Isolation and characterization of functional insertion sequences of rhizobia

Ismael Hernandez-Lucas[1], Jose Augusto Ramirez-Trujillo[1], Miguel Angel Gaitan[1], Xianwu Guo[2], Margarita Flores[2], Esperanza Martinez-Romero[2], Ernesto Perez-Rueda[1] & Patrick Mavingui[3]

[1]Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México; [2]Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México; and [3]UMR CNRS 5557, Ecologie Microbienne, IFR41 Bio-Environnement et Santé, Université Lyon 1, Villeurbanne, France

Correspondence: Ismael Hernandez-Lucas, Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México. Av. Universidad 2001, Cuernavaca, Morelos 62210, México. Tel.: +52 777 329-1627; fax: +52 777 313-8673; e-mail: ismaelh@ibt.unam.mx

Received 3 March 2006; revised 5 May 2006; accepted 11 May 2006.
First published online 6 June 2006.

doi:10.1111/j.1574-6968.2006.00319.x

Editor: Yaacov Okon

Keywords
rhizobia; insertion sequences; transposition.

Abstract

Rhizobia are a group of bacteria that form nodules on the roots of legume host plants. The sequenced genomes of the rhizobia are characterized by the presence of many putative insertion sequences (IS) elements. However, it is unknown whether these IS elements are functional and it is therefore relevant to assess their transposition activity. In this work, several functional insertion sequences belonging to the IS*1256*, IS*3*, IS*5*, IS*166*, and IS*21* families were captured from *Rhizobium tropici*, *Rhizobium* sp. NGR234 and *Sinorhizobium meliloti*, using pGBG1 as a trapping system. *In silico* analysis shows that homologs of rhizobia mobile elements are present in distantly related genomes, suggesting that *Rhizobium* IS elements are prone to genetic transfer.

Introduction

Since the original discovery of transposable elements by Barbara McClintock in the late 1940s, many mobile elements have been identified in eukaryotes and prokaryotes. These genetic elements mediate numerous molecular and genetic phenomena such as gene activation, repression, duplication, deletion, inversion, and transfer. These DNA rearrangements can be advantageous, deleterious, or neutral for the host organism. Among transposable elements, insertion sequence (IS) elements are the most widespread. Their sizes range from 200 bp to more than 2500 bp. IS elements share a common structure consisting of short inverted repeats at their ends, and at least one ORF coding for a transposase (Galas & Chandler, 1989). The transposase binds to both inverted repeats and the target DNA. It then performs DNA breakage to remove the element and insert it into its new site. The inserted IS element is flanked by short duplicated sequences at the target site. While transposition of most IS elements occurs by this cut and paste mechanism, several bacterial transposons move by replicative transposition during which two copies of the transposable element are generated. IS elements have been found in different bacterial genomes including those of rhizobia. For instance, the genomes of *Agrobacterium tumefaciens* (Wood *et al.*, 2001), *Bradyrhizobium japonicum* (Kaneko *et al.*, 2002), *Mesorhizobium loti* (Kaneko *et al.*, 2000), and *Sinorhizobium meliloti* (Galibert *et al.*, 2001) contain IS elements on both chromosome and plasmids. Comparative genomic studies have inferred that these IS elements have shaped the mosaic structure of *Rhizobium* replicons (Freiberg *et al.*, 1997; Galibert *et al.*, 2001; Gonzalez *et al.*, 2003). In some cases, participation of IS elements in recombination leading to genomic rearrangements has been evidenced by experimental studies on distinct *Rhizobium* species (Mavingui *et al.*, 1998, 2002; Flores *et al.*, 2000). Although several works have reported the presence of IS elements in rhizobia (Wheatcroft & Watson, 1988; Kosier *et al.*, 1993; Selbitschka *et al.*, 1995, 1999; Rochepeau *et al.*, 1997; Schneiker *et al.*, 1999), transposition activity was demonstrated for only some of them (Priefer *et al.*, 1981, 1989; Ruvkun *et al.*, 1982; Gay *et al.*, 1985; Dusha *et al.*, 1987; Simon *et al.*, 1991; Zekri & Toro, 1996). In this work, we report several functional IS elements in *R. tropici* CFN299, *Rhizobium* sp. NGR234, and *S. meliloti* 1021.

Materials and methods

Bacterial strains, plasmids, culture conditions and matings

The bacterial strains and plasmids used are listed in Table 1. Rhizobia were grown on PY medium (0.5% peptone, 0.3% yeast extract, 10 mM $CaCl_2$) supplemented with the following antibiotics ($\mu g\,mL^{-1}$) when required: chloramphenicol (Cm, 10), tetracycline (Tc, 10) and nalidixic acid (Nal, 20). *Escherichia coli* strains were grown in LB medium (1% peptone, 0.5% yeast extract, 1% NaCl) supplemented with chloramphenicol (10), kanamycin (Km, 25) and ampicillin (Amp, 100) when needed. *Escherichia coli* and rhizobia were grown at 30 °C. Conjugation experiments were performed in the presence of the helper strain *E. coli* HB101 containing pRK2013 (Figurski & Helinski, 1979) as described previously (Charles & Finan, 1991).

IS entrapment procedure

To isolate active IS elements from rhizobia, the plasmid pGBG1 was used as a trap (Schneider *et al.*, 2000). pGBG1 confers resistance to chloramphenicol and contains the mutagenesis target that consists of a *tetA* gene under the control of the *pR* promoter from bacteriophage λ which is repressed by λ CI repressor. Insertion of mobile elements in the repressor activates the *tetA* gene and allows the positive selection of tetracycline-resistant colonies.

The plasmid pGBG1 was introduced by triparental mating into rhizobia. Transconjugants were obtained at frequency of *c.* 10^{-5} on PY medium supplemented with chloramphenicol. The presence of pGBG1 was ascertained by plasmid profiles. For *Rhizobium* sp. NRG234 and *Sinorhizobium meliloti* 1021, five individual Cm^R clones containing pGBG1 were selected and were grown separately to late log phase in liquid PY medium supplemented with chloramphenicol. For *Rhizobium tropici* CNF299, one Cm^R colony of RH165 bearing pGBG1 was grown overnight in liquid minimal medium (MM) (Romanov *et al.*, 1994) supplemented with 10 mM succinate. To isolate derivative clones bearing IS inserted into the target, each culture was diluted and spread on the corresponding agar plate medium containing tetracycline. To isolate IS elements from bacteria inside nodules, seeds of *Phaseolus vulgaris* were inoculated with RH165. Nodules from two plants, harvested 18 days after inoculation, were crushed, diluted and 0.1 mL was spread on PY plates containing tetracycline. To estimate the frequencies of transconjugant appearance, cultures were also titered by plating on PY agar containing chloramphenicol alone.

Table 1. Bacterial strains and plasmids used

Strains or plasmids	Relevant characteristics	Source or reference
Strains		
Rhizobium tropici CFN299	Wild-type strain which contains plasmid pSym, plasmid B, plasmid A	Martínez-Romero *et al.* (1991)
Rhizobium tropici CFN299-10	CFN299 derivative lacking pB and 300 kb fragment of pSym	Martínez-Romero & Rosenblueth (1990)
Rhizobium sp. NGR234	Broad host range, Rif^R	Trinick (1980)
Sinorhizobium meliloti 1021	SU47 str-21	Meade *et al.* (1982)
Escherichia coli DH5α	*recA1* ϕ*80 lac*ZΔM15, *gyrA96*	Gibco BRL
Escherichia coli HB101	*recA13 rspL20* (Sm^R)	Gibco BRL
Plasmids		
pRK2013	ColE1 replicon with RK2 transfer region, Km^R Nm^R	Figurski *et al.* (1979)
pGBG1	Origin pBBR1MCS, *oriT*, Cm^R, silent *tetA*, *pR* promoter, CI repressor	Schneider *et al.* (2000)

DNA manipulations

Plasmid purification and genomic DNA extraction were performed according to published protocols (Sambrook *et al.*, 1989). Plasmid profiles were obtained by the Eckhardt method as modified by Hynes & McGregor (1990). For hybridization, DNA was digested with restriction enzymes then transferred from agarose gels to nylon membranes. Probes (internal fragments of IS elements) were labelled with ^{32}P by polymerase extension using random primers and hybridization was carried out under high stringency conditions (Southern, 1975). For sequencing, double stranded DNA was purified with the High Pure Plasmid Isolation Kit (Roche, Germany) and sequencing was performed with an automatic Perkin Elmer/Applied Biosystems 377-18 or Genome Express (Meylan, France) sequencer.

PCR amplification

Primers G11 (5′-TAT CAG CTA TGC GCC GAC CAG AAC-3′) and G12 (5′-GCC AAT CCC CAT GGC ATC GAG TAA-3′) were used to amplify IS elements inserted into the trapping vector pGBG1 (Schneider *et al.*, 2000). PCR assays were carried out in a 25 μL reaction containing the template genomic DNA (250 ng) in 1 × polymerase

reaction XL buffer II (Perkin-Elmer), 1.1 mM $Mg(OAc)_2$, 200 μM dNTPs, 5 pmol of each primer, and 1 U of *rTth* polymerase (Perkin-Elmer). PCR amplifications were performed in a 9700 Thermocycler (Perkin-Elmer) with the following conditions: an initial denaturation at 94 °C for 1 min; then 35 cycles of denaturation (94 °C, 15 s), annealing and extension (65 °C, 3 min); and a final extension at 72 °C for 7 min. PCR samples were electrophoresed through 0.8–1% agarose gels in TAE buffer, and stained with ethidium bromide.

Sequence analysis

ORF identification was performed with the BLASTX program (http://www.ncbi.nlm.nih.gov/BLAST/). In order to determine the domain organization of the insertion sequences, predicted ORFs were scanned against databases SUPERFAMILY v1.63 (Gough *et al*., 2001) and PFAM. To identify homologs of the IS elements reported in this work, a total of 152 genomes, 129 bacteria, 15 archaea, and eight eukarya (Entrez genome database ftp://ftp.ncbi.nlm.nih.gov/genomes/bacteria) were analyzed. The novel IS elements identified in this work were registered in the IS database at http://www-is.biotoul.fr/is.html

Nodulation assays

Plant nodulation assays were performed in flasks containing agar as described previously (Martínez *et al*., 1985). Seedlings of *Phaseolus vulgaris* cv. Negro Jamapa were inoculated with 10^5 *Rhizobium* cells.

Results and discussion

Isolation of IS elements from rhizobia

Active IS elements were isolated from rhizobia using the plasmid pGBG1 as a trap (see Materials and methods). The frequency of Tc^R transconjugants among Cm^R colonies was between 10^{-5} and 10^{-6} for both *Rhizobium* sp. NGR234 and *Sinorhizobium meliloti* 1021 strains. The plasmid profiles of a total of 100 Tc^R clones (20 per culture) from each strain were analyzed. For *Rhizobium tropici*, colonies resistant to tetracycline appeared at a frequency of around 10^{-6} from culture-based selection, without significant differences between PY and MM-succinate media, whereas those from nodules emerged at a frequency of 3×10^{-7}. A total of 46 Tc^R clones obtained from nodules and 85 Tc^R clones selected in culture media were screened by plasmid profiles. Presence of IS elements in the target was inferred by an increase in size of pGBG1 in plasmid patterns of Tc^R transconjugants. The proportion of resistant transconjugants harboring larger sized-pGBG1 plasmid was 22% for *Rhizobium* sp. NGR234 and 15% for *S. meliloti* 1021, whereas for *R. tropici* it was 6.5% from synthetic media and 8.5% from nodules. Plasmid patterns of the remaining Tc^R clones showed no detectable changes in the size of pGBG1, indicating point mutations, and microdeletion or microinsertion events. All positive Tc^R clones bearing potential mobile elements were further analyzed by PCR amplification using primers G11 and G12 (Schneider *et al*., 2000). Sequence analyses of the PCR products revealed eight distinct IS elements, five from *R. tropici*, designated as IS*Rtr1*, IS*Rtr2*, IS*Rtr3*, IS*Rtr4* and IS*Rtr5*, two from *Rhizobium* sp. NGR234 (Freiberg *et al*., 1997; Perret *et al*., 1997) designated here as IS*NGR3* and IS*NGR4*, and one from *S. meliloti* 1021 previously reported as IS*Rm1* (Watson & Wheatcroft, 1991). IS*Rtr1*, IS*NGR3*, IS*NGR4* and IS*Rm1* were isolated from bacteria grown in PY medium, and their sequences were identical to the published data. IS*Rtr2* and IS*Rtr3* were obtained from bacteria grown in MM-succinate. IS*Rtr4* and IS*Rtr5* were isolated from bacteria recovered from nodules inoculated with the *R. tropici* RH165 transconjugant.

Characterization of functional IS elements from rhizobia

A structural analysis of the isolated IS elements was performed. The relevant characteristics such as length (bp), number of ORFs, transposase size, G+C content, IS family, copy number and distribution of rhizobia IS elements in different genomes are shown in Table 2. As the functional IS elements from *Rhizobium* sp NGR234 and *S. meliloti* 1021, as well as the IS*Rtr1* of *R. tropici* isolated here were previously reported, only novel mobile IS obtained from *R. tropici* will be further described. In addition to the salient characteristics mentioned in Table 2, we found other relevant features of *R. tropici* IS elements. For instance, IS*Rtr2* contains a DDE motif typical of the IS*3* family (Mahillon & Chandler, 1998). IS*Rtr2* is highly similar (81%) to IS*868* of *Agrobacterium tumefaciens* biotype III (Paulus *et al*., 1991). Furthermore, the GC content of IS*Rtr2* and IS*868* are similar, 59.2% and 59.3%, respectively, suggesting a common origin. By sequencing the region adjacent to IS*Rtr2* located on the symbiotic plasmid of *R. tropici* CFN299, we determined that this mobile element is flanked by genes for citrate synthase and isocitrate lyase, with a GC content of 56.4% and 57.7%, respectively. This result implies that IS*Rtr2* may have an external origin. The copy numbers (i.e. number of hybridization bands) of IS*Rtr2* in *R. tropici* was determined by Southern hybridization of *Eco*RI-digested genomic DNA (Fig. 1). As no cleavage site for *Eco*RI is present in the IS element the number of hybridizing bands reflect the copy number. The resulting IS profiles of CFN299 and CFN299-10 (a derivative of 299 lacking plasmid B and 300 kb fragment of the symbiotic plasmid) shows that

Table 2. Mobile IS elements trapped from rhizobia

Strain	IS isolated	Length (bp)	G+C %	ORFs	IR/ mismatch (bp)	Tnp size aa	IS family*	Copy number†	Gene Bank accession number	Similarity (%) of rhizobia Tnps with other organisms
Rhizobium tropici CFN299	IS*Rtr1*	1364	61	2	28/4	172	IS*1256*	6	AF041379	45–85% α, Gram+, Archaeobacteria
	IS*Rtr2*	1321	59	2	26	301	IS*3*	1	AY751757	46–81% α, β, γ, Gram+
	IS*Rtr3*	933	57	1	28/7	273	IS*5*	1	AY753543	37–85% β, γ, Gram+, Archaeobacteria, Cyanobacteria
	IS*Rtr4*	932	58	1	28/12	291	IS*5*	6	AY753544	44–87% α, β, Cyanobacteria, Archaeobacteria
	IS*Rtr5*	2699	60	3	27	451	IS*166*	1	DQ499058	38–70% α, β, γ, Gram+, Cyanobacteria, Archaeobacteria
Rhizobium sp. NGR234	IS*NGR3*	2625	62	2	31/2; 34/13	516	IS*21*	3	U00090	38–65% α, β, γ, Gram+, Cyanobacteria, Archaeobacteria
	IS*NGR4*	3324	63	4	0	694	ND	9	U00090	49–81% α, β, Gram+
Sinorhizobium meliloti 1021	IS*Rm1*	1321	60	2	23/3	276	IS*3*	9	X56563	54–60% α, β

*Based on protein and catalytic domain analyses.
†Determined by hybridization.
IS, insertion sequence; ND, not determined; Tnp, transposase; IR, inverted repeats.

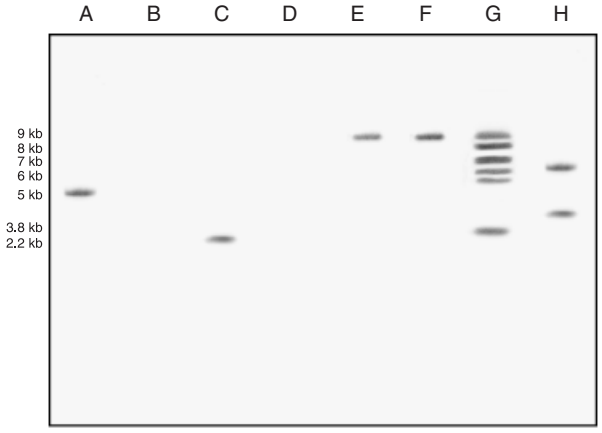

Fig. 1. Presence of the IS elements in *Rhizobium tropici*. Autoradiogram of Southern blot of digested genomic DNA hybridized with IS probes. In lane A, CFN299 shows a 5 kb signal. In lane B, no signal was detected in CFN299-10. In lane C, CFN299 shows a 2.2 kb hybridization band. In lane D, no hybridization band was detected in CFN299-10. In lane E, CFN299, and lane F, CFN299-10 identical hybridization profile was obtained. In lane G, CFN299 shows six bands ranging from 3.8 to 9 kb. In lane H, CFN299-10 present two bands of 4.5 and 6.5 kb. Genomic DNAs were restricted with endonucleases *Eco*R1 (lanes A and B) and *Pst*I (lanes C–H). Lanes A–B, C–D, E–F, and G–H correspond to DNA pattern hybridized with probe IS*Rtr2*, IS*Rtr3*, IS*Rtr5*, and IS*Rtr4*, respectively.

CFN299 presents a single hybridization band, while no signal was obtained with CFN299-10, suggesting that the mobile element is located on a plasmid. Furthermore, the sequencing project of the *R. tropici* symbiotic plasmid confirms that IS*Rtr2* is located on this plasmid. No hybridization signal was detected in the genome of *Agrobacterium*. sp. K-Ag-3, *Agrobacterium*. sp. Ch-Ag-4, *Mesorhizobium huakuii* CCBAU2609, *Rhizobium etli* CFN42, *Rhizobium* sp. CFN234, *Sinorhizobium terangae* USDA4102, and *Mesorhizobium mediterraneum* USDA 3392 (data not shown). The scarcity of IS*Rtr2* among rhizobia suggests that this mobile element was recently acquired in its host genome.

IS*Rtr3* encodes a putative transposase of 273 aa which is highly similar (85%) to the IS element IS*RlF7-2* of *R. leguminosarum* (W. Selbitschka *et al*., unpublished results), and it contains the N3 (DGGY) and C1 (LPRRWVVERT-FAWLG) domains of the IS*5* family (Rezsohazy *et al*., 1993). In order to determine the copy number of IS*Rtr3* in *R. tropici*, *Pst*I restricted genomic DNA was hybridized with an internal fragment of IS*Rtr3* as a probe (Fig. 1). One hybridizing band of 2.2 kb was detected in CFN299, however no signal was observed in CFN299-10. This result indicates that IS*Rtr3* is present in *R. tropici* plasmid B or in the symbiotic plasmid.

IS*Rtr4* is predicted to encode a transposase of 291 aa that is similar (87%) with IS*Rm4-1* from *S. meliloti* GR4 (Soto *et al*., 1992; Zekri *et al*., 1998). IS*Rtr4* contains the N3 and C1 domains, typical of the IS*5* family (Mahillon & Chandler, 1998). The copy number (Fig. 1) of this element was determined and the result shows that it is present in a single copy in both CFN299 and CFN299-10, therefore the exactly location of IS*Rtr4* remains to be established.

IS*Rtr5* presents the features typical of the IS*66* family, such as the presence of multiple ORFs. IS*Rtr5* encodes products of 159, 117 and 451 aa, respectively. The 159 and 117 aa proteins are 67% and 92% identical to *ynH* and *ynI* of *A. tumefaciens* (Winans *et al*., unpublished results), whereas the 451 aa protein is 70% identical to a transposase from *Nitrobacter hamburgensis* (Copeland *et al*., unpublished

results). It is important to mention that *ynH* and *ynI* ORFs overlap by 17 bp that include the termination codon of *ynH* and the initiation codon of *ynI*, suggesting that translational coupling is involved in expression of YnI; this genetic signature is common in members of the IS*66* family (Chang-Gyun *et al*., 2001). Southern hybridization experiments show that IS*Rtr5* was present in six copies in the genome of *R. tropici* CFN299 (Fig. 1). The hybridization profiles of CFN299-10 suggest that at least four copies of IS*Rtr5* are located on the symbiotic plasmid or in plasmid B.

In this work, the mobility of the IS elements from rhizobia was evaluated in minimal and rich media as well as in nodules using a plasmid trapping system. The percentage of IS capture was similar in free-living conditions and in nodules. The mobile elements isolated have a smaller size (932–3324 bp) and present conservative properties of IS elements such as inverted repeats and the presence of at least one ORF encoding for a transposase. Some of the mobile elements captured in this work seem to be specific for each rhizobial strain, since as they are not present in the completely sequenced genomes of other members of the *Rhizobiaceae* family. Therefore, these mobile elements can be used as a molecular tool to identify strains belonged to *R. tropici*, *S. meliloti* and *Rhizobium* sp. NGR234. It is important to mention that a moderate number of IS elements were trapped, even though some of them not detected here are known to be present in high copy numbers (e.g. in the genome of *S. meliloti* 1021 IS*Rm5* is present in 10 copies and IS*Rm11* in 12 copies). The reason could be that such elements have a low frequency of transposition, or that under the physiological conditions tested the mobilization of these elements is not occurring, or these elements have preferential target sequences, or a conditional function. It is also possible that some of these IS are no longer functional. However, their ability to move should be tested by other methods.

IS elements are involved in several process such as transposition and rearrangements. IS elements may confer a selective advantage to their host. For instance, it was reported that increased motility in *Escherichia coli* is obtained by spontaneous insertion of an IS*1* upstream of the *flhD* operon (Barker *et al*., 2004). It is also known that IS*150* mediates deletions of the ribose operon, and such deletions are beneficial to *E. coli* in glucose minimal medium (Cooper *et al*., 2001). Previously we reported that several nodulation and nitrogen fixation genes are delimited by two IS*Rtr1* elements and recombination between these elements generate a circular 60 kb fragment. *Agrobacterium tumefaciens* containing the 60 kb region acquires the ability to produce Nod factors and to nodulate *Phaseolus vulgaris*. In this example, a biological role of IS*Rtr1* was demonstrated (Mavingui *et al*., 1998). However many important issues such as the functional role of rhizobial IS's in free-living cells and in symbiosis, as well as their transposition mechanism remain to be established. In this regard, preliminary results obtained in our laboratory suggest that the IS*166* family represented by IS*Rtr5* move by replicative transposition whereas the IS*5* family represented by IS*Rtr3* move by a cut and paste mechanism in *R. tropici*.

In conclusion, we demonstrated the functionality of six IS elements (IS*Rtr2*, IS*Rtr3*, IS*Rtr4*, IS*Rtr5*, IS*NGR3* and IS*NGR4*) and confirmed the transposition activity of two previously reported mobile elements (IS*Rtr1* and IS*Rm1*). The mobile elements isolated are present in single or multiple copies in host genomes and homologous transposases were detected in several other bacterial genera as well as in archaea branch suggesting that rhizobia IS elements are prone to genetic transfer. In addition to their mobility and potential associated mutator phenotype, IS elements are sites for homologous recombination that generate genomic rearrangements, thus participating in shaping the overall structure of rhizobia genomes.

Acknowledgements

We thank D.J. De Los Santos-Ramos, K.K. Guzmán-Rabadan, A. Gamez-Reyes, G. Uribe-Figueroa, M. Rodriguez-Padilla, M.A. Moreno-Ocampo, R.M. Ocampo-Vargas, J. Espiritu-Salazar, for technical help and discussion, and M. Dunn, J. Caballero-Mellado and W.J. Broughton for useful scientific comments to improve the manuscript. This research was supported partly by grants to I.H.L from DGAPA/UNAM (IN206802 and IN206705).

References

Barker CS, Prüß BM & Matsumura P (2004) Increased motility of *Escherichia coli* by insertion sequence element integration into the regulatory region of the *flhD* operon. *J Bacteriol* **186**: 7529–7537.

Chang-Gyun H, Shiga Y, Tobe T, Sasakawa C & Ohtsubo E (2001) Structural and functional characterization of IS*679* and IS*66*-family elements. *J Bacteriol* **183**: 4296–4304.

Charles TC & Finan TM (1991) Analysis of a 1600-kilobase *Rhizobium meliloti* megaplasmid using defined deletions generated *in vivo*. *Genetics* **127**: 5–20.

Cooper VS, Schneider D, Blot M & Lenski RE (2001) Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. *J Bacteriol* **183**: 2834–2841.

Dusha I, Kovalenko S, Bonfalvi Z & Kondorosi A (1987) *Rhizobium meliloti* insertion element IS*Rm2* and its use for identification of the *fixX* gene. *J Bacteriol* **169**: 1403–1409.

Figurski DH & Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci USA* **76**: 1648–1652.

Flores M, Mavingui P, Perret X, Broughton WJ, Romero D, Hernández G, Davila G & Palacios R (2000) Prediction, identification, and artificial selection of DNA rearrangements in *Rhizobium*: toward a natural genomic design. *Proc Natl Acad Sci USA* **97**: 9138–9143.

Freiberg C, Fellay R, Bairoch A, Broughton WJ, Rosenthal A & Perret X (1997) Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* **387**: 394–401.

Galas DJ & Chandler M (1989) Bacterial insertion sequences. *Mobile DNA* (Berg DE & Howe MM, eds), pp. 109–162. American Society for Microbiology, Washington, DC.

Galibert F, Finan TM, Long SR, *et al.* (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**: 668–672.

Gay P, Le coq D, Steinmetz M, Berkelman T & Kado CI (1985) Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *J Bacteriol* **164**: 918–921.

Gonzalez V, Bustos P, Ramirez-Romero MA, *et al.* (2003) The mosaic structure of the symbiotic plasmid of *Rhizobium etli* CFN42. *Genome Biol* **4**: R36.

Gough J, Karplus K, Hughey R & Chothia C (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J Mol Biol* **313**: 903–919.

Hynes MF & McGregor NF (1990) Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen fixing nodules by *Rhizobium leguminosarum*. *Mol Microbiol* **4**: 567–574.

Kaneko T, Nakamura Y, Sato S, *et al.* (2000) Complete genome sequence of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res* **7**: 331–338.

Kaneko T, Nakamura Y, Sato S, *et al.* (2002) Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res* **9**: 189–197.

Kosier B, Puhler A & Simon R (1993) Monitoring the diversity of *Rhizobium meliloti* field and microcosm isolates with a novel rapid genotyping method using insertion elements. *Mol Ecol* **2**: 35–46.

Mahillon J & Chandler M (1998) Insertion sequences. *Microbiol Mol Biol Rev* **62**: 725–774.

Martínez E, Pardo MA, Palacios R & Cevallos MA (1985) Reiteration of nitrogen fixation gene sequences and specificity of *Rhizobium* in nodulation and nitrogen fixation in *Phaseolus vulgaris*. *J Gen Microbiol* **131**: 1779–1786.

Martinez-Romero E & Rosenblueth M (1990) Increased bean (*Phaseolus vulgaris* L.) nodulation competitiveness of genetically modified *Rhizobium* strains. *Appl Environ Microbiol* **56**: 2384–2388.

Martínez-Romero E, Segovia L, Mercante M, Franco A, Graham P & Pardo MA (1991) *Rhizobium tropici*: a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. *Int J Syst Bacteriol* **41**: 417–426.

Mavingui P, Laeremans T, Flores M, Romero D, Martínez-Romero E & Palacios R (1998) Genes essential for nod factor production and nodulation are located on a symbiotic amplicon (AMPRtrCFN299pc60) in *Rhizobium tropici*. *J Bacteriol* **180**: 2866–2874.

Mavingui P, Flores M, Guo X, Davila G, Perret X, Broughton WJ & Palacios R (2002) Dynamics of genome architecture in *Rhizobium* sp. strain NGR234. *J Bacteriol* **184**: 171–176.

Meade HM, Long SR, Ruvkun GB, Brown SE & Ausubel FM (1982) Physical and genetic characterisation of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J Bacteriol* **149**: 114–122.

Paulus F, Canaday J & Otten L (1991) Limited host range Ti plasmids: recent origin from wide host range Ti plasmid and involvement of a novel IS element, IS*868*. *Mol Plant Microbe Interact* **4**: 190–197.

Perret X, Viprey V, Freiberg C & Broughton WJ (1997) Structure and evolution of NGRRS-1, a complex, repeated element in the genome of *Rhizobium* sp. strain NGR234. *J Bacteriol* **179**: 7488–7496.

Priefer UB, Burkardt HJ, Klipp W & Puhler A (1981) ISR1 an insertion element isolated from the soil bacterium *Rhizobium lupini*. *Movable Genetic Elements, Cold Spring Harbor Symposium on Quantum Biology, Vol. XLV*, pp. 87–91. Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.

Priefer UB, Kalinowski J, Ruger B, Heumann W & Puhler A (1989) ISR1 a transposable DNA sequence resident in *Rhizobium* class IV strains, shows structural characteristics of classical insertion elements. *Plasmid* **21**: 120–128.

Rezsohazy R, Hallet B, Delcour J & Mahillon J (1993) The IS*4* family of insertion sequences: evidence for a conserved transposase motif. *Mol Microbiol* **9**: 1283–1295.

Rochepeau P, Selinger LB & Hynes MF (1997) Transposon-like structure of a new plasmid-encoded restriction-modification system in *Rhizobium leguminosarum* VF39SM. *Mol Gen Genet* **256**: 387–396.

Romanov VI, Hernández-Lucas I & Martínez-Romero E (1994) Carbon metabolism enzymes of *Rhizobium tropici* cultures and bacteroids. *Appl Environ Microbiol* **60**: 2339–2342.

Ruvkun GB, Long SR, Meade HM, van den Bos RC & Ausubel FM (1982) IS*Rm1*: a *Rhizobium meliloti* insertion sequence that transposes preferentially into nitrogen fixation genes. *J Mol Appl Genet* **1**: 405–418.

Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schneider D, Faure D, Noirclerc-Savoye M, Barriere A, Coursange E & Blot M (2000) A broad-host-range plasmid for isolating mobile genetic E1 in gram-negative bacteria. *Plasmid* **44**: 201–207.

Schneiker S, Kosier B, Puhler A & Selbitschka W (1999) The *Sinorhizobium meliloti* insertion sequence (IS) element IS*Rm14* is related to a previously unrecognized IS element located adjacent to the *Escherichia coli* locus of enterocyte effacement (LEE) pathogenicity island. *Curr Microbiol* **39**: 274–281.

Selbitschka W, Arnold W, Jording D, Kosier B, Toro N & Puhler A (1995) The insertion sequence element IS*Rm2011-2* belongs to the IS*630*-Tc1 family of transposable elements and is abundant in *Rhizobium meliloti*. *Gene* **163**: 59–64.

Selbitschka W, Zekri S, Schroder G, Puhler A & Toro N (1999) The *Sinorhizobium meliloti* insertion sequence (IS) elements IS*Rm102F34*-1/IS*Rm7* and IS*Rm220-13-5* belong to a new family of insertion sequence elements. *FEMS Microbiol Lett* **172**: 1–7.

Simon R, Hotte B, Klauke B & Kosier B (1991) Isolation and characterization of insertion sequence elements from gram-negative bacteria by using new broad-host-range, positive selection vectors. *J Bacteriol* **173**: 1502–1508.

Soto MJ, Zorzano O, Olivares J & Toro N (1992) Sequence of IS*Rm4* from *Rhizobium meliloti* strain GR4. *Gene* **120**: 125–126.

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**: 503–517.

Trinick MJ (1980) Relationships among the fast-growing rhizobia of *Lablab purpureus*, *Leucaena leucochephala*, *Mimosa* sp., *Acacia farnesiana*, and *Sesbania grandiflora* and their affinities with other rhizobia groups. *J Appl Bacteriol* **49**: 39–53.

Watson RB & Wheatcroft R (1991) Nucleotide sequence of *Rhizobium meliloti* insertion sequence IS*Rm1*: homology to IS*2* from *Escherichia coli* and IS*426* from *Agrobacterium tumefaciens*. *DNA Seq* **2**: 163–172.

Wheatcroft R & Watson RB (1988) Distribution of insertion sequence IS*Rm1* in *Rhizobium meliloti* and other gram-negative bacteria. *J Gen Microbiol* **134**: 113–121.

Wood DW, Setubal JC, Kaul R, *et al.* (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**: 2317–2323.

Zekri S & Toro N (1996) Identification and nucleotide sequence of *Rhizobium meliloti* insertion sequence IS*Rm6*, a small transposable element that belongs to the IS*3* family. *Gene* **175**: 43–48.

Zekri S, Soto MJ & Toro N (1998) IS*Rm4-1* and IS*Rm9*, two novel insertion sequences from *Sinorhizobium meliloti*. *Gene* **207**: 93–96.