4338-4346.

**Fauron, C., Casper, M., Gao, Y. and Moore, B.** (1995). The maize mitochondrial genome: dynamic, yet functional. *Trends Genet.* **11:** 228-235.

**Ghigo, J.M.** (2001). Natural conjugative plasmids induce bacterial biofilm development. *Nature* **412:** 442-445.

**Godon, J.J., Jury, K., Shearman, C.A. and Gasson, M.J.** (1994). The *Lactococcus lactis* sex-factor aggregation gene cluA. *Mol. Microbiol.* **12:** 655-663.

**Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M. and Rochaix, J.D.** (1991). A small chloroplast RNA may be required for *trans*-splicing in *Chlamydomonas reinhardtii*. *Cell* **65:** 135-143.

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

**Heinemann, J.A. and Sprague, G.F. Jr.** (1989). Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340:** 205-209.

**Ichiyanagi, 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.

**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.

**Knoop, V.** (2004). The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.* **46:** 123-139.

**Knoop, V., Altwasser, M. and Brennicke, A.** (1997). A tripartite group II intron in mitochondria of an angiosperm plant. *Mol. Gen. Genet.* **255:** 269-276.

**Lambowitz, A.M. and Belfort, M.** (1993). Introns as mobile genetic elements. *Annu. Rev. Biochem.* **62:** 587-622.

**Lorenz, M.G. and Wackernagel, W.** (1994). Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol. Rev.* **58:** 563-602.

**Luo, H., Wan, K. and Wang, H.H.** (2005). High-frequency conjugation system facilitates biofilm formation and pAMbeta1 transmission by *Lactococcus lactis*. *Appl. Environ. Microbiol.* **71:** 2970-2978.

**Martinez-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.

**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.

**Mills, D.A., Choi, C.K., Dunny, G.M. and McKay, L.L.** (1994). Genetic analysis of regions of the Lactococcus lactis subsp. lactis plasmid pRS01 involved in conjugative transfer. *Appl. Environ. Microbiol.* **60:** 4413-4420.

**Molin, S. and Tolker-Nielsen, T.** (2003). Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Curr. Opin. Biotechnol.* **14:** 255-261.

**Nisa-Martinez, R., Jimenez-Zurdo, J.I., Martinez-Abarca, F., Munoz-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.

**Ochman, H., Lawrence, J.G. and Groisman, E.A.** (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405:** 299-304.

**Romero, D.A. and Klaenhammer, T.R.** (1990). Characterization of insertion sequence IS946, an Iso-ISS1 element, isolated from the conjugative lactococcal plasmid pTR2030. *J. Bacteriol.* **172:** 4151-4160.

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

**Staddon, J.H., Bryan, E.M., Manias, D.A. and Dunny, G.M.** (2004) Conserved target for group II intron insertion in relaxase genes of conjugative elements of gram-positive bacteria. *J. Bacteriol.* **186:** 2393-2401

**Stentz, R., Jury, K., Eaton, T., Parker, M., Narbad, A., Gasson, M. and Shearman, C.** (2004). Controlled expression of CluA in *Lactococcus lactis* and its role in conjugation. *Microbiology* **150:** 2503-2512.

# Appendix I

## Contribution to original knowledge

Research conducted during this thesis degree generated substantial novelty to knowledge in the fields of group II introns and bacterial conjugation.

In Chapter Two, we built the first experimental model system to investigate lateral transfer of a bacterial group II intron. We demonstrated for the first time that group II introns can be disseminated between different strains and to different bacterial species by conjugation. We showed that following conjugative transfer, a group II intron can invade new DNA sites either by retrohoming or by retrotransposition. We showed that conjugation of a non-mobilizable plasmid can also promote intron dissemination, albeit at a low efficiency. This study constituted the first experimental evidence that bacterial group II introns can be laterally transferred by the dissemination of their host mobile elements. It provided a key experimental support to the long-standing evolutionary theory suggesting that group II introns can be laterally transferred.

In Chapter Three, we studied dissemination of the Ll.LtrB group II intron by conjugation of the *L. lactis* chromosomal sex factor. We characterized retrohoming of the Ll.LtrB group II intron following transfer and integration of the sex factor into the chromosome of the recipient cell. We provided the first experimental demonstration of co-transfer of the chromosomal sex factor along

with a mobilizable plasmid by conjugation. We showed for the first time that conjugation of the *L. lactis* sex factor promotes dissemination of the Ll.LtrB group II intron between different *L. lactis* strains. We observed and characterized the competitive homing site invasion by two populations of group II introns expressed from two different sources in the same cellular environment.

In Chapter Four, we engineered the first *L. lactis* sex factor containing a marked variant of the Ll.LtrB intron. We showed for the first time that the *L. lactis* sex factor can be transferred to another bacterial species. This particular finding revises previous knowledge on interspecies conjugation of the *L. lactis* sex factor. Indeed, it had been reported that conjugation of this integrative and conjugative element to other species was impossible. Our finding reveals the potential of this *L. lactis* conjugative element to be transferred to other species. Moreover, we showed that the pRS01 plasmid can also be transferred to another bacterial species and promote efficient dissemination of the group II intron it harbours. This result broadens the extent of our previous findings. Altogether, these three studies constitute a significant contribution to understand the dissemination of group II introns in the bacterial kingdom.

In Chapter Five, we built the first *in vivo* model system to study *trans*-splicing of group II introns. In an innovative initiative, we used bacterial conjugation as a read-out for *trans*-splicing of the Ll.LtrB intron. Our system was engineered in the *L. lactis* bacterium, thereby providing the first experimental demonstration that *trans*-splicing of a group II intron can take place in bacteria. We showed for the first time that a bacterial group II intron can be independently fragmented at several locations (domains I, III and IV) and retain its splicing

activity *in trans*. We also demonstrated that the intron-encoded protein is essential for *trans*-splicing of a group II intron. This study constituted the first advance in several years in the field of *trans*-splicing. The system that we engineered will be instrumental to address multiple questions about *trans*-splicing in optimal experimental conditions.

Finally, in Chapter Six, we conducted a thorough survey of potential group II intron fragmentations that could allow *trans*-splicing. We developed and used a unique and innovative transposon-based genetic screen to generate a pool of Ll.LtrB variants fragmented at all possible locations. These variants were tested for their *trans*-splicing potential using a conjugation-based system. We demonstrated for the first time that a group II intron can be fragmented at other locations that those observed in nature, i.e. domains I, III and IV. These observations unravelled the unexpected versatility of the group II intron *trans*-splicing reaction. Moreover, the isolated fragmentation sites were clustered between regions analogous to snRNAs. Therefore, our results provided a solid experimental support to the evolutionary theory linking group II introns to nuclear introns and the five snRNAs part of their splicing machinery.