Formation of an inverted repeat junction in the transposition of insertion sequence ISLC3 isolated from *Lactobacillus casei*

Ta-Chun Lo, Hung-Wen Chen, Yu-Kuo Tsai, Yang-Cheng Kuo, Chao-Fen Lin, Ssu-Ying Kuo and Thy-Hou Lin

Correspondence

Thy-Hou Lin

thlin@life.nthu.edu.tw

Institute of Molecular Medicine and Department of Life Science, National Tsing Hua University, Hsinchu, Taiwan, ROC

Received 17 September 2007

Revised 20 December 2007

Accepted 12 January 2008

An insertion sequence, ISLC3, of 1351 bp has been isolated from *Lactobacillus casei*. Formation of IS circles containing a 3 bp spacer (complete junction) or deletion of 25 bp at the left inverted repeat (IRL) between the abutted IS ends of the ISLC3 junction region (deleted junction) was also discovered in the lactobacilli and *Escherichia coli* system studied. We found that the promoter formed by the complete junction P_{jun} was more active than that formed by the 25 bp deleted junction P_{djun} or the indigenous promoter P_{IRL} . The corresponding transcription start sites for both promoter P_{jun} and P_{IRL} as well as P_{djun} were subsequently determined using a primer extension assay. The activity of transposase OrfAB of ISLC3 was also assayed using an *in vitro* system. It was found that this transposase preferred to cleave a single DNA strand at the IRR over the IRL end in the transposition process, suggesting that attack of one end by the other was oriented from IRR to IRL.

INTRODUCTION

Insertion sequences (ISs) are mobile genetic elements (<2500 bp) flanked by inverted repeats, and their own transposition functions are generally encoded on the same sequences. They can alter the organization of prokaryotic genomes at frequencies comparable to or greater than the spontaneous mutation rates. ISs are grouped into more than 20 families based on the conservation of motifs of the putative transposase and terminal nucleotide sequences. The IS3 family is the largest group identified so far, with more than 90 members isolated from over 26 genera and 51 species (Mahillon & Chandler, 1998; Rousseau *et al.*, 2002). To date, only a few members of the IS3 family have been characterized in lactobacilli, e.g. IS1223 (Walker & Klaenhammer, 1994) in *Lactobacillus johnsoni* and IS1163 (Skaugen & Nes, 1994) and IS1520 (Skaugen & Nes, 2000) in *Lactobacillus sakei*. Members of this family are characterized by containing two overlapping open reading frames, namely *orfA* and *orfB*, leading to the fusion protein OrfAB produced by programmed translational frameshift between *orfA* and *orfB*. A DD(35)E catalytic motif has been identified at the C-terminal region of OrfAB (Fayet *et al.*, 1990). The IS3 family elements are produced by transposition in a simple cut-and-paste manner involving several steps. In the first step of transposition, the transposase OrfAB cleaves a single DNA

strand at one IS end to generate a free 3'-OH group and then transfers this end to an open end of the same strand which is several nucleotides distal from the other IS end (Haas & Rak, 2002; Sekine *et al.*, 1999). A 'figure of eight' is formed by the IS sequence following repair by the host repair machinery and replication by the IS itself at the junction region (Turlan & Chandler, 1995). In the second step of transposition, the figure of eight is processed into a transposon circle. However, the left inverted repeat (IRL) and right inverted repeat (IRR) junction is created by either intra-IS or occasionally inter-IS recombination to form a head-to-tail IS tandem dimer (Szeverenyi *et al.*, 2003; Ton-Hoang *et al.*, 1998).

The expression of transposase is controlled by an original promoter P_{IRL} upstream of *orfA*. However, in several cases of the IS3 family, an inwardly directed -10 hexamer has been detected in the IRL (Galas & Chandler, 1989). New promoters, such as P_{jun} , may be generated by a combination of the -10 hexamer in the IRL with a -35 hexamer in the neighbouring IRR when two ends from an IS are juxtaposed as head-to-tail dimers or circles. Compared with the original P_{IRL} , this new promoter is a relatively strong one in some members of IS3 family, e.g. IS21 (Reimann, *et al.*, 1989), IS30 (Dalrymple, 1987) and IS911 (Ton-Hoang *et al.*, 1997), which leads to high transposase expression and consequently an increase in the transposition activity. Moreover, Duval-Valentin *et al.* (2001) demonstrated that inactivation of P_{jun} strongly decreases IS911 transposition when the corresponding

Abbreviations: DR, direct repeat; 6-FAM, 6-carboxyfluorescein; IPCR, inverse PCR; IRL, left inverted repeat; IRR, right inverted repeat; IS, insertion sequence.

transposase is produced by its natural configuration P_{IRL} . Transpositionally active junctions have been reported experimentally for several other IS families, e.g. IS21 (Reimann & Haas, 1990), IS30 (Kiss & Olsz, 1999; Olsz *et al.*, 1993) and IS1 (Shiga *et al.*, 1999; Ton-Hoang *et al.*, 1997).

In previous studies, we have shown that ISLC3 (GenBank accession no. AF445084) is inserted into frame *orf14* of the temperate bacteriophage ϕ AT3 of *Lactobacillus casei* ATCC 393, resulting in the insertion of a premature stop codon (Lo *et al.*, 2005). This insertion causes a deletion of 36 amino acids in Orf14 from the C-terminus of the original protein. Sequence analysis of ISLC3 has revealed that this DNA element belongs to the IS3 family. A high level of sequence similarity is also found between ISLC3 and the putative mobile elements of *Lactobacillus casei* ATCC 334 (accession no. CP000423, 99% identity) (Makarova *et al.*, 2006), *Lactobacillus coryniformis* DSM 20001^T (accession no. AJ605769, 99% identity), *Pediococcus pentosaceus* ATCC 25745 (accession no. CP000422, 98% identity) and IS153 of *Lactobacillus sanfranciscensis* (accession no. AJ239042, 82.1% identity) (Ehrmann & Vogel, 2001). The total length of ISLC3 has been determined as 1351 bp, and 37 bp IRs are found at each end. Moreover, the flanking sequence of ISLC3 has been determined as a 3 bp (5'-ACC-3') direct repeat (DR) on its target sequence. There are two open reading frames, *orfA* (from 82 to 333 bp) and *orfB* (from 387 to 1229 bp) identified on the ISLC3 sequence, which are in phase 0 and -1, respectively. A conserved DD(35)E triad has also been identified in the middle of the OrfB sequence, suggesting the existence of a catalytic domain on the sequence (Rousseau *et al.*, 2002).

In this study, we investigated the excision process of the ISLC3 transposition reaction. We found that an *in vivo* genetic rearrangement has occurred at the IRL and IRR of ISLC3 which results in the generation of circles or head-to-tail tandem dimers containing a 3 bp spacer or a deletion of 25 bp at the IRL between the abutted IS ends of the junction region. We also determined the transcription start sites for P_{jun} and P_{IRL} in several lactobacilli by using a primer extension assay. An extra transcription start site was detected by the same assay as well, indicating that a potential promoter is located in the proximity of P_{jun} or P_{IRL} . In other primer extension experiments where the complete ISLC3 sequence was cloned into *Escherichia coli* using plasmid pSA1, we found that this potential promoter was formed by the 25 bp deleted IRL (IRL Δ 25) of the inverted repeat junction sequence, namely, P_{djun} . We also found via an *in vitro* transposition assay that the transposase OrfAB of ISLC3 cleaved a single DNA strand at the IRR but not IRL end in the first step of the transposition process. Moreover, OrfAB activity was also found to be strongly enhanced in the presence of Mn^{2+} and Mg^{2+} .

METHODS

Bacterial strains and culture conditions. The bacterial strains and plasmids used in this study are listed in Table 1. All the lactobacilli

used in the study were obtained from the Culture Collection and Research Center (CCRC) of Taiwan and were maintained at -80°C and subcultured on MRS (Difco) agar plates or in MRS broth anaerobically at 37°C . The *E. coli* host used in the cloning procedures was DH5 α , which was grown in Luria-Bertani (LB) broth at 37°C with agitation. Antibiotics were added at the following concentrations: kanamycin (Km), $50\text{ }\mu\text{g ml}^{-1}$; chloramphenicol (Cm), $50\text{ }\mu\text{g ml}^{-1}$; and ampicillin (Amp), $50\text{ }\mu\text{g ml}^{-1}$. The concentration of X-Gal used in all the cloning procedures was $20\text{ }\mu\text{g ml}^{-1}$. Unless otherwise stated, all the chemicals used were purchased from Sigma. Klenow DNA polymerase, T4 DNA ligase and all restriction enzymes used were purchased from New England Biolabs (NEB) and the manufacturer's instructions were followed.

Plasmid construction and PCR assay. PCR amplifications were performed with 2 units *Pfu* DNA polymerase (Promega) or 1 unit *ext-Taq* DNA polymerase (TakaRa), 25 pmol oligonucleotide primers, $1\text{ }\mu\text{g}$ chromosomal (or plasmid) DNA, $200\text{ }\mu\text{M}$ deoxynucleoside triphosphates and 2.5 mM MgCl_2 in each reaction buffer as recommended by the manufacturer. Each PCR was cycled 30 times through the following conditions: denaturation at 94°C for 2 min, annealing at 55°C (or 2.5°C less than the T_m of the oligonucleotide used) for 1 min, and extension at 72°C for 30 s to 2 min. All the PCR products were purified by phenol extraction and then ethanol precipitation. The ligation reactions were incubated at 16°C for 8 h with T4 DNA ligase (NEB). The resulting ligation mixtures were then used to transform *E. coli* DH5 α and were plated on LB agar containing appropriate antibiotics. All the DNA sequences were determined by Mission Biotech (Taiwan) using ABI377-96 automated sequencers (Applied Biosystems). The BLAST network service of the National Center for Biotechnology Information (NCBI) was used to compare the homologous sequences searched from the GenBank/EMBL/DBJ and SWISS-PROT databases. The ISLC3 DNA fragment was isolated from the temperate bacteriophage ϕ AT3 (GenBank accession no. AY605066) of *L. casei* ATCC 393 and was subcloned into vector pBluescriptSK⁺ (pSK⁺) (Stratagene) to generate plasmid p2END. Plasmid pSA1 was constructed as follows: a *Hind*III fragment carrying the intact ISLC3 sequence was amplified by PCR from plasmid p2END using primer Hin-R (5'-GCAAGCTTTGTGGCTAATACCAAACTTATTTCAA-3'; 15054–15080 bp of phage ϕ AT3 genome) and Hin-L (5'-GCAAGCTTGATGGCATGAGCTGAATGATG-3'; 16525–16529 bp of phage ϕ AT3 genome). The artificial *Hind*III site on both 5' ends of the primers is shown in italics. The PCR fragment obtained was cloned into the *Hind*III site of vector pACYC184 or directly into vector pGEM-T-Easy (Promega) for sequencing. The selection markers were the chloramphenicol (Cm^r) or ampicillin (Amp^r) resistance genes in *E. coli* DH5 α and the resultant plasmids were designated pSA1 or pGIS, respectively. The alkaline lysis method described by Sambrook *et al.* (1989) was used for the extraction of plasmid DNA. The QIAquick Gel Extraction kit (Qiagen) was used to extract and purify DNA following separation by the agarose gel electrophoresis. Two ISLC3 complementary primers, 149out (5'-GCTGATTGCGCTTGCTTATATAGG-3'; 149–126 bp of ISLC3) and 999out (5'-GCATGCTTTTACAATAGGAAGAG-3'; 1130–1153 bp of ISLC3), were used in inverse PCR (IPCR) to amplify the IS inserts from cell lysates of *L. casei* ATCC 393, *L. casei* ATCC 334, *L. casei* ATCC 27139 and *E. coli* DH5 α (pSA1), and the PCR products were cloned into vector pGEM-T-Easy for sequencing. The uncut chromosomal DNA of lactobacilli or plasmids was separated by agarose gel electrophoresis and purified using the QIAquick Gel Extraction kit. DNA molecules between 1000 and 4000 bp were designated part A and those smaller than 1000 bp were designated part B. Both parts of DNA were then utilized in IPCR or Southern blot assay.

Plasmid pTL, containing a truncated IRL end, used in the *in vitro* activity assay for OrfAB, was made by IPCR using the following primers: 5'-GGGTATAATGAATAAATCGTCTC-3' (38–60 bp of

Table 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Reference or source
Strains		
<i>E. coli</i> DH5 α	F ⁻ <i>hsdR17 recA1 supE44 endA1 gyrA96 relA1</i> Δ (<i>argF-lacZYA</i>)U169 <i>thi-1</i> ϕ 80d <i>lacZ</i> Δ M15	BRL
<i>E. coli</i> BL21(DE3)	F ⁻ <i>ompT hsdSB</i> (<i>r_B⁻ m_B⁻</i>) <i>gal dcm</i> λ (DE3)	Novagen
<i>Lactobacillus casei</i>	ATCC 393	ATCC*
<i>Lactobacillus casei</i>	ATCC 393NT, mitomycin-C resistant strain	This study
<i>Lactobacillus casei</i>	ATCC 334	ATCC
<i>Lactobacillus casei</i>	ATCC 27139	ATCC
<i>Lactobacillus rhamnosus</i> GG	ATCC 53103	ATCC
<i>Lactobacillus sanfranciscensis</i>	ATCC 27651	ATCC
<i>Lactobacillus delbrueckii</i> subsp. <i>bulgaricus</i>	HER1308	Peake & Stanley (1978)
Plasmids		
pACYC184	Cloning vector, Cm ^r , Tc ^r	New England Biolabs
pGEM-T-Easy	TA Cloning vector, Ap ^r , CoE1 ori	Promega
pBlue-TOPO	LacZ reporter vector, Ap ^r , pUC ori	Invitrogen
pBluescriptSK ⁺	Cloning vector, Ap ^r	Stratagene
p2END	Contains wild-type ISLC3 with <i>Hind</i> III site in pBluescriptSK ⁺	This study
pSA1	Contains wild-type ISLC3 in pACYC184, p15A ori, <i>bla</i> ⁺	This study
pGIS	Contains wild-type ISLC3 in pGEM-T Easy	This study
pTR	Truncated IRR end in pGIS	This study
pTL	Truncated IRL end in pGIS	This study
pEAB	Expression of OrfAB in pET101 directional TOPO Expression kits	This study
pRDRL2	Carrying the complete 5'-IRR-DR(ACC)-IRL-3' fragment in pGEM-T Easy	This study
pIRLz	Contains promoters P _{IRL} to <i>lacZ</i> gene in pBlue-TOPO	This study
pDSLz	Contains promoters P _{djun} to <i>lacZ</i> gene in pBlue-TOPO	This study
pACLz	Contains promoters P _{jun} to <i>lacZ</i> gene in pBlue-TOPO	This study
pUV5	Control promoter lacUV5 in pBlue-TOPO	This study

*ATCC, American Type Culture Collection.

ISLC3) and 5'-TGCCTAATGTATGCAAAGGGCGATC-3' (15115–15091 bp of ϕ AT3), on plasmid pGIS as a template. Similarly, the truncated-IRR plasmid pTR was made by PCR from plasmid pGIS using the primers 5'-CATTGAGAACCAGCGTTTACCTTCC-3' (1314–1291 bp of ISLC3) and 5'-GTGGGACAAGTTTCTTTTGCCACCC-3' (16474–16498 bp of ϕ AT3). The site-directed mutagenesis kit Phusion (Finnzymes) was used in these experiments.

Southern blotting. Total cellular DNA was extracted from the lactobacilli by treating them with SDS/proteinase K and the DNA was phenol/chloroform treated and then precipitated with ethanol as described by Forsman & Alatossavam (1994). For Southern hybridization analysis, equal amounts of chromosomal DNA from each strain were digested with *Eco*RV and subjected to electrophoresis on a 1% agarose gel. The DNA fragments were transferred to a Hybond-N⁺ membrane (Amersham) and cross-linked by a UV cross-linker (Spectronics). The membrane was pre-hybridized with 0.5 ml of a 1 mg ml⁻¹ solution of sonicated salmon sperm DNA at 42 °C for 1 h. The DNA fragments amplified by PCR from plasmid p2END using primers ISR50 and ISL50 were used as the probe (labelled with [α -³²P]dATP). The membrane was hybridized with the labelled probe at 42 °C for 12 h. After hybridization, the membrane was washed under high-stringency conditions: 30 min with 2 $\times$ SSC/0.1% SDS (1 $\times$ SSC is 150 mM NaCl plus 15 mM sodium citrate) followed by another 30 min with 0.5 $\times$ SSC/0.1% SDS. Both washes were at room temperature and were followed by a final 30 min wash at 55 °C with 0.1 $\times$ SSC/0.1% SDS.

RNA extraction and primer extension. The lactobacilli were cultivated in 40 ml MRS broth at 37 °C. Bacteria were collected at the mid-exponential phase and rapidly pulverized in liquid nitrogen. The total RNA was isolated by using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. The RNA pellets obtained were suspended in diethyl pyrocarbonate (DEPC) water treated with RQ1 RNase-free DNase (Promega). The RNA was recovered by extraction with TRIzol reagent again. The quantity of RNA was determined by spectrophotometry (*A*₂₆₀ and *A*₂₈₀). Analysis of the 5' ends of the OrfA mRNA transcripts was performed by primer extension using 6-FAM-labelled primers (5'-TTGAATTCCTT-TGTCGTAACGA-3'; 110–190 bp of ISLC3) (Merighi *et al.*, 2006). A total of 100 pmol primer was annealed to 20 μ g total RNA. Synthesis of cDNA was performed at 50 °C using the SuperScript III reverse transcriptase (Invitrogen) according to the manufacturer's instructions. These first-strand cDNAs were analysed in an ABI 3730 capillary electrophoresis sequencer and the corresponding fluorescence intensity was quantified by the GeneMapper V3.7 software (Applied Biosystems).

Circle junction promoter activity assay. Sequences for promoters P_{IRL}, P_{djun} (the 25 bp truncated junction promoter) and P_{jun} (the complete junction promoter) were obtained by the following respective PCR reactions: (i) primer ISL (5'-TGGATGTAAGT-TGATTCCTGAGA-3'; 1–23 bp of ISLC3) and HiaR (5'-AAA-AATTCCTTCCTTTTGTAG-3'; 79–60 bp of ISLC3) on plasmid pSA1; (ii) primer 999out and HiaR on the genomic DNA of *L. casei*

ATCC 393; and (iii) primer 999out and HiaR on plasmid pRDRL2 as the template DNA. Plasmid pRDRL2 carrying the complete 5'-IRR-DR(ACC)-IRL-3' fragment was obtained from one of the subclones of the IPCR assay on the *E. coli* DH5 α (pSA1) cell lysates. These PCR fragments were cloned into a TA vector, the pBlue-TOPO vector (Invitrogen), and then fused with the *lacZ* gene to generate plasmids pIRLz, pDSLz and pACLz, respectively. The lacUV5 promoter on plasmid pUV5 was made by using the oligonucleotides UV5+ (5'-ACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGAA-3') and UV5- (5'-TCCACACATTATACGAGCCGGAAGC-ATAAAGTGTAAGCCTGGGGTA-3') via annealing and then direct TA cloning on vector pBlue-TOPO. The P_{UV5} -35 and -10 hexamers are shown in bold in the sequence of primer UV5+. Each of these plasmids was used to transform *E. coli* DH5 α (RecA⁻) and the transformed cells were grown at 37 °C overnight in LB medium supplemented with appropriate antibiotic(s). The overnight culture was diluted 1:50 with fresh medium and grown until the mid-exponential phase (OD₆₀₀ 0.4–0.8) was reached. The cells were collected by centrifugation and lysed; then the β -galactosidase activity was measured as A₄₂₀, as described by Miller (1992). The β -galactosidase activity measured was converted to Miller units using the following equation: β -galactosidase activity (Miller units) = (A₄₂₀ × 1000)/(OD₆₀₀ × t × v), where t is the incubation time (min) and v is the volume of culture (ml).

Protein expression and purification. The site-directed mutagenesis for the -1 frameshift region upstream of the *orfA* stop codon was performed by IPCR on the entire *orfA* and *orfB* genes on plasmid pGIS with the 5' primer 5'-CTTCTTGCGCAGAAGTAACCAATA-3' (315–339 bp of ISLC3) and 3' primer 5'-CCTAGGGCTTTTTT-TAGGATTTCGTT-3' (289–314 bp of ISLC3), where the latter was complementary to the 3' end of the *orfA* gene including an extra T base (bold) of the *orfA* gene. The site-directed mutagenesis kit Phusion (Finnzymes) was used for this experiment. The mutant plasmid product was purified and used as the PCR template for testing the expression of OrfAB using the Champion pET101 directional TOPO Expression kits (Invitrogen). Overexpression of OrfAB (1–382 aa) was achieved by IPTG (Boehringer Mannheim) induction of an exponentially growing culture of *E. coli* BL21(DE3) transformed with plasmid pET101/D-TOPO(OrfAB). The OrfAB derivatives were partially purified using the C-terminal histidine tag as indicated by the manufacturer. Induction was achieved by the addition of IPTG to a final concentration of 1 mM and the cells were incubated for a further 90 min before being harvested and then lysed by sonication. The protein purification was performed at 4 °C as follows: the crude extracts (the total protein concentration was about 10 mg ml⁻¹) were subjected to centrifugation at 8000 g for 10 min and then the overexpressed OrfAB protein was purified by a Ni-NTA resin column.

In vitro activity assay for OrfAB. Typical activity assays for OrfAB (50 μ l) contained 500 ng transposase OrfAB and 200–500 ng plasmids and were performed in the following cocktail: 20 mM HEPES pH 7.5, 5 mM DTT, 50 mM MgCl₂, 200 mM KCl and 10% (v/v) glycerol. Assays were incubated at 37 °C. The concentrations of metal ions used are given in the Results. The products of all the *in vitro* assays were treated with 0.5 mg proteinase K ml⁻¹ for 30 min at 50 °C before being purified with Qiaquick purification minicolumns (Qiagen).

RESULTS

Detection of the genetic rearrangement of ISLC3 in lactobacilli

The distribution and copy number of ISLC3 in *L. casei* ATCC 393, *L. fermentum* ATCC 14931, *L. sanfranciscensis* ATCC

27651, *L. casei* subsp. *rhamnosus* ATCC 14957, *L. casei* ATCC 27139, and *L. delbrueckii* subsp. *bulgaricus* HER1308 were determined by Southern hybridization analyses. The amount of genomic DNA extracted for each strain was 3 μ g; it was digested with *EcoRV* and *HindIII*, which have one and no site on the ISLC3 sequence, respectively. The Southern hybridization result indicated that more than four copies of ISLC3 were harboured by *L. casei* ATCC 393 while the other lactobacilli harboured one to three copies (data not shown). An IPCR using two outward-oriented IS-specific oligonucleotides, 149out and 999out (Fig. 1a, b), was conducted to examine whether the free circular form or head-to-tail tandem copies were present extrachromosomally in *L. casei* ATCC 393, where the presence of ISLC3 was confirmed by IPCR analysis. A small fragment of nearly 350 bp was detected by IPCR (Fig. 1a), suggesting that ISLC3 may circularize itself as observed in other IS elements identified from several Gram-positive bacteria, such as the flanking IS256 copies of Tn4001 tandem dimer in *Staphylococcus aureus* (Prudhomme *et al.*, 2002). In fact, a fragment of the same size was also detected by IPCR using the same primer set on the extracted chromosomes of *L. casei* ATCC 27139, *L. sanfranciscensis* ATCC 27651 and *L. delbrueckii* subsp. *bulgaricus* HER1308. The binding sites of the 149out and 999out primers were 98 bp or 161 bp distal from the ends of ISLC3 (Fig. 2). Through sequencing on the subcloned fragment on plasmid pGEM-T-Easy, we obtained some IPCR products that varied around 342 bp in the lactobacilli studied. Furthermore, the difference in such sequences detected for the complete 370 bp circles in *L. sanfranciscensis* ATCC 27651 or *L. delbrueckii* subsp. *bulgaricus* HER1308 was 2 or 5 bp respectively. The sequence of the suspected complete circle formed by ISLC3 was detected as a fragment of 370 bp as shown in Fig. 2. The truncated circle of the 342 bp fragment is also shown.

Since the binding sites of the 149out and 999out primers were 149 bp and 218 bp distal from both ends of ISLC3, the size of the IPCR product deduced was 370 bp, which included the 3 bp linker sequence and IRs. This 3 bp linker sequence was derived from the flanking sequences for ISLC3 in the parental bacteriophage ϕ AT3 genome. The 3 bp is one of the two copies of the 3 bp DR(5'-ACC-3') target sequence produced in the initial insertion event. However, the complete 5'-IRR-DR-IRL-3' junction sequence characterized by a DR at the insertion site by ISLC3 such as that in the bacteriophage ϕ AT3 genome was not detected from the IPCR products of the *L. casei* ATCC 393 genome or other DR-separated IRs of the junction regions of *L. casei* ATCC 27139, *L. sanfranciscensis* ATCC 27651 and *L. delbrueckii* subsp. *bulgaricus* HER1308. Instead, the IPCR product was detected as circles or head-to-tail tandem dimers where 25 bp from the IRL junction region was found to be deleted (5'-IRR-IRLA25-3').

Detection of ISLC3 circles in *E. coli*

It is difficult to detect the original circular forms in lactobacilli studied using the Southern hybridization

Fig. 1. (a) The presence of ISLC3 circles in several lactobacilli was investigated using an IPCR assay with primers 149out and 999out. The IPCR products were separated by electrophoresis. Lanes 1, 2, 3 and 4 contain DNA samples from *L. casei* ATCC 393, *L. casei* ATCC 27139, *L. casei* ATCC 393NT and *L. rhamnosus* GG, respectively. Lane M contains a 100 bp ladder size maker. (b) Plasmid pSA1, constructed to carry the wild-type ISLC3. Primers ISR50, ISL50, 149out and 999out are indicated by arrows.

typ. cir.	5'-gcttttcttca	ataggaagag	AATTCATAGT	TCACTGGGCT	ACCAGACCCC	CtTaCAAGTT
<i>L. s.</i>	-----	-----	-----	-----	-----	-c-t-----
<i>L. d.</i>	-----	-----	-----	-----	-----	-c-t-----
GAAATTGCAA	CACttACGAG	CCAAATGGCC	GCCTGACTTA	ATGcttTTCCA	GGGTTCAAAT	AAGTTATTAA
-----	-----	-----	-----	-----	-----	-----
-----	-----	-----	-----	-----	-----	-----
TCGCATATAA	TGATTTATTT	GAGCTGTGGA	AGGTAAACGC	GGTTCTGAAT	GTCTtTTAAA	ATCTGTCTCA
-----	-----	-----	-----	-----	-----	-----
-----	-----	-----	-----	-----	-----	-----
*****	*****	***	***	***	***	***
AATATTGACT	TCAATCCAAC	CTGGATGTAA	GCTGATTC	CT GAGACAACTT	TTAAGAGAGG	G[TATAAT]GAA
-----	-----	-----	-----	-----	-----	-----
Δ25 bp cir.	TATTGACTTC	AATCCAACCTT	TTAAGAGAGG	-----	-----	-----
-----	-----	-----	-----	-----	-----	-----
TAAATCGtCT	CTCAAAAGGA	AGGAATTTTT	ACATGACAAG	TCGTTACGAC	AAAGAATTCA	AACAAAACAT
-----	-----	-----	-----	-----	-----	-----
-----	-----	-----	-----	-----	-----	-----
TATCAAccta	tataagcaag	gcgaatcagc-3'	-----	-----	-----	-----
-----	-----	-----	-----	-----	-----	-----
-----	-----	-----	-----	-----	-----	-----

Fig. 2. Sequence of the 370 bp complete circle (typ. cir.) formed by ISLC3. The sequence consists of primer 999out (lower-case letters), 98 bp flanking sequence, IRR (underlined), DR (5'-ACC-3', italic), IRL (underlined), 61 bp flanking sequence and primer 149out (lower-case letters). The formation of a truncated circle (Δ25 bp cir.) in which DR plus 25 bp in IRL (bold letters) were deleted is highlighted with shadow. The indigenous promoter (P_{IRL}) comprises -10 (TATAAT) and -35 (CTGAGA) regions (boxed) while the truncated one (P_{jun}) contains a new -10 (TATTGA) and -35 (TGTCTC) region (dashed-line boxes). The typical promoter formed by the complete circle at the -10 (TAAGCT) and -35 (TTGACT) regions is also marked with asterisks. Differences in the sequences from *L. sanfranciscensis* ATCC 27651 (*L. s.*) or *L. delbrueckii* subsp. *bulgaricus* HER1308 (*L. d.*) are shown with bold letters.

technique due to low copy number. Therefore, we employed a technique which was used by Perkins-Balding *et al.* (1999) to detect the IS492 circle. The size of digested DNA between 1000 to 4000 bp (part A) was separated from that below 1000 bp (part B) in an agarose gel and each part of DNA was purified using the QIAquick Gel Extraction kit. The DNA sample extracted from each gel slice (parts A and B) was utilized in an IPCR assay (Fig. 3a). Plasmid pSA1 was constructed from pACYC184 to carry a complete wild-type ISLC3 sequence flanked respectively by 50 bp of IRs of *orf14* target sequence on the left and right ends (Fig. 3a). This constructed vector was used to transform *E. coli* DH5 α (RecA⁻) and the corresponding plasmid DNA was extracted using the alkaline lysis method. The DNA samples were separated into two parts by electrophoresis and purified using the QIAquick Gel Extraction kit as described (Fig. 3a, parts A and B). To confirm the formation of the circular form, plasmid DNA was digested with *Pst*I, since the restriction site was located

right in the middle of the ISLC3 sequence. Without the presence of intact plasmid pSA1, the amount of circular forms may be enriched by the separation and collection procedures described in Fig. 3(a). The *Pst*I-digested DNA sample of part A (Fig. 3a) was analysed by Southern hybridization using the ISLC3-specific probe. Detection of a signal of 1.3 kb (lane 1 of Fig. 3b) confirms the presence of the circular form of ISLC3 in *E. coli*. The 5.5 kb signal was the intact plasmid pSA1 leaked out from part A during the electrophoresis process (lane 1 of Fig. 3b). However, both signals (the ISLC3 circles and intact pSA1) were much weaker than those of the undigested pSA1 fraction (lane 2 of Fig. 3b) or *L. casei* ATCC 393 chromosomal DNA (lane 3 of Fig. 3b). The formation of ISLC3 circles was further examined at the circle junction region using IPCR with primer 149out and 999out for DNA samples collected in part A and B (Fig. 3a), respectively. The presence of circle junction in parts B (lane 4 of Fig. 3b) and A (lane 5 of Fig. 3b) was evident and this further confirmed the

Fig. 3. (a) Technique used to identify the ISLC3 circular forms. Slices were cut from the electrophoresis gel based on the DNA molecular size standards (1 kb ladder) marked. DNA samples of these slices were used as the template DNAs for PCR or Southern hybridization analysis to detect the ISLC3 circle junction. The sizes of PCR products obtained from the part A and B gel slices were 4 to 1.0 kb and <1.0 kb, respectively. (b) Detection of the ISLC3 circular forms by ISLC3-specific Southern hybridization on extrachromosomal DNA extracted from *E. coli* DH5 α (pSA1) and *L. casei* ATCC 393. Lane 1, *Pst*I-digested fraction from part A eluted in (a); the signal of 1.3 kb is indicated by an arrow. Lane 2, undigested plasmid fraction of pSA1; lane 3, undigested *L. casei* ATCC 393 chromosomal DNA. Lanes 4–6, formation of circles in *E. coli* DH5 α (pSA1) and *L. casei* ATCC 393 chromosome detected by PCR: lane 4, from part B eluted in (a); lane 5, from part A eluted in (a); lane 6, undigested *L. casei* ATCC 393 chromosome. Lane M is a 100 bp ladder size marker.

presence of ISLC3 circles in *E. coli*. The IPCR products collected from both parts were subcloned and sequenced. It was found that nearly 10% (2 out of 21 independent clones) of DNA samples sequenced has the complete 5'-IRR-DR-IRL-3' junction region, and a deletion of 25 bp including the DR of the junction region (IRR-IRL Δ 25) was also found.

Assembly of a new promoter by formation of a circle junction

Studies on the transposition of IS2, IS911 and IS492 have revealed that these IS circles are derived by joining the IRR and IRL ends (Polard *et al.*, 1992). It is also known that a promoter generated by the joining reaction will control the expression of the transposase. Some promoter-like sequences resembling the typical *E. coli* σ^{70} -dependent promoter sequences, namely TTGACA of the -35 element, and TATAAT spaced by 16–18 bp of the -10 element (Harley & Reynolds, 1987), were found by analysing the transcription regions of the complete 5'-IRR-DR(ACC)-IRL-3' junction region of ISLC3 (Fig. 4a) (Burr *et al.*, 2000). Some primer extension experiments were conducted to analyse the 5' ends of the OrfA mRNA transcripts prepared from *L. casei* ATCC 393, ATCC 334, ATCC 27139 and *E. coli*(pDSLz), respectively. The method was used to map the transcription start sites for the OrfA mRNA extracted from these bacteria. Using 6-FAM end-labelled oligonucleotide primers and the total RNA extracted for these lactobacilli grown at 37 °C, three reverse-transcribed DNA fragments, of 55 bp, 80 bp and 90 bp, were detected for each. We found that the transcription start site of the 90 bp fragment of the complete 5'-IRR-DR-IRL-3' junction region of ISLC3 was at the guanine base 61 bp upstream of the OrfA translation start codon (Fig. 4a). This finding confirmed our conjecture that a putative σ^{70} promoter (TTGACT-N₁₇-TAAGCT) including an untypical extended TGn -10 motif (5'-TGTAAGCT-3', where an additional nucleotide 'n' is presumably missing) was generated at the 5'-IRR-DR-IRL-3' junction region of ISLC3 (Fig. 4a) (Keilty & Rosenberg, 1987; Ponnambalam *et al.*, 1986; Sabelnikov *et al.*, 1995). In other words, the transcripts made from this promoter started at the guanine base which was eight bases downstream from the -10 hexamer (5'-TAAGCT-3') of the predicted junction promoter P_{jun} while the -35 hexamer (5'-TTGACT-3') was initiated within the IRR. The transcription start site for the 55 bp reverse-transcribed DNA fragment was 10 bases downstream from a potential -10 hexamer (5'-TATAAT-3') at the guanine base of the indigenous OrfA promoter P_{IRL} (Fig. 4a). The corresponding fluorescence intensity of the 90 bp fragment of P_{jun} was about 2.5-fold stronger than that of the 55 bp fragment of P_{IRL} (Fig. 4a). However, the transcription start site for the 80 bp reverse-transcribed DNA fragment could not be unambiguously defined. We suspected that this transcript could be made from an alternative transcription start site of promoter P_{jun} or it could be produced by a new promoter formed by 25 bp

deletion of IRL (IRL Δ 25) at the junction region, namely P_{djun}. Therefore, we performed another primer extension experiment using the same 6-FAM end-labelled oligonucleotide primers on the total RNA extracted from *E. coli* transformed with plasmid pDSLz, where the 5'-IRR-IRL Δ 25-3' fragment was ligated with the *lacZ* gene on a pBlue-TOPO vector. As shown in Fig. 4(a), a cDNA fragment of nearly 80 bp was generated in the *E. coli*-based system. However, the fluorescence intensity of transcripts made in *E. coli* for the 80 bp fragment was only about one-tenth that of the 90 bp reverse-transcribed DNA fragment made in lactobacilli (Fig. 4a). These *E. coli* transcripts detected confirmed the existence of promoter P_{djun}. This putative promoter P_{djun} (TGTCTC-N₁₇-TATTGA) made at the 5'-IRR-IRL Δ 25-3' region of ISLC3 is a non-typical one since it lacks a recognizable -35 region and its -10 hexamer is far from complete since it relies on an extended and TGn -10 motif. The transcription start site of this unusual promoter was determined to be at the thymine base 16 bases downstream from the -10 hexamer of the promoter (Fig. 4b).

We also constructed plasmids pIRLz, pDSLz and pACLz, which were pBlue-TOPO-based vectors made by ligating the sequences of the indigenous P_{IRL}, P_{djun} and P_{jun} promoters with the *lacZ* gene (Knudtson & Minion, 1994). All these plasmids plus plasmid pUV5, where a lacUV5 constitutive promoter was used to drive the expression of β -galactosidase, were transformed into *E. coli* DH5 α (RecA⁻). The transformed bacteria were incubated in LB with the addition of X-Gal so that the activity of β -galactosidase could be measured as A₄₂₀. The β -galactosidase activities measured for promoters P_{jun}, P_{IRL} and P_{djun} were 3.5-, 0.7- and 0.5-fold of that of promoter lacUV5 (data not shown). This result indicated that the promoter activity of P_{jun} was much stronger than that of both P_{IRL} and P_{djun} in *E. coli*.

OrfAB activity *in vitro*

To examine the activity of OrfAB *in vitro*, plasmid pGIS carrying the wild-type ISLC3 and plasmid pTL or pTR carrying the truncated IRL or IRR ends, respectively, were incubated with a highly enriched fusion protein OrfAB expressed by plasmid pEAB. The protein was made as an ISLC3 derivative by inserting 1 bp in the sequence of the A₇G frameshifting site between *orfA* and *orfB*, which rendered the transposase gene *orfAB* in-frame. The *in vitro* experiments performed by incubating OrfAB and plasmid pGIS together in the presence of various divalent metal ions revealed that the protein was able to linearize pGIS to give a 4.5 kb product (Fig. 5a and b). Some sparsely present high-molecular-mass species were recognized as the autointegration products (Fig. 5a, b) (Ton-Hoang *et al.*, 1998). As shown in Fig. 5(c) by a densitometer tracing of the gel images, about 85% of pGIS was influenced by OrfAB while the autointegration products were gradually increased to nearly 20% within the 30 min incubation with

Fig. 4. (a) Primer extension experiments to confirm the formation of promoters at the circle junction. Locations of the transcription start sites of the *orfA* gene of *L. casei* ATCC 393, ATCC 27139 and ATCC 334 were mapped by the reverse-transcribed DNA fragments with fluorescent 6-FAM-labelled specific primers (underlined by arrows) and quantified by the GeneMapper software on the reverse-transcribed DNA fragments, and the results were displayed as electrophoretograms. Three reverse-transcribed DNA fragments, of 55 bp, 80 bp and 90 bp, were detected in each case. The putative promoter P_{jun} (TTGACT-N₁₇-TAAGCT) at the 5'-IRR-DR(ACC)-IRL-3' junction region is marked and underlined; the indigenous *orfA* promoter P_{IRL} (CTGAGA-N₁₇-TATAAT) is also highlighted. The 17 bp spacer is underlined in bold. The transcription start sites determined were 55 bp G, 80 bp T and 90 bp G, respectively; and each is marked by an arrowed boldface letter. A knuckle arrow line marks the position of the 90 bp G on the *L. casei* ATCC 393 primer extension map. The TGN -10 motif of promoter P_{jun} is indicated by black stars. The OrfA translation start codon ATG is also indicated by a knuckle arrow. The 25 bp deletion in the IRL region is boxed. Note that the transcription start site at the 80 bp T could not be unambiguously defined on promoter P_{jun} (TTGACT-N₁₇-TAAGCT) at the 5'-IRR-DR(ACC)-IRL-3' junction region. (b) The putative promoter P_{djun} (TGTCTC-N₁₇-TATTGA) at the 5'-IRR-IRL Δ 25-IRL-3' junction region was determined and is underlined. The 17 bp spacer is underlined in bold. The transcription start site at the 80 bp T is indicated by an arrowed boldface letter. This figure is drawn to show the location of the transcription start site 80 bp T on promoter P_{djun} . (c) This plot shows that a reverse-transcribed DNA fragment of about 80 bp was determined in the *E. coli*-based system. Note that the fluorescence intensity quantified by the GeneMapper software (y axis) in (c) was not identical with that shown in (a); the maximum intensity of the 80 bp peak of the former was nearly the same as that of the minimum peak at 80 bp of the latter.

OrfAB. In contrast, no linearization product was found for the case where plasmid pTR, carrying a 37 bp truncation at the IRR end, was incubated with OrfAB (Fig. 5d). However, OrfAB was able to linearize pTL, where the

37 bp truncation was at the IRL end (Fig. 5d) and no corresponding high-molecular-mass autointegration product was detected in this case (Fig. 5d). These results showed that IRR was the first target chosen by OrfAB when the

Fig. 5. (a) *In vitro* activity and metal ion requirement of OrfAB. Digested products of the autointegration (Auto), linearized (L), relaxed circle (RC) or supercoiled (SC) forms of plasmid pGIS (500 ng) were incubated with purified OrfAB (500 ng). Lane 1, molecular mass standards (1 kb ladder); lane 2, plasmid pGIS digested with *Pst*I. Lanes 3–7 are reactions with the plasmid incubated for 0, 0.5, 3 and 15 min, respectively, with 50 nM OrfAB. (b) Digested products of plasmid pGIS (500 ng) after being incubated with OrfAB with or without the presence of various divalent metal ions. Lane 1 shows the case where no divalent metal ion was added. Lanes 2–4, pGIS incubated with OrfAB in the presence of Ca²⁺, Mn²⁺ or Mg²⁺, respectively. Lane 5, molecular mass standards (1 kb ladder). (c) The gel images of the Auto and SC digested products from (a) were traced by a densitometer and the corresponding densities were plotted as a function of incubation time (note non-linear scale). The density measured for SC pGIS without the addition of OrfAB was defined as 100 %. (d) Digested products of plasmids pTL (200 ng) and pTR (500 ng) after being incubated with OrfAB (500 ng) in the presence of Mg²⁺. Lanes 2–6 show pTL after being incubated with 0, 50, 100, 200 and 400 ng OrfAB, respectively. Lanes 7–9 show pTR after being incubated with 0, 200 and 400 ng of OrfAB, respectively. Molecular mass standard (1 kb ladder) was loaded in lane 1. The DNA fragments were separated in 0.8 % agarose gel in TBE buffer and the inverted images of the corresponding ethidium bromide-stained agarose gels are presented.

transposase cleaved a single DNA strand at one IS end to generate a free 3'-OH group in the first step of transposition. The effect of divalent metal ions on the OrfAB activity was examined at pH 7.5. The results showed that OrfAB was almost inactive in the presence of EDTA or in the absence of metal ions. The OrfAB activity was strongly enhanced in the presence of Mn^{2+} and Mg^{2+} *in vitro* but was severely impaired if these metal ions were replaced by Ca^{2+} (Fig. 5b). We also observed that the transposase preferred Mg^{2+} over Mn^{2+} *in vitro* since the concentration of the former and latter required by the enzyme to achieve the optimal activity was 10 and 50 mM, respectively.

DISCUSSION

In this study, we have identified an unusual ISLC3 circular intermediate which is not similar to that ordinarily found in the IS3-like elements. We have purified OrfAB and demonstrated that it cuts the DNA at the IRR as predicted. We have also shown that there are three promoters, namely strong promoter P_{jun} (TTGACT- N_{17} -TAAGCT), comparatively weaker promoter P_{djun} (TGTCTC- N_{17} -TATTGA), plus indigenous promoter P_{IRL} , formed at the junction of the circular form, which is a rather uncommon feature found in the members of this family. Some previous studies have shown that formation of the circle junction in a few IS3 members will create only a new promoter P_{jun} where the IRL and IRR ends are joined by some linking bases (Duval-Valentin *et al.*, 2001; Szeverenyi *et al.*, 1996; Lewis *et al.*, 2004). It has been shown for IS911, IS2 or other IS elements of Gram-negative or positive bacteria that formation of the IRL-IRR circle junction will generate a promoter whose activity may be stronger than that of the indigenous one in driving the transcription of transposase (Polard *et al.*, 1992; Turlan & Chandler, 1995). However, no intact IRR-DR-IRL junction has been found in the lactobacilli studied here, although we did observe the presence of an intact IRR-DR-IRL junction in *E. coli* transformed by a plasmid carrying a wild-type ISLC3 copy. We also found that the copy number of intact IRR-DR-IRL junction in the lactobacilli studied was much lower than that of the 25 bp deletion junction (IRR-IRLA25), the major product detected in an IPCR assay.

It is currently unknown why the intra-molecular recombination by the transposition of ISLC3 in lactobacilli will cause a frequent deletion of 25 bp at the IRL region. While the head-to-tail tandem dimer carrying the truncated IRR-IRLA25 junction predominantly existed as a stable copy (or copies) on the chromosome of lactobacilli studied here, the formation of intact IRR-DR-IRL circular form could be an extrachromosomal intermediate occasionally generated during the transposition of the IS. This is evidenced by significant difference either in the amount of the two types of structures detected or in the corresponding promoter activities measured here or reported elsewhere by others (Olasz *et al.*, 1993; Szeverenyi *et al.*, 2003) using the same type of promoter activity assays.

The corresponding promoter activity of P_{jun} was found to be much stronger than that of P_{djun} , which is a putative promoter formed by the truncated IRR-IRLA25. Moreover, P_{jun} was the strongest promoter among the three promoters studied. Although some variations in the -10 region (TAAGCT) of P_{jun} with the consensus -10 hexamer box were also found (Fig. 4a), the presence of TGn motifs directly upstream of the -10 region of P_{jun} may enhance its binding with the RNA polymerase, which renders P_{jun} a stronger promoter than P_{IRL} or P_{djun} (Burr *et al.*, 2000). The sequence of the -10 region (TATAAT) of P_{IRL} appears to be the most conserved one among the three promoters studied, but the -35 region (CTGAGA) of P_{IRL} was not conserved with the consensus TTGACA box. On the other hand, the -35 hexamer (TGTCTC) of P_{djun} was less conserved, since only 1 bp matched the consensus TTGACA box (McCracken *et al.*, 2000). However, a number of promoters have now been reported where the specific -35 regions are not required for transcription initiation (Bown *et al.*, 1997; Chan *et al.*, 1990). Although no TGn motif was detected in P_{djun} , the sequence of the -10 region (TATAAT) of P_{djun} appears to be the most conserved one, which gave rise to its transcription activity (Fig. 4b).

It is known that in the initial step of circle formation by IS3 family members one IS end acts as the donor and the other one as the target. The donor DNA is cleaved and then transferred to the DNA adjacent to the other end. However, both ends of most of the IS3-related elements act equally well as a donor or target (Lewis & Grindley, 1997). The intra-molecular recombination between two ends of an IS element is the key for formation of circles or tandem repeats (Szeverenyi *et al.*, 2003). In the transposition by IS911, an initial single-strand cleavage at the donor end resulting from intra-molecular recombination creates a free 3' OH group which will attack the target end on the same strand of DNA.

The first step in the transposition reaction by the IS3 family is an asymmetric single-stranded cleavage at the active donor end, followed by the DNA strand transfer process. This was confirmed here through an *in vitro* activity assay for the overexpressed transposase OrfAB, where the cleavage of IRR by the enzyme was found to be prior to that of IRL. Furthermore, the donor activity of the initial asymmetric DNA strand transfer by the transposase virtually vanished if the IRR was deleted. For the transposition of IS911 studied by others (Turlan *et al.*, 2000), both IRR and IRL ends are found to be used in the intermolecular transposition, while in some minor cases only a single end is used. It is known that an IR-damaged IS can not transpose alone and will eventually lead to elimination of the transposon if the DNA sequence at the recipient end is mutated before the strand transfer process (Mahillon & Chandler, 1998; Wagner, 2006). On the other hand, each IS can either transpose alone or it can transpose together with the other IS in a composite transposon made by two identical ISs. This composite transposon is

evolutionarily stable and the corresponding ISs will act in concert in the transposition process to nearby genomic locations. The single IRLΔ25 structure detected here in the head-to-tail tandem dimer could be evolutionarily unstable since it was unable to form a composite transposon.

ACKNOWLEDGEMENTS

This work was supported in parts by grants from the National Science Council Taiwan (NSC94-2313-B007-001 and NSC95-2313-B007-002).

REFERENCES

- Bown, J., Barne, K., Minchin, S. & Busby, S. (1997). Extended -10 promoters. *Nucleic Acids Mol Biol* **11**, 41–52.
- Burr, T., Mitchell, J., Kolb, A., Minchin, S. & Busby, S. (2000). DNA sequence elements located immediately upstream of the -10 hexamer in *Escherichia coli* promoters: a systematic study. *Nucleic Acids Res* **28**, 1864–1870.
- Chan, B., Spassky, A. & Busby, S. (1990). The organization of open complexes between *Escherichia coli* RNA polymerase and DNA fragments carrying promoters either with or without consensus -35 region sequences. *Biochem J* **270**, 141–148.
- Dalrymple, B. (1987). Novel rearrangements of IS30 carrying plasmids leading to the reactivation of gene expression. *Mol Gen Genet* **207**, 413–420.
- Duval-Valentin, G., Normand, C., Khemici, V., Marty, B. & Chandler, M. (2001). Transient promoter formation: a new feedback mechanism for regulation of IS911 transposition. *EMBO J* **20**, 5802–5811.
- Ehrmann, M. A. & Vogel, R. E. (2001). Characterisation of IS153, an IS3-family insertion sequence isolated from *Lactobacillus sanfranciscensis* and its use for strain differentiation. *Syst Appl Microbiol* **24**, 443–450.
- Fayet, O., Ramond, P., Polard, P., Prere, M. F. & Chandler, M. (1990). Functional similarities between retrovirus and the IS3 family of bacterial insertion sequences? *Mol Microbiol* **4**, 1771–1777.
- Forsman, P. & Alatossavam, T. (1994). Repeated sequences and the sites of genome rearrangements in bacteriophages of *Lactobacillus delbrueckii* subsp. *lactis*. *Arch Virol* **137**, 43–54.
- Galas, D. J. & Chandler, M. (1989). Bacterial insertion sequences, pp. 109–162. In *Mobile DNA*. Edited by D. E. Berg & M. M. Howe. Washington, DC: American Society for Microbiology.
- Haas, M. & Rak, B. (2002). *Escherichia coli* insertion sequence IS150: transposition via circular and linear intermediates. *J Bacteriol* **184**, 5833–5841.
- Harley, C. B. & Reynolds, R. P. (1987). Analysis of *E. coli* promoter sequences. *Nucleic Acids Res* **15**, 2343–2361.
- Keilty, S. & Rosenberg, M. (1987). Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J Biol Chem* **262**, 6389–6395.
- Kiss, J. & Olsasz, F. (1999). Formation and transposition of the covalently closed IS30 circle: the relation between tandem dimers and monomeric circles. *Mol Microbiol* **34**, 37–52.
- Knudtson, K. L. & Minion, F. C. (1994). Use of *lac* gene fusions in the analysis of *Acholeplasma* upstream gene regulatory sequences. *J Bacteriol* **176**, 2763–2766.
- Lewis, L. A. & Grindley, N. D. (1997). Two abundant intramolecular transposition products, resulting from reactions initiated at a single end, suggest that IS2 transposes by an unconventional pathway. *Mol Microbiol* **25**, 517–529.
- Lewis, L. A., Cylin, E., Lee, H. K., Saby, R., Wong, W. & Grindley, N. D. (2004). The left end of IS2: a compromise between transpositional activity and an essential promoter function that regulates the transposition pathway. *J Bacteriol* **186**, 858–865.
- Lo, T.-C., Shih, T. C., Lin, C. F., Chen, H. W. & Lin, T. H. (2005). Complete genomic sequence of the temperate bacteriophage φAT3 isolated from *Lactobacillus casei* ATCC 393. *Virology* **339**, 42–55.
- Mahillon, J. & Chandler, M. (1998). Insertion sequences. *Microbiol Mol Biol Rev* **62**, 725–774.
- Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V. & other authors (2006). Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* **103**, 15611–15616.
- McCracken, A., Turner, M. S., Giffard, P., Hafner, L. M. & Timms, P. (2000). Analysis of promoter sequences from *Lactobacillus* and *Lactococcus* and their activity in several *Lactobacillus* species. *Arch Microbiol* **173**, 383–389.
- Merighi, M., Majerczak, D. R., Zianni, M., Tessanne, K. & Coplin, D. L. (2006). Molecular characterization of *Pantoea stewartii* subsp. *stewartii* HrpY, a conserved response regulator of the Hrp type III secretion system, and its interaction with the *hrpS* promoter. *J Bacteriol* **188**, 5089–5100.
- Miller, J. H. (1992). *A Short Course in Bacterial Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Olsasz, F., Stalder, R. & Arber, W. (1993). Formation of the tandem repeat (IS30)₂ and its role in IS30-mediated transpositional DNA rearrangements. *Mol Gen Genet* **239**, 177–187.
- Peake, S. E. & Stanley, G. (1978). Partial characterization of a bacteriophage of *Lactobacillus bulgaricus* isolated from yoghurt. *J Appl Bacteriol* **44**, 321–323.
- Perkins-Balding, D., Duval-Valentin, G. & Glasgow, A. C. (1999). Excision of IS492 requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*. *J Bacteriol* **181**, 4937–4948.
- Polard, P., Prere, M. F., Fayet, O. & Chandler, M. (1992). Transposase-induced excision and circularization of the bacterial insertion sequence IS911. *EMBO J* **11**, 5079–5090.
- Ponnambalam, S., Webster, C., Bingham, A. & Busby, S. (1986). Transcription initiation at the *Escherichia coli* galactose operon promoters in the absence of the normal -35 region sequences. *J Biol Chem* **261**, 16043–16048.
- Prudhomme, M., Turlan, C., Claverys, J. P. & Chandler, M. (2002). Diversity of Tn4001 transposition products: the flanking IS256 elements can form tandem dimers and IS circles. *J Bacteriol* **184**, 433–443.
- Reimann, C. & Haas, D. (1990). The *istA* gene of insertion sequence IS21 is essential for cleavage at the inner 3' ends of tandemly repeated IS21 elements *in vitro*. *EMBO J* **9**, 4055–4063.
- Reimann, C., Moore, R., Little, S., Savioz, A., Willetts, N. S. & Haas, D. (1989). Genetic structure, function and regulation of the transposable element IS21. *Mol Gen Genet* **215**, 416–424.
- Rousseau, P., Normand, C., Loot, C., Turlan, C., Alazard, R., Duval-Valentin, G. & Chandler, M. (2002). Transposition of IS911. In *Mobile DNA II*, pp. 367–383. Edited by N. L. Craig, R. Craigie, M. Gellert & A. M. Lambowitz. Washington, DC: American Society for Microbiology.
- Sabelnikov, A. G., Greenberg, B. & Lacks, S. A. (1995). An extended -10 promoter alone directs transcription of the DpnII operon of *Streptococcus pneumoniae*. *J Mol Biol* **250**, 144–155.

- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989).** *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sekine, Y., Aihara, K. & Ohtsubo, E. (1999).** Linearization and transposition of circular molecules of insertion sequence IS3. *J Mol Biol* **294**, 21–34.
- Shiga, Y., Sekine, Y. & Ohtsubo, E. (1999).** Transposition of IS1 circles. *Genes Cells* **4**, 551–561.
- Skaugen, M. & Nes, I. F. (1994).** Transposition in *Lactobacillus sake* and its abolition of lactocin S production by insertion of IS1163, a new member of the IS3 family. *Appl Environ Microbiol* **60**, 2818–2825.
- Skaugen, M. & Nes, I. F. (2000).** Transposition in *Lactobacillus sakei*: inactivation of a second lactocin S operon by the insertion of IS1520, a new member of the IS3 family of insertion sequences. *Microbiology* **146**, 1163–1169.
- Szeverenyi, I., Bodoky, T. & Olsasz, F. (1996).** Isolation, characterization and transposition of an (IS2)₂ intermediate. *Mol Gen Genet* **251**, 281–289.
- Szeverenyi, I., Nagy, Z., Farkas, T., Olsasz, F. & Kiss, J. (2003).** Detection and analysis of transpositionally active head-to-tail dimers in three additional *Escherichia coli* IS elements. *Microbiology* **149**, 1297–1310.
- Ton-Hoang, B., Betermier, M., Polard, P. & Chandler, M. (1997).** Assembly of a strong promoter following IS911 circularization and the role of circles in transposition. *EMBO J* **16**, 3357–3371.
- Ton-Hoang, B., Polard, P. & Chandler, M. (1998).** Efficient transposition of IS911 circles in vitro. *EMBO J* **17**, 1169–1181.
- Turlan, C. & Chandler, M. (1995).** IS1-mediated intramolecular rearrangements: formation of excised transposon circles and replicative deletions. *EMBO J* **14**, 5410–5421.
- Turlan, C., Ton-Hoang, B. & Chandler, M. (2000).** The role of tandem IS dimers in IS911 transposition. *Mol Microbiol* **35**, 1312–1325.
- Wagner, A. (2006).** Cooperation is fleeting in the world of transposable elements. *PLOS Comput Biol* **2**, e162.
- Walker, D. C. & Klaenhammer, T. R. (1994).** Isolation of a novel IS3 group insertion element and construction of an integration vector for *Lactobacillus* spp. *J Bacteriol* **176**, 5330–5340.

Edited by: M. Kleerebezem