

Identification of Rhizobium-Specific Intergenic Mosaic Elements within an Essential Two-Component Regulatory System of *Rhizobium* Species

MAGNE ØSTERÅS,¹† JOHN STANLEY,² AND TURLOUGH M. FINAN^{1*}

Department of Biology, McMaster University, Hamilton, Ontario L8S 4K1, Canada,¹ and Plasmid and Molecular Genetics Unit, National Collection of Type Cultures, Central Public Health Laboratory, London NW9 5HT, United Kingdom²

Received 12 May 1995/Accepted 2 August 1995

Analysis of the DNA regions upstream of the phosphoenolpyruvate carboxykinase gene (*pckA*) in *Rhizobium meliloti* and *Rhizobium* sp. strain NGR234 identified an open reading frame which was highly homologous to the *Agrobacterium tumefaciens* chromosomal virulence gene product ChvI. A second gene product, 500 bp downstream of the *chvI*-like gene in *R. meliloti*, was homologous to the *A. tumefaciens* ChvG protein. The homology between the *R. meliloti* and *A. tumefaciens* genes was confirmed, because the *R. meliloti* *chvI* and *chvG* genes complemented *A. tumefaciens* *chvI* and *chvG* mutants for growth on complex media. We were unable to construct *chvI* or *chvG* insertion mutants of *R. meliloti*, whereas mutants carrying insertions outside of these genes were readily obtained. A 108-bp repeat element characterized by two large palindromes was identified in the *chvI* and *chvG* intergenic regions of both *Rhizobium* species. This element was duplicated in *Rhizobium* sp. strain NGR234. Another structurally similar element with a size of 109 bp was present in *R. meliloti* but not in *Rhizobium* sp. strain NGR234. These elements were named rhizobium-specific intergenic mosaic elements (RIMEs), because their distribution seems to be limited to members of the family *Rhizobiaceae*. A homology search in GenBank detected six more copies of the first element (RIME1), all in *Rhizobium* species, and three extra copies of the second element (RIME2), only in *R. meliloti*. Southern blot analysis with a probe specific to RIME1 showed the presence of several copies of the element in the genome of *R. meliloti*, *Rhizobium* sp. strain NGR234, *Rhizobium leguminosarum*, and *Agrobacterium rhizogenes*, but none was present in *A. tumefaciens* and *Bradyrhizobium japonicum*.

Repetitive DNA sequence elements were first identified and characterized in eukaryotic genomes. The *Alu* family of repeats found in mammals are the best known of such elements (51). Repeated elements were believed to be absent from the smaller prokaryotic genomes until DNA sequence comparison of *Escherichia coli* and *Salmonella typhimurium* intergenic regions identified an element referred to as the repetitive extragenic palindrome (REP), or palindromic unit (23, 35, 58). This 38-bp element is characterized by a largely conserved inverted repeat flanking a 5-bp variable loop (58). REP can form more complex elements by incorporating different motifs to form bacterial interspersed mosaic elements (BIMEs), whose size can be up to 500 nucleotides long (26). More recently, another family of repetitive elements, called the enterobacterial repetitive intergenic consensus (ERIC), or intergenic repeat unit, was identified again by comparison of *E. coli* and *S. typhimurium* DNA sequences (29, 53). The ERIC element is 126 bp long and contains a large inverted repeat flanking a 4-bp central region. Both the REP and ERIC element families are found in the intergenic regions of genes, not necessarily at the same position between species, and REP can represent up to 1% of the total DNA in the genome (53, 58). These two families of repetitive sequences are widely distributed among

eubacteria, as was shown by hybridization and PCR fingerprinting (62). Although these elements are widely distributed within a genome, their mechanisms of spreading and their putative functions remain obscure. It has been shown that their presence can affect the level of gene expression (41). Another proposed role is in relation to chromosomal structure, because both DNA gyrase and DNA polymerase I can bind to REP elements (4, 25, 68). However, no satisfactory single role explaining the wide distribution and the sequence conservation of these elements has been determined (35).

As part of our study of the regulation of the phosphoenolpyruvate carboxykinase gene (*pckA*) of the alfalfa symbiont *Rhizobium meliloti* and a broad-host-range *Rhizobium* sp. strain, NGR234, we determined the nucleotide sequences of the *pckA* structural genes (42, 43). The transcription start site of the *R. meliloti* *pckA* gene was determined, and expression of *pckA* was shown to be regulated by gluconeogenic substrates such as succinate (42). Here we report the DNA sequences of the regions upstream of the *pckA* genes of *R. meliloti* and *Rhizobium* sp. strain NGR234. We identify two genes in this region which are homologs of the *chvG* and *chvI* genes of *Agrobacterium tumefaciens*. The *A. tumefaciens* *chvG* and *chvI* genes were previously shown to be homologous to the environmental sensor-regulator proteins, and mutations in these genes generate an avirulent phenotype (8, 36). In *R. meliloti*, we were unable to disrupt either of these genes, and thus under our growth conditions, these genes appear to be essential for growth.

Comparison of the two *Rhizobium* DNA sequences revealed the presence of a 108-bp element in the intergenic region between the *chvI* and *chvG* genes, which appears to be present

* Corresponding author. Mailing address: Department of Biology, McMaster University, 1280 Main St. West, Hamilton, Ontario L8S 4K1, Canada. Phone: (905) 525-9140, ext. 22932. Fax: (905) 522-6066. Electronic mail address: finan@mcmil.cis.mcmaster.ca.

† Present address: Laboratoire de Biologie Végétale et Microbiologie, URA CNRS 1114, Université de Nice Sophia-Antipolis, Parc Valrose, 06108 Nice Cedex 02, France.

TABLE 1. Bacterial strains and plasmids used in this study

Strain, plasmid or phage	Relevant characteristic(s)	Source or reference
Strains		
<i>Rhizobium</i> sp.		
Rm1021	<i>R. meliloti</i> SU47, <i>str-21</i>	37
Rm5000	<i>R. meliloti</i> SU47, <i>rif-5</i>	15
Rm5065	Rm1021, <i>pck-1::Tn5-132</i>	17
AK631	<i>R. meliloti</i> 41, <i>exoB631</i>	19
102F34	<i>R. meliloti</i> , wild type	13
RCR2012	<i>R. meliloti</i> , wild type	Rothamsted Experimental Station
NGR234R	<i>Rhizobium</i> sp. strain NGR234, <i>rif-1</i>	56
ANU265	<i>Rhizobium</i> sp. strain NGR234, pSym ⁻	40
GF160	<i>R. leguminosarum</i> bv. <i>viciae</i>	18
<i>Bradyrhizobium</i> sp. strain USDA110	<i>B. japonicum</i> , wild type	32
<i>Agrobacterium</i> sp.		
A348	<i>A. tumefaciens</i> C58 pTiC58 cured, with pTiA6NC	52
At11063	<i>A. tumefaciens</i> C58 <i>chvG::aacI</i>	8
At11064	<i>A. tumefaciens</i> C58 <i>chvI::aacI</i>	8
GMI9023	<i>A. tumefaciens</i> C58 pTiC58 cured, pAtC58 cured, Sm ^r , Rif ^r	47
At125	GMI9023 pRmeSU47b	16
At128	GMI9023 pRmeSU47a	16
At167	<i>A. rhizogenes</i> ATCC 15834	Patrice Dion
<i>E. coli</i>		
DH5 α	F ⁻ <i>endA1 hsdR17 supE44 thi-1 recA1 gyrA96 relA1Δ(argF-lacZYA)</i>	Bethesda Research Laboratories
XL-1Blue	<i>supE44 hsdR17 endA1 gyrA46 relA1 thi recA</i> [F' <i>proAB, lacI^q ΔM15 Tn10(Tc^r)</i>]	7
MT607	<i>pro-82 thi-1 hsdR17 supE44 recA56</i>	16
MT616	MT607 pRK600	16
Plasmids		
pLAFR1	IncP cosmid cloning vector, Tc ^r	21
pRK7813	IncP cosmid cloning vector, Tc ^r	31
pHP45 Ω	pBR322 derivative (Ap ^r) with Ω Sp ^r	20
pPH1JI	IncP Gm ^r Sp ^r Cm ^r	5
R751-pGM2	IncP Sm ^r Gm ^r	30
pBluescript +/-	Derivative of pUC19 with fl(+) or (-) <i>oriR</i>	Stratagene
pUC118	ColE1 <i>oriV</i> cloning vector, Ap ^r	63
pRmT103	pLAFR1, <i>R. meliloti</i> cosmid clone with <i>pckA</i>	17
pTH84	pRK7813, 7.5-kb <i>Bam</i> HI fragment containing <i>R. meliloti pckA, chvI, and chvG</i> genes	42
pMOP5	pRK7813, 5-kb <i>Pst</i> I fragment containing NGR234 <i>pckA</i> and <i>chvI</i> genes	43
pMOP8-1	pBluescript+, 1.3-kb <i>Hind</i> III- <i>Pst</i> I fragment with NGR234 <i>chvI</i>	This work
pMOP8-2	pBluescript-, 1.3-kb <i>Hind</i> III- <i>Pst</i> I fragment with NGR234 <i>chvI</i>	This work
pTH118	pUC118, 1.7-kb <i>Eco</i> RI fragment with <i>R. meliloti chvI</i>	This work
pTH119	pUC118, 1.7-kb <i>Eco</i> RI fragment with <i>R. meliloti chvI</i>	This work
pACL1	pRK7813, <i>A. tumefaciens</i> cosmid clone with <i>chvIG</i>	8
pTH225	pRK7813, 1.7-kb <i>Eco</i> RI fragment from pTH118 with Ω Sp ^r inserted in the <i>Sst</i> I site	This work
Phage M13K07	M13 derivative, helper phage	63

only in members of the family *Rhizobiaceae*. Another element, with a size of 109 bp, in the same region is present in *R. meliloti* only. We have designated these elements as rhizobium-specific intergenic mosaic elements (RIMES) because their size, structure, and location are reminiscent of those of the BIMES of enteric bacteria.

MATERIALS AND METHODS

Bacterial strains, plasmids, and media. The strains and plasmids used in this study are listed in Table 1. Manipulations of the *pckA* gene region in *R. meliloti* were carried out with strain SU47 derivatives. The following media were used for the growth of various bacterial strains: *E. coli*, Luria broth (LB); *R. meliloti* and *A. tumefaciens*, LB containing 2.5 mM CaCl₂ and 2.5 mM MgSO₄; and *Rhizo-*

bium sp. strain NGR234, *Rhizobium leguminosarum*, *Bradyrhizobium japonicum*, and *Agrobacterium rhizogenes*, TY (5, 39). When required, antibiotics were added at concentrations previously described (16).

Genetic techniques and genomic recombination. Bacterial matings were performed as previously described (16). In most experiments with genomic recombination, the tetracycline-resistant (Tc^r) IncP group plasmid pRK7813, containing *R. meliloti* DNA with neomycin-resistant (Nm^r) Tn5 or spectinomycin-resistant (Sp^r) Ω insertions, was introduced into the rifampin-resistant (Rif^r) SU47 derivative Rm5000. The Ruvkun and Ausubel plasmid incompatibility technique (48) was then used to detect strains in which the insertions had recombined from plasmids into the *R. meliloti* genome. Five-milliliter overnight cultures of the donor and recipient strains were centrifuged and resuspended in 0.5 ml of LB before being mixed 1:1. One hundred-microliter aliquots of the mixtures were spotted onto LB plates. After incubation overnight at 28°C, the spots were resuspended in 5 ml of sterile 0.85% NaCl solution. One hundred microliters of the desired dilutions was plated onto selective media. Viable

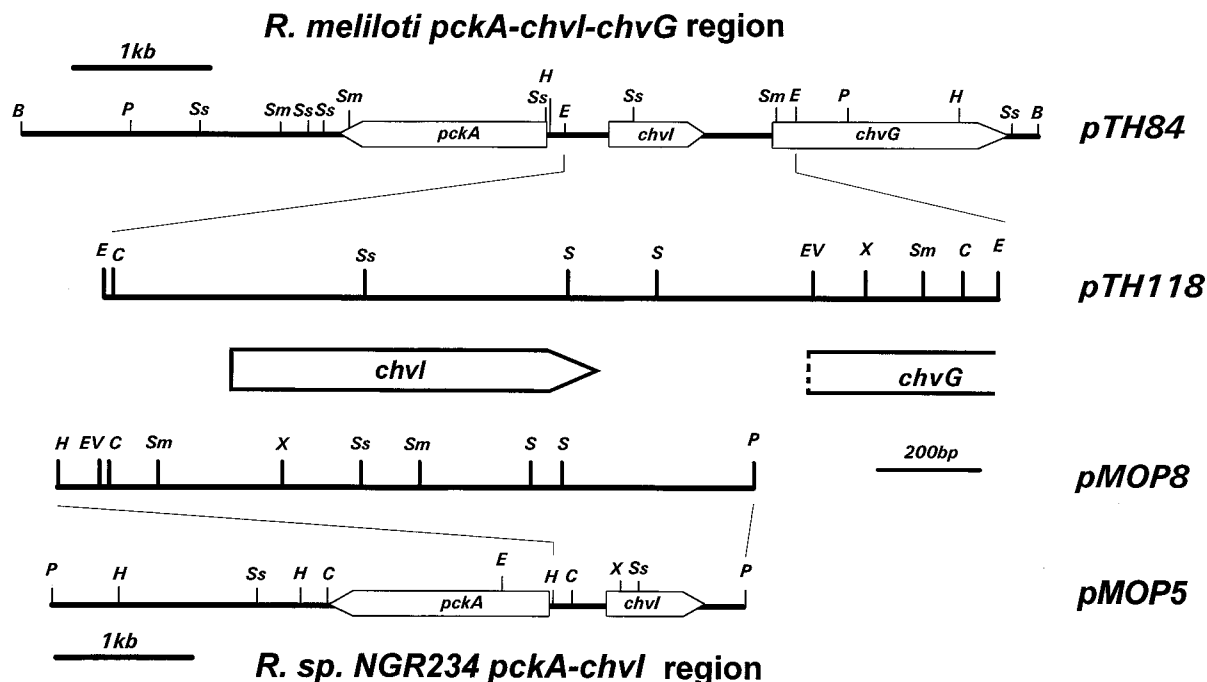


FIG. 1. Physical and genetic map of the plasmids overlapping the region containing *pckA* and the two-component regulatory system *chvI-chvG* in *R. meliloti* (pTH84 and pTH118) and in *Rhizobium* sp. strain NGR234 (pMOP5 and pMOP8). B, *Bam*HI; C, *Cla*I; E, *Eco*RI; EV, *Eco*RV; H, *Hind*III; P, *Pst*I; S, *Sal*I; Sm, *Sma*I; Ss, *Sst*I; X, *Xho*I.

counts for donors and recipients were determined by plating on LB containing gentamicin and LB containing rifampin, respectively. Recombinants carrying Tn5 were selected with the gentamicin and spectinomycin resistance (Gm^r Sp^r) plasmid pPH1J1 (5). The ΩSp^r insertion recombinants were selected after transfer of the gentamicin-resistant (Gm^r) plasmid R751-pGM2 (30).

DNA manipulations. Standard methods were used for plasmid DNA isolation, restriction analysis, agarose and polyacrylamide gel electrophoresis, Southern blotting, DNA ligation, and transformation (49). Bacterial genomic DNA was isolated by the method described previously for *R. meliloti* (37). Hybridizations were performed with digoxigenin-labelled probe (Boehringer Mannheim) for analysis of potential recombinants and with ^{32}P -labelled DNA probe for the detection of RIME1. Unbound probe was removed by washing the filters twice at room temperature for 15 min with $5\times$ SSC ($1\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate (SDS) for low stringency, followed by two washes for 15 min at $65^\circ C$ with $0.1\times$ SSC–0.1% SDS for high stringency. High-stringency washes for the RIME1 probe were done at room temperature because of the small size of the probe (90 bp).

DNA sequencing. The 1.3-kb *Hind*III–*Pst*I fragment containing the region upstream of *pckA* in *Rhizobium* sp. strain NGR234 was subcloned from pMOP5 into pBluescript+ and pBluescript– (pMOP8-1 and pMOP8-2). The 1.8-kb *Eco*RI fragment containing the region upstream of *pckA* in *R. meliloti* was subcloned from pTH84 into pUC118 in both orientations (pTH118 and pTH119). A series of unidirectional nested deletions were isolated for each plasmid by exonuclease III treatment followed by S1 nuclease digestion as described by Sambrook et al. (49). Single-stranded DNA was obtained by using the host strain XL-1Blue and the helper phage M13K07 (63). The universal –20 primer was employed for the annealing procedure for pBluescript+ and pUC118-derived plasmids. The reverse primer was used for pBluescript–-derived plasmids. The Tn5 insertions in pTH84 were subcloned as a *Bam*HI fragment into pUC118, and the site of insertion was determined with a Tn5-specific primer (5'-TCACATGGAAGTCAGATCCT-3' [65]). DNA sequencing was performed by dideoxy chain termination (50) according to the protocol of U.S. Biochemicals for the Sequenase 2.0 enzyme, with [α - ^{35}S]dATP (DuPont, NEN) and 7-deaza dGTP (Pharmacia). Both strands of DNA were sequenced. DNA and derived protein sequences were analyzed with the PC/Genie (Intelligenetics), BLAST (2), and CLUSTAL V (28) software packages.

PCR amplification of the RIME1 element. Two primers complementary to the ends of the element were synthesized (left, 22-mer, 5'-ACTGCATGTTTCCTT TAATCGT-3'; right, 19-mer, 5'-TCTTACCAGACGCGCAAAG-3') (Mobix, McMaster University). Amplification with the two primers would give a 90-bp-long product internal to the element. The PCR was performed on a Perkin-Elmer Cetus thermocycler as described by Sambrook et al. (49). The conditions were 30 cycles with annealing at $48^\circ C$ for 45 s, elongation at $72^\circ C$ for 2 min, and denaturation at $94^\circ C$ for 1 min. The reactions were performed in a final volume

of 50 μ l with 1.5 mM $MgCl_2$ and 2.5 U of *Ampli*Taq (Perkin-Elmer). The template DNA used was pTH118. Because direct amplification of the fragment was unsuccessful, a two-step amplification was performed. The first amplification was performed with the left primer specific to RIME1 and the reverse primer that recognized the vector DNA just outside of the polylinker. The amplified product (590 bp) was diluted 1,000-fold and used for amplification with the left and right primers specific to RIME1. The final amplification product (90 bp) was purified from an 8% nondenaturing acrylamide gel by overnight elution at $4^\circ C$ followed by ethanol precipitation. The fragment was then labelled with [α - ^{32}P] dCTP (ICN) with the left and right primers and Klenow enzyme (Boehringer Mannheim).

Nucleotide sequence accession number. The DNA sequences of the *R. meliloti* and *Rhizobium* sp. strain NGR234 *chvI* and *chvG* genes have been submitted to the GenBank and EMBL data banks and have been assigned accession numbers U32941 and U32869, respectively.

RESULTS

Nucleotide sequence of the region upstream of *pckA*. The DNA sequences of a 1.7-kb *Eco*RI fragment from *R. meliloti* (pTH118) and a 1.3-kb *Hind*III–*Pst*I fragment from *Rhizobium* sp. strain NGR234 (pMOP8) spanning the regions upstream of their respective *pckA* genes were determined (Fig. 1 and 2). Analysis of the six translation frames in the *R. meliloti* sequence revealed the presence of one potential open reading frame 240 amino acids in length which is transcribed divergently from the *pckA* gene. The codon usage for this protein is very similar to what has been determined from other *R. meliloti* genes (6), with a strong G+C bias in the third position (88.3% G+C content compared with 63.6% G+C content in general). The same open reading frame was also present in the *Rhizobium* sp. strain NGR234 sequence but was slightly shorter with 238 amino acid residues. For reasons described below, we have designated this open reading frame as *chvI*. The ATG start codons for *chvI* were located 342 and 336 bp upstream of the *pckA* translational start site in *R. meliloti* and *Rhizobium* sp. strain NGR234, respectively. High levels of homology at the nucleotide level between the two *Rhizobium* species were ob-

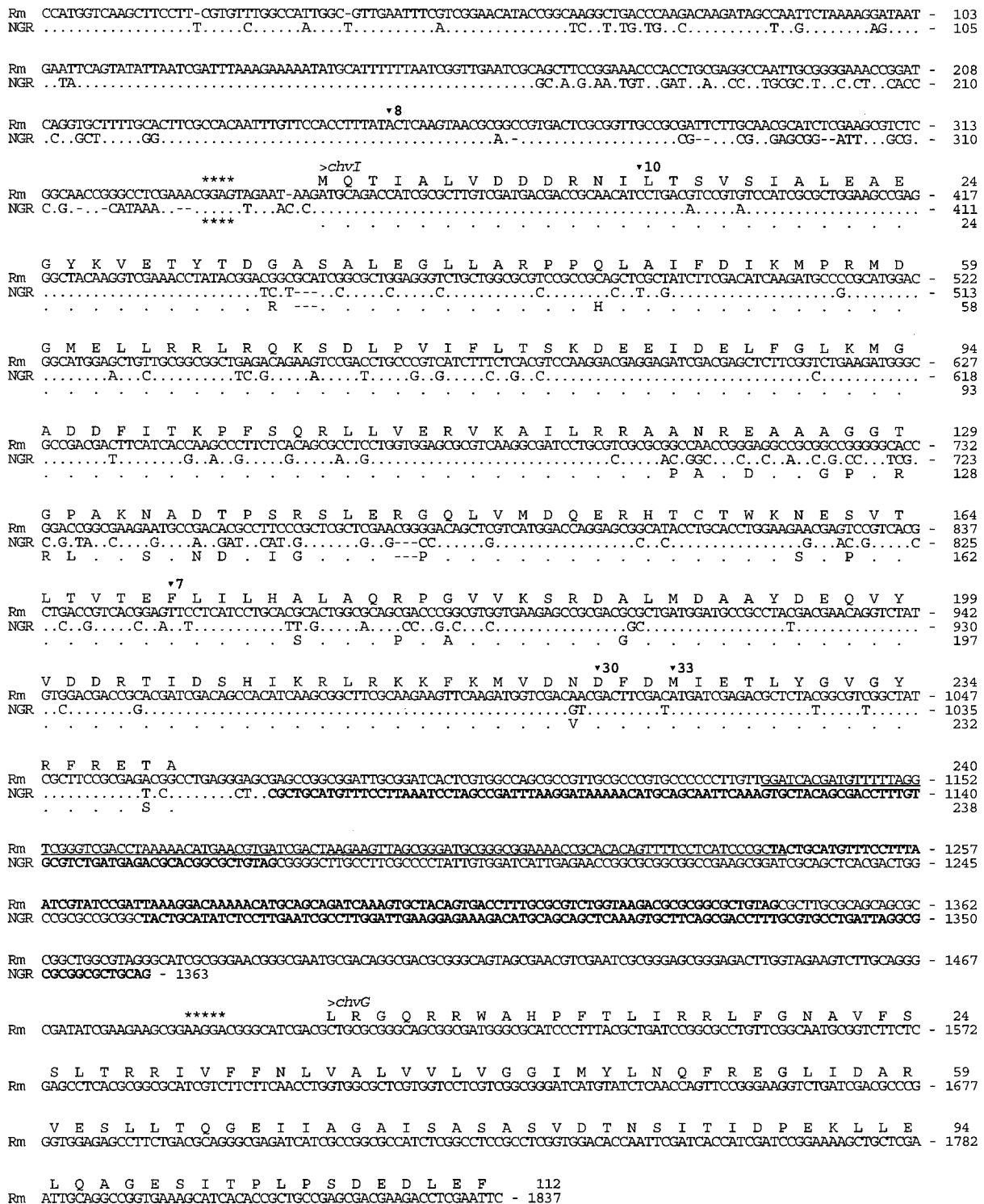


FIG. 2. Alignment of the nucleotide sequences of *R. meliloti* (Rm) and *Rhizobium* sp. strain NGR234 (NGR) from the *Nco*I site containing the *pckA* gene start codon to the second *Eco*RI site in *R. meliloti* (1,837 bp) and to the *Pst*I site in *Rhizobium* sp. strain NGR234 (1,363 bp). The deduced protein translations for ChvI are shown over (*R. meliloti*) and under (*Rhizobium* sp. strain NGR234) the nucleotide sequences. The ChvG protein is given over the *R. meliloti* nucleotide sequence. Conserved nucleotides and amino acids between the two species are replaced by a dot in the *Rhizobium* sp. strain NGR234 sequence until the end of the ChvI coding region (nucleotide 1077 in *R. meliloti*). Gaps, represented by dashes, were inserted to permit the correct alignment. The exact positions of the Tn5 insertions in pTH84 as determined by the DNA sequence are shown by inverted solid triangles followed by the mutation numbers. The repeat elements found in the intergenic region are shown in boldface for RIME1 and are underlined for RIME2. The asterisks indicate putative ribosome binding sites for *chvI* and *chvG*.

A

RmChvI	M-QTIALVDDDRNILTSSVLSIALEAEGYKVVETVYTDGASALEGLLARPPOLA	49
RsChvI-R.....H...	48
AtChvIR.....VD...I...	49
RmPhoB	.LPK..V.E.EEALSIVLLRYN....FE.D.ILR.DE.EIR.QE.L.D.L	50
RmChvI	IFDIKMPRMDGMELLRLRQKSD--LPVIFLTSKDEEIDELPGLKMGAD	96
RsChvII.....	95
AtChvII.....	96
RmPhoB	.L.WML.GVS.I.C.....RPETER.I.M.ARG.SERVR..AT...	100
RmChvI	DFITKPFQSRLLVERVKAILRRAANREAAAG-GTGPAKNADTPSRSLERG	145
RsChvIPA.D..GP..RRL..S.ND..IG..P...	143
AtChvISS..S.AT.GTLKPT..QQA..T...	146
RmPhoB	.YVV....TPE.MA....M.....KPEVL.TL.RC.	135
RmChvI	QLVMDQERHTCTWRNESVTLTVTEFLILHALAQRPGVVKSRDALMDAAVD	195
RsChvIS.P.....S..P.A.....G.....	193
AtChvIA.....G.P.....S..PAA.RGEK.....	196
RmPhoB	DIEL.R.T.RVHRRSRE.R.GP...RL.EF.MSS..R.F..SQ.L.GVWG	185
RmChvI	EQVYVDDRTIDSHIKRLRKKFKMVDNDFMIETLYGVGYFRFETA	240
RsChvIV.....S.....	238
AtChvIL.G.....A.A.....	241
RmPhoB	HDI...E..V.V.VG...ALNFSNMP-.V.R.VR.A..SL--ES	227

B

RmChvG	LRGQRRWAHPFTLIRRLFGNAVFSLSLRRIVFFNLVALVVLVGGIM	46
AtChvG	MPKNRSER.HRI..L.I...I.....L..VA.T.....L	50
RmChvG	YLNQFREGILIDARVESLSLTQGEIIAGAISASASVDINSITIDPEKLLLELQ	96
AtChvGV.R.....N.....	100
RmChvG	AGESITPLPSDEDLDF	112
AtChvG	..Q...A.N...S.	116

FIG. 3. (A) Protein alignment of *R. meliloti* and *Rhizobium* sp. strain NGR234 ChvI (RmChvI and RsChvI, respectively) with *A. tumefaciens* ChvI (AtChvI) and *R. meliloti* PhoB (RmPhoB). Amino acid residues in RsChvI, AtChvI, and RmPhoB conserved with RmChvI are indicated by dots. Dashes indicate gaps inserted to permit a correct alignment. (B) Protein alignment between *R. meliloti* ChvG (RmChvG) N-terminal fragment and the *A. tumefaciens* ChvG (AtChvG) N-terminal region.

served in the *chvI* promoter (nucleotides 216 to 346 in *R. meliloti*) and in the two *chvI* coding regions. Within the *chvI* translated sequences, about one-third of the codons showed base changes. Seventy percent of the changes occurred in the third position, and overall 70% of the changes did not alter the encoded amino acid (silent mutations). This is consistent with the profile one expects from a translated gene. Strikingly, the high level of homology between the two *chvI* genes was lost immediately after the 3' end of the genes (nucleotide 1070 in *R. meliloti*) (Fig. 2).

Characterization of the *Rhizobium* ChvI proteins. A comparison of the amino acid sequences of the *chvI*-encoded proteins from *R. meliloti* and *Rhizobium* sp. strain NGR234 revealed that they are highly conserved (89%), especially in the N- and C-terminal regions (residues 1 to 119 and 146 to 240, with 97 and 92% identity, respectively). The 30-amino-acid central region (residues 120 to 145) shows the highest level of variability between the two species (Fig. 3A). A search of

GenBank with the BLASTP program revealed that the two ChvI proteins showed homology to regulatory proteins of the large family of two-component regulator-sensor systems (59). These regulators are defined by a conserved domain of about 100 amino acids at the N-terminal region and are divided into subfamilies on the basis of the similarity observed in the C-terminal region. The size and sequence of the rhizobium ChvI proteins indicate that they belong to the group of regulators thought to interact with promoters recognized by the Eσ70 polymerase. The highest level of homology was found to the *A. tumefaciens* ChvI protein. *A. tumefaciens* is classified as a member of the *Rhizobiaceae* family. The *A. tumefaciens* ChvI protein is believed to be a response regulator of a two-component regulatory system recently characterized in *A. tumefaciens* (36). Among the other response regulators, *R. meliloti* PhoB was the next most homologous protein to *R. meliloti* ChvI identified. The *R. meliloti* ChvI and PhoB proteins were 44% identical, whereas *R. meliloti* ChvI and *A. tumefaciens* ChvI were 85% identical (Fig. 3A).

In *A. tumefaciens*, the sensor protein, ChvG, of this two-component system is encoded by a gene located 100 bp downstream of *chvI* (8). In *R. meliloti*, we found an incomplete open reading frame which started 400 bp downstream of the *chvI* gene. The translated sequence encoded the N terminus of a protein which showed 80% amino acid identity to the N terminus of the *A. tumefaciens* ChvG protein (Fig. 3B). In view of the homology and similar location to the *A. tumefaciens* *chvG* gene, we designated the *R. meliloti* gene as *chvG*. An unusual feature of this gene was the absence of either the usual initiation codon ATG or the rare initiation codons GTG and TTG. A verification of the DNA sequence in this region did not detect any mistakes. Instead, the sequence showed CTG at nucleotide 1532 as the probable start codon (Fig. 2).

The high level of homology at the DNA sequence level between the *Rhizobium* regulator and *A. tumefaciens* *chvI* suggested that these two genes were identical. To confirm this, we introduced the cosmid clone pACL1, which carries the *A. tumefaciens* *chvIG* genes into the *R. meliloti* *pckA* mutant Rm5065 (17). Unlike Rm5065, the Rm5065 (pACL1) transconjugants grew on succinate, showing that this plasmid, pACL1, complemented the *pckA* mutation. Thus, on the basis of gene organization, similarity in DNA and protein sequences, and functional complementation (described below), we have designated the *Rhizobium* genes upstream of *pckA* as *chvI* and *chvG*.

Attempted construction of *chvIG* mutants. To help identify the function of the *chvIG* genes, we wished to recombine plasmid-borne Tn5 and ΩSp insertions located in different positions across the *pckA-chvI-chvG* region into the *R. meliloti* genome (7, 8, 10, 16, 17, 21, 27, 31, 34, 36, 44, and 50) (Fig. 4 and pTH225). Many of the insertions generated few recombinants, and Southern blot analysis revealed that only Tn5 insertions 8, 35, and 43, which lay outside of *chvIG*, formed true

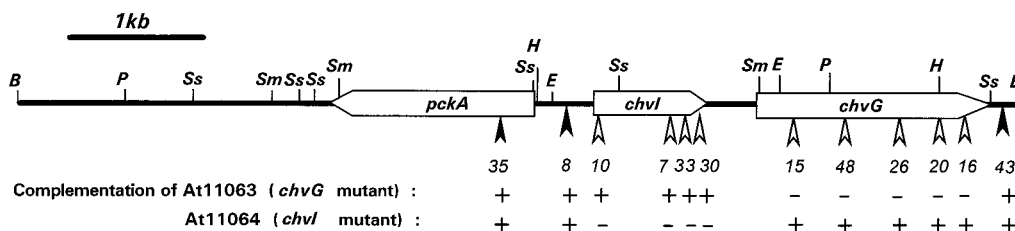


FIG. 4. Location on the pTH84 plasmid of the Tn5 insertions (arrowheads) used for the recombination attempts and for complementation. The complementation phenotypes of the *A. tumefaciens* *chvG* (At11063) and *chvI* (At11064) mutants are indicated by + for restoration of growth on complex medium or - for no complementation. B, BamHI; P, PstI; Ss, SstI; Sm, SmaI; H, HindIII; E, EcoRI.

recombinants. All other putative recombinants retained the wild-type *chvI* and *chvG* genes.

Because *chvI* and *chvG* mutants of *A. tumefaciens* were shown to be unable to grow on complex medium, we also attempted all of the gene disruption experiments described above with M9 minimal medium. Again, no recombinants carrying insertions in the *chvI* or *chvG* genes were detected.

Complementation of *A. tumefaciens chvI* and *chvG* mutants. The pLAFR1 cosmid pRmT103, which contains *R. meliloti pckA* and the upstream genes identified in the present study, was able to complement *A. tumefaciens chvI* and *chvG* mutants (8). In order to confirm that the complementation of *A. tumefaciens chvI* and *chvG* mutants by pRmT103 was specifically caused by the *R. meliloti chvIG* homologs, the ability of pTH84 and the pTH84 derivatives with Tn5 insertions located in different positions across the *pckA-chvI-chvG* region (42) (Fig. 4) to complement At11063 and At11064 for growth on complex medium was determined. pTH84 insertions 7, 10, 30, and 33 complemented At11063 (*chvG* mutant) but failed to complement At11064 (*chvI* mutant) for growth on complex medium. Similarly, insertions 15, 17, 20, 26, and 48 were unable to complement At11063 but complemented At11064 for growth on complex medium. Insertions 8, 35, and 43 were able to complement both mutants (Fig. 4). These results, together with the sequence data, establish that the *R. meliloti* genes referred to as *chvI* and *chvG* are in fact the same genes as those present in *A. tumefaciens*. In addition, the data suggest that *chvI* and *chvG* are transcribed independently of each other in *R. meliloti*.

Identification of a repeated element in the intergenic region between *chvI* and *chvG*. The sensor gene *chvG* is located 100 bp downstream of *chvI* in *A. tumefaciens* (8). In both *Rhizobium* species, the intergenic region is significantly larger, with more than 300 bp in NGR234 and 400 bp in *R. meliloti*. Inspection of the DNA sequence in the 3' noncoding region of the two *Rhizobium chvI* genes showed no conservation. This contrasts with the high degree of conservation observed between the promoter regions of the *Rhizobium chvI* genes and both the promoter and termination regions of their *pckA* genes (42). Further analysis of this intergenic region in the *Rhizobium* species revealed a complex secondary structure with several large inverted repeats (Fig. 2). A dot plot alignment between the two *Rhizobium* intergenic sequences revealed a conserved 108-bp region. This region is duplicated in NGR234. The relative positions of the 108-bp elements within the intergenic regions differ (Fig. 5 and 6A). The repeat element is characterized by two large and different palindromic sequences of 48 and 44 bp that are AT rich (38% G+C) and GC rich (61% G+C), respectively, which flank a central conserved region with a length of 12 bp (Fig. 6A). A BLASTN search with this 108-bp element specifically detected six other similar sequences in the data bank. All were located in DNA sequences obtained from *Rhizobium* species. One element was located 59 bp upstream of the *rpoN* gene of *Rhizobium* sp. strain NGR234 (61); four were found in *R. meliloti*, 188 bp downstream of *ureC* (38) and 11, 23, and 168 bp upstream of *actR* (GenBank accession no. L39938), *cya2* (EMBL accession no. X80991) (3), and *phbC* (GenBank accession no. U17227), respectively; and the sixth was present 62 bp downstream of *glnA* in *R. leguminosarum* (9). The latter element was incomplete. The orientation of the elements in regard to the direction of transcription does not seem to be relevant, because six copies are present in one orientation and the other three copies are present in the opposite orientation (RmphbC, RsrpoN, and RlglmA) relative to the closest transcribed gene (Fig. 5). In all cases, the homology observed between the elements did not extend to

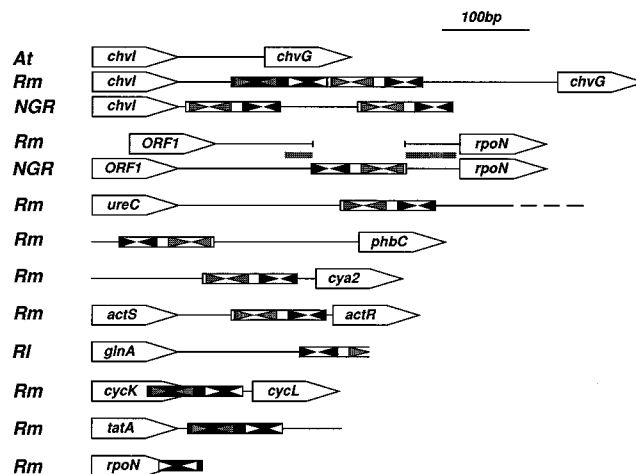


FIG. 5. Physical position and orientation of the identified repeat elements in relation to their neighboring genes. The boxed genes indicate the orientation of transcription and the position of the 5' or 3' end of the gene. The gray boxes between the *rpoN* promoters of *R. meliloti* and *Rhizobium* sp. strain NGR234 indicate homologous sequence between the two species as described in the legend to Fig. 6C. RIME1 has no background shading, and RIME2 has a black background. At, *A. tumefaciens*; Rl, *R. leguminosarum*; Rm, *R. meliloti*; Rs, *Rhizobium* sp. strain NGR234.

the flanking sequences (Fig. 6A). Interestingly, conservation of the palindromic sequences must be significant to the element, because several of the base changes observed within one arm of the inverted repeats also had compensatory base modifications in the other arm (see circled residues in Fig. 6A). Two of these elements have a large deletion either in the left palindrome (RmcyA2) or in the central region (RmactR), which does not affect the overall structure of the element (Fig. 6A).

A second element is present in the *R. meliloti chvIG* intergenic region. The RIME represents one-fourth of the intergenic region between the *R. meliloti chvI* and *chvG* genes and cannot explain by itself the large difference in size observed between *R. meliloti* and *A. tumefaciens*. The sequence determined downstream of *chvI* in *Rhizobium* sp. strain NGR234 is constituted almost entirely by the two RIME copies. Two large palindromes (each 48 bp) located upstream of the RIME in *R. meliloti* showed a structure similar to that of this element, although the DNA sequences were very different (Fig. 2). A BLASTN search with this 109-bp fragment (nucleotides 1133 to 1241) identified three homologous sequences, all in *R. meliloti*, suggesting that this is also a repeat element (Fig. 5 and 6B). One is located within the *cycHJKL* genes, more precisely between *cycK* and *cycL*, and appears to contain the stop codon for the *cycK* gene (accession no. X82560). Another was found 5 bp downstream of the aminotransferase structural gene *tatA* (45, 65), and the third (incomplete) was downstream of the *rpoN* gene in strain 104A114 (GenBank accession no. M28846). The latter also overlapped the end of the coding region of the gene (Fig. 6B). This element was named RIME2 to separate it from the first element, renamed RIME1, as described above. The available data about RIME2 are limited to only three copies, all in one species, but it seems to share most of the characteristics observed for RIME1. Some mutations in the palindromic sequences were compensated for by another modification, and no homology could be found in the regions flanking RIME2. The central region showed some variation in size (11 to 15 bp).

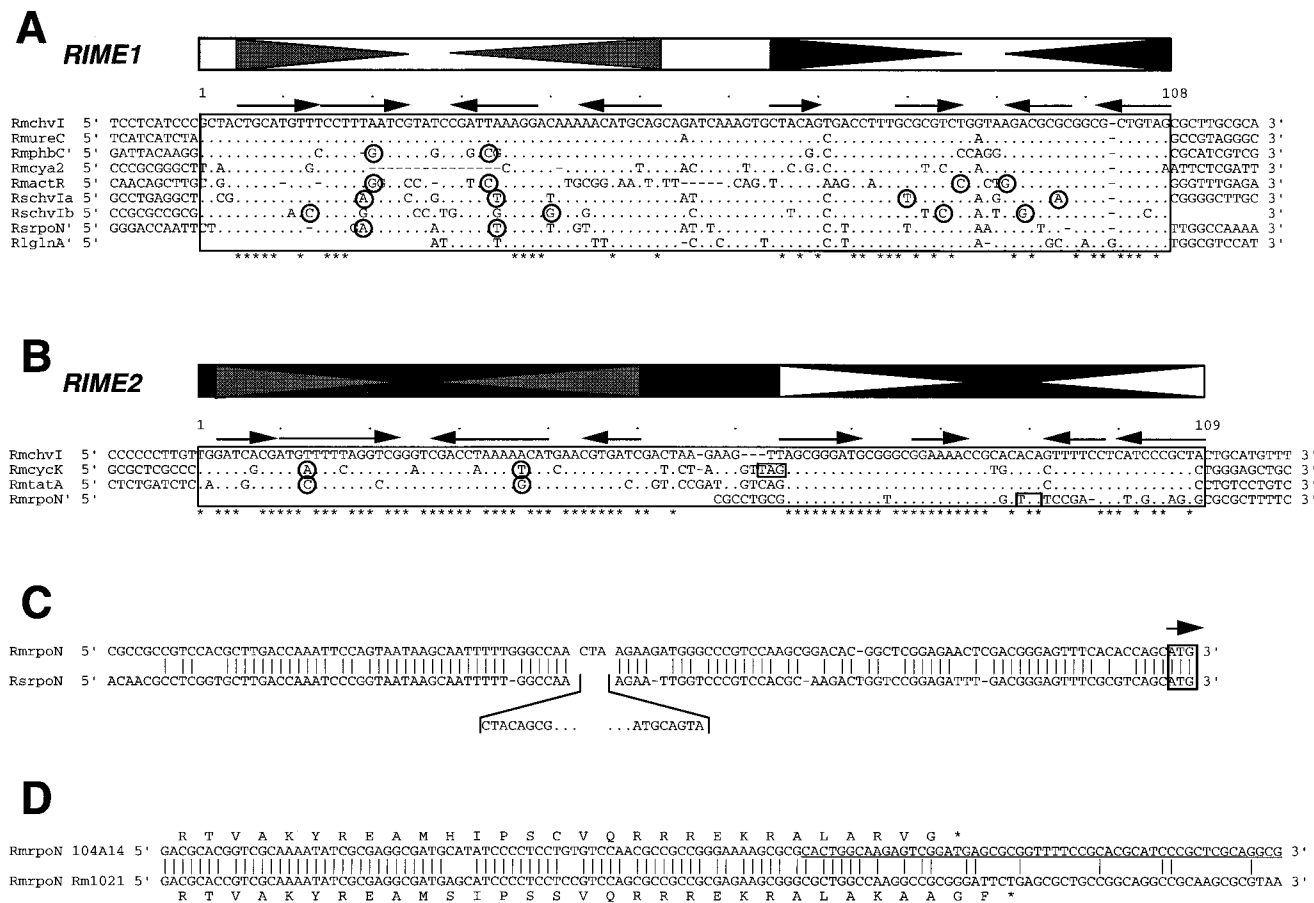


FIG. 6. Sequence alignment of the nine RIME1 (A) and four RIME2 (B) identified, with 10 bp of flanking sequence on each side. All residues conserved in relation to *R. meliloti* (Rm) ChvI are shown by dots. Rl, *R. leguminosarum*; Rs, *Rhizobium* sp. strain NGR234. Stars under the alignment indicate nucleotides conserved in all elements. The two palindromic sequences are indicated by arrows. Base changes that conserve the inverted repeat are circled. The stop codons of *cycK* and *rpoN* located within RIME2 are boxed. (C) DNA sequence alignment of the promoter region of *rpoN* in *R. meliloti* (1) and *Rhizobium* sp. strain NGR234 (61) indicating the insertion position of the element in the NGR234 sequence. The start codon (ATG) and the orientation of *rpoN* are shown by an arrow. The elements oriented in the opposite orientation to the closest gene are indicated by a prime. (D) DNA sequence alignment of the 3' end of the *rpoN* genes in *R. meliloti* 104A14 (54) and 1021 (46) with the translated proteins. The fragment of RIME2 present in strain 104A14 is underlined.

Multiple copies of the 108-bp RIME1 are present in several species of the family Rhizobiaceae. The sequence homology search of the DNA sequence data bank detected this repeated element only in *Rhizobium* species. The absence of homologous sequence in *E. coli* and some other eubacteria in which a much larger fraction of the genome has been sequenced compared with *Rhizobium* species, suggested that this element did not have a broad distribution among prokaryotes. To investigate this further, Southern blots of *EcoRI*-restricted total DNA from several bacterial species within the *Rhizobiaceae* family and from *E. coli* were hybridized with a ³²P-labelled 90-bp fragment from within the element. This fragment was amplified in two steps by PCR with primers from both ends of the element and from the vector (see Materials and Methods). Several attempts to directly amplify RIME1 by changing the PCR conditions were unsuccessful. The cause of this failure is unclear, because no difficulties were encountered when PCR was performed with one primer specific to RIME1 and another primer specific to the vector DNA located either 590 or 1,230 bp from the element. A second amplification with the product of the first PCR, in which the flanking DNA on one side of the element was absent, and the RIME1-specific primers gave the expected fragment. The autoradiograph results obtained when the hybridized filters were washed under low-stringency con-

ditions (5× SSC) or high-stringency conditions (0.1× SSC) were similar. The data showed that up to 21 copies of RIME1 are present in the genomes of *R. meliloti* and *Rhizobium* sp. strain NGR234, 7 copies are found in that of *R. leguminosarum*, 4 copies are found in that of *A. rhizogenes*, and one homologous sequence is present in that of *E. coli*. No hybridization signal was obtained with DNA from *B. japonicum* or *A. tumefaciens* (Fig. 7). Similar results were obtained when total DNA digested with *PstI* was probed with the element (data not shown).

R. meliloti SU47, like many *R. meliloti* strains, contains two large megaplasmids named pRmSU47a (pNOD) and pRmSU47b (pEXO) that represent a large fraction of the genome (30 to 40%). Each plasmid has been mobilized into *A. tumefaciens*, and total DNA from the resulting hybrid strains can be used to determine on which replicon a DNA fragment is located by hybridization with a specific probe (16). This methodology applied to RIME1 revealed that this element is also present on the megaplasmids, with one copy on pNOD and up to five copies on pEXO (data not shown). Similarly, comparison between the *PstI* restriction fragment length polymorphisms obtained with total DNA from NGR234R and those obtained with DNA from its pSym-cured derivative, ANU265, showed

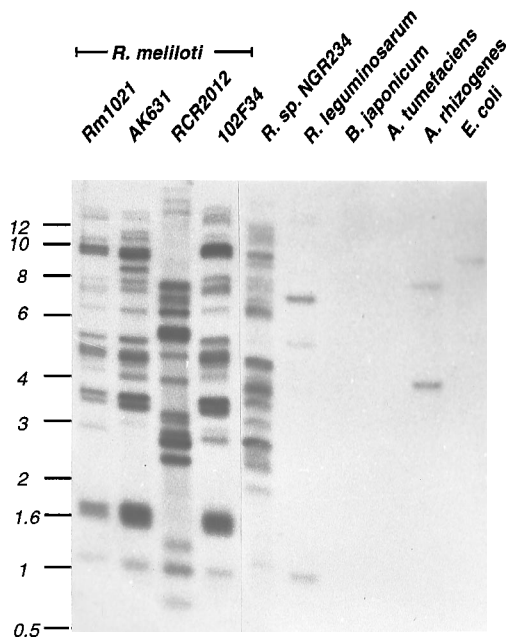


FIG. 7. Southern blot of *Eco*RI-restricted DNA from several members of the family *Rhizobiaceae* and from *E. coli*, probed with a 32 P-labelled fragment internal to RIME1. The numbers at the left of the autoradiogram are marker sizes given in kilobases. The strains used were, in addition to the four *R. meliloti* strains, ANU265 for *Rhizobium* sp. strain NGR234, GF160 for *R. leguminosarum*, USDA110 for *B. japonicum*, A348 for *A. tumefaciens*, ATCC 15834 for *A. rhizogenes*, and DH5 α for *E. coli* (see Table 1).

no visible difference (data not shown), which suggests that no RIME is present on the 500-kb pSym plasmid.

DISCUSSION

In this paper, we have characterized the DNA region which lies upstream of the phosphoenolpyruvate carboxykinase structural gene, *pckA*, in two *Rhizobium* strains. The strain examined in most detail, *R. meliloti* SU47, was originally isolated from an alfalfa root nodule in New South Wales, Australia (64). The other isolate, *Rhizobium* sp. strain NGR234, has a very broad host range and was isolated from a root nodule of *Lablab purpureus* in New Guinea (60). These two strains were shown to be genetically related, although their symbiotic properties are different (55). We found that the region upstream of *pckA* in both strains was highly conserved and encoded a protein homologous to *A. tumefaciens* ChvI. The DNA sequence from *R. meliloti* extended further than that from *Rhizobium* sp. strain NGR234, and analysis of the additional sequence revealed the N terminus of an open reading frame whose predicted protein was homologous to *A. tumefaciens* ChvG. The sequence data together with the observation that the *R. meliloti* genes complemented the growth phenotype of *A. tumefaciens* *chvIG* mutants demonstrated that the *Rhizobium* genes were homologs of *chvI* and *chvG*. The homology of this gene pair to the environmental sensor (*chvG*)-response regulator (*chvI*) family of two-component regulators has been previously described and discussed (8, 36). Neither the stimulus sensed by ChvG nor the genes regulated by ChvI have been identified. In view of the location of the *chvIG* genes, we were particularly interested in determining the effect of mutations in these genes on the regulation of *pckA* gene expression in *R. meliloti*. Unfortunately, despite numerous attempts with both complex and minimal media, we failed to isolate recombinants in which the

chvI or *chvG* genes were disrupted. Moreover, recombinants carrying insertions on either side of the *chvIG* genes or within the *pckA* gene were readily obtained. We conclude that under our experimental conditions, *R. meliloti* *chvG* and *chvI* are essential for the viability of *R. meliloti* cells. We have included the phrase "under our experimental conditions" because of the possibility that the antibiotics employed in our selections may kill the desired recombinants. This is particularly possible in view of the known detergent- and antibiotic-sensitive phenotype of *A. tumefaciens* *chvI* and *chvG* mutants (8, 36). The observation that *A. tumefaciens* *chvIG* mutants grow like the wild type on minimal medium containing succinate (data not shown) suggests that *pckA* expression is not dependent on the *chvIG* products.

In contrast to the high level of homology observed within the coding regions, the intergenic region is much larger in *Rhizobium* species than in *Agrobacterium* species. Lying between the *chvI* and *chvG* genes of *Rhizobium* sp. were two elements which in overall structure are very similar to the repeat elements described for enteric bacteria (35). These elements are 108 and 109 bp long and consist of two inverted repeats flanking a small central region. The presence of two palindromic sequences differs from such elements as REP and ERIC that contain one large inverted repeat (23, 29, 53, 58). The new elements are closer in structure to the BIME, which consists of several palindromic units (26), and were named RIMES. RIMES share most of the characteristics of short interspersed repeat elements. The nine RIME1 identified by DNA sequence are all located outside coding regions, either upstream or downstream of a gene, and their orientation in relation to the direction of transcription seems to be irrelevant. Unlike what was observed with RIME1, two of the four RIME2 appear to overlap the 3' end of the closest coding region. The two inverted repeats are important because a base pair change in one palindromic region is usually accompanied by a compensatory modification in the symmetric portion of the sequence (Fig. 6A), as has been previously observed in REP and ERIC elements (24, 29). The two distinct RIMES are linked in the *chvIG* intergenic region, but because this is not observed in any of the other cases observed in this work, we therefore assume that RIME1 and RIME2 are independent of each other.

The RIME1 were found by DNA sequence homology search with BLASTN in three different *Rhizobium* species only: *R. meliloti*, *Rhizobium* sp. strain NGR234, and *R. leguminosarum* bv. *viciae*. By Southern blotting and hybridization, the RIME1 was detected, in addition to the three species previously mentioned, in *A. rhizogenes*, but not in *A. tumefaciens* or *B. japonicum*. By comparison of this observation with those found through phylogenetic analysis of members of the family *Rhizobiaceae* performed by 16S rRNA nucleotide sequence or RFLP (33, 67), the RIME1 seem to be limited to a group of related species, in which *R. meliloti* is closely linked to *Rhizobium* sp. strain NGR234 and *R. leguminosarum* bv. *viciae* is linked to *A. rhizogenes*. The number of copies of RIME1 detected by hybridization also follows this order, because more signals (>20) were obtained in the *R. meliloti* subgroup than in *R. leguminosarum* (4–6). The smaller number of copies detected in the *R. leguminosarum* subgroup and the apparent absence of RIME1 in *A. tumefaciens* and *B. japonicum* could also be explained by a higher level of divergence in the consensus sequence of such elements compared with the *R. meliloti* RIME1 used as a probe, which then would not hybridize. Effectively, there seems to be a strain-specific conservation of the consensus sequence of the repeat element that is well documented for REP between *E. coli* and *S. typhimurium* (28). Detailed questions regarding the distribution, copy number,

and sequence conservation of RIME will not be resolved without a more complete survey. In addition, the number of available sequences are as yet too few to identify any clear differences between the RIME sequences of *R. meliloti* and *Rhizobium* sp. strain NGR234. The one, incomplete *R. leguminosarum* RIME sequence shows some differences from the *R. meliloti* and *Rhizobium* sp. strain NGR234 sequences, especially in the right palindrome.

The distribution of repetitive extragenic elements within the genome is usually random (12), and the known positions of the RIME sequences described in this study also suggest diverse gene locations (*actR*, *chv1G*, *cyo2*, *cycK*, *glnA*, *phbC*, *rpoN*, *ureC*, and *tataA*). However, only a few copies of RIME1 were detected on the megaplasmids of *R. meliloti*, and none appeared to be present on the 500-kb pSym of NGR234. This suggests that the RIME sequences are not equally distributed between the replicons in *Rhizobium* species. We speculate that dispersion of RIME1 occurred in the chromosome prior to the acquisition of these plasmids by the bacteria. The latter possibility is particularly interesting in view of the poorly understood origins of the megaplasmids of *R. meliloti*.

The mechanism of dissemination of bacterial intergenic mosaic elements is not known. In this respect, we were fortunate to find a RIME1 present in the *rpoN* promoter region of *Rhizobium* sp. strain NGR234, while no such element was present in the same region of *R. meliloti* (Fig. 6C). Comparison of these two sequences reveals no duplication of a possible RIME1 target sequence. Similar cases have been observed for the other repeat elements, which suggests that their mechanism of distribution differs from the one used by insertion elements which involve duplication of flanking base pairs at the insertion site (24). An explanation presently considered is duplication by gene conversion (27), but no evidence is yet available. Although the DNA sequences of RIME1 and RIME2 are different, the homology observed in their structure could be caused by a similar mode of propagation or conservation. The high level of homology observed between different strains of *R. meliloti* would indicate a very low frequency in the dispersion of RIME. Our data were similar to previous results from phylogenetic analysis of *R. meliloti* strains performed with *ISRm1*-specific probes, multilocus enzyme electrophoresis, or PCR fingerprinting (11, 14, 66).

RIME is not the first example of a repeat element apparently limited to a genus because it has been observed in *Neisseria*, *Myxococcus*, and *Deinococcus* species (10, 22, 34). The potential use of such sequence in the identification of specific strains by DNA fingerprinting has already been mentioned (35) and demonstrated by REP- and ERIC-specific PCR fingerprinting (62). Similar methodologies are regularly exploited to detect and monitor microorganisms in the environment (57) and have recently been applied to *Rhizobium* species (11). RIME may have the advantage of being a rhizobium-specific sequence and can then be used in environmental studies without interference from nonrelated bacteria. Furthermore, and as shown in Fig. 7, RIME has a very high level of potential for phylogenetic analysis of closely related *Rhizobium* strains because of its limited distribution and because of the variations observed between strains of the same species, a level at which 16S rRNA comparison cannot be used.

ACKNOWLEDGMENTS

We are grateful to Brian Golding for helpful discussions and advice about PCR techniques and T. Charles for bacterial strains and plasmids. We thank R. J. Watson for critical comments on the manuscript and for sharing unpublished results.

This work was supported by operating grants from the Natural Sciences and Engineering Research Council of Canada to T.M.F. M.Ø. was supported by a postdoctoral fellowship from the Fonds National Suisse de la Recherche Scientifique. This work was initially supported by a fellowship under the OECD Project on Food Production and Preservation to T.M.F.

REFERENCES

- Albright, L. M., C. W. Ronson, B. T. Nixon, and F. M. Ausubel. 1989. Identification of a gene linked to *Rhizobium meliloti ntrA* whose product is homologous to a family to ATP-binding proteins. *J. Bacteriol.* **171**:1932-1941.
- Altschul, S. F., W. Gish, W. Miller, E. W. Meyers, and D. J. Lipman. 1990. Basic local alignment research tool. *J. Mol. Biol.* **215**:403-410.
- Archdeacon, J., J. Talty, B. Boeston, and F. O'Gara. 1995. Cloning of a second adenylate cyclase (*cyo2*) gene from *Rhizobium meliloti* F34: sequence similarity to eukaryotic cyclases. *FEMS Microbiol. Lett.* **128**:177-184.
- Bachelier, S., W. Saurin, D. Perrin, M. Hofnung, and E. Gilson. 1994. Structural and functional diversity among bacterial interspersed mosaic elements (BIMEs). *Mol. Microbiol.* **12**:61-70.
- Beringer, J. E., J. L. Beynon, A. V. Buchanan-Wollaston, and A. W. B. Johnston. 1978. Transfer of the drug resistance transposon Tn5 to *Rhizobium*. *Nature (London)* **276**:633-634.
- Brown, C. M., P. A. Stockwell, M. E. Dalphin, and W. P. Tate. 1994. The translational termination signal database (TransTerm) now also includes initiation contexts. *Nucleic Acids Res.* **22**:3620-3624.
- Bullock, W. O., J. M. Fernandez, and J. M. Short. 1987. XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with β -galactosidase selection. *BioTechniques* **5**:376-378.
- Charles, T. C., and E. W. Nester. 1993. A chromosomally encoded two-component sensory transduction system is required for virulence of *Agrobacterium tumefaciens*. *J. Bacteriol.* **175**:6614-6625.
- Colonna-Romano, S., A. Riccio, M. Guida, R. Defez, A. Lamberti, M. Icarino, W. Arnold, U. Priefer, and A. Puhler. 1987. Tight linkage of *glnA* and a putative regulatory gene in *Rhizobium leguminosarum*. *Nucleic Acids Res.* **15**:1951-1964.
- Correia, F. F., S. Inouye, and M. Inouye. 1986. A 26-base-pair repetitive sequence specific for *Neisseria gonorrhoeae* and *Neisseria meningitidis* genomic DNA. *J. Bacteriol.* **167**:1009-1015.
- de Bruijn, F. J. 1992. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Appl. Environ. Microbiol.* **58**:2180-2187.
- Dimri, G. P., K. E. Rudd, M. K. Morgan, H. Bayat, and G. F.-L. Ames. 1992. Physical mapping of repetitive extragenic palindromic sequences in *Escherichia coli* and phylogenetic distribution among *Escherichia coli* strains and other enteric bacteria. *J. Bacteriol.* **174**:4583-4593.
- Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski. 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. USA* **77**:7347-7351.
- Eardly, B. D., L. A. Materon, N. H. Smith, D. A. Johnson, M. D. Rumbaugh, and R. K. Selander. 1990. Genetic structure of natural populations of the nitrogen-fixing bacterium *Rhizobium meliloti*. *Appl. Environ. Microbiol.* **56**:187-194.
- Finan, T. M., E. Hartweg, K. LeMieux, K. Bergman, G. C. Walker, and E. R. Signer. 1984. General transduction in *Rhizobium meliloti*. *J. Bacteriol.* **159**:120-124.
- Finan, T. M., B. Kunkel, G. F. DeVos, and E. R. Signer. 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J. Bacteriol.* **167**:66-72.
- Finan, T. M., I. Oresnik, and A. Bottacin. 1988. Mutants of *Rhizobium meliloti* defective in succinate metabolism. *J. Bacteriol.* **170**:3396-3403.
- Finan, T. M., J. M. Wood, and D. C. Jordan. 1981. Succinate transport in *Rhizobium leguminosarum*. *J. Bacteriol.* **148**:193-202.
- Forrai, T., E. Vincze, Z. Banfalvi, G. B. Kiss, G. S. Randhawa, and A. Kondorosi. 1983. Localization of symbiotic mutations in *Rhizobium meliloti*. *J. Bacteriol.* **153**:635-643.
- Frey, J., and H. M. Krisch. 1985. Ω mutagenesis in gram-negative bacteria: a selectable interposon which is strongly polar in a wide range of bacterial species. *Gene* **36**:143-150.
- Friedman, A. M., S. R. Long, S. E. Brown, W. J. Buikema, and F. M. Ausubel. 1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. *Gene* **18**:289-296.
- Fujitani, S., T. Komano, and S. Inouye. 1991. A unique repetitive DNA sequence in the *Myxococcus xanthus* genome. *J. Bacteriol.* **173**:2125-2127.
- Gilson, E., J. M. Clément, D. Brutlag, and M. Hofnung. 1984. A family of dispersed repetitive extragenic palindromic DNA sequences in *E. coli*. *EMBO J.* **3**:1417-1421.
- Gilson, E., J. M. Clément, D. Perrin, and M. Hofnung. 1987. Palindromic units: a case of highly repetitive DNA sequences in bacteria. *Trends Genet.* **3**:226-230.

25. Gilson, E., D. Perrin, and M. Hofnung. 1990. DNA polymerase I and a protein complex bind specifically to *E. coli* palindromic unit highly repetitive DNA: implications for bacterial chromosome organization. *Nucleic Acids Res.* **18**:3941–3952.
26. Gilson, E., W. Saurin, D. Perrin, S. Bachelier, and M. Hofnung. 1991. Palindromic units are part of a new bacterial interspersed mosaic element. *Nucleic Acids Res.* **19**:1375–1383.
27. Higgins, C. F., R. S. McLaren, and S. F. Newbury. 1988. Repetitive extragenic palindromic sequences, mRNA stability and gene expression: evolution by gene conversion? A review. *Gene* **72**:3–14.
28. Higgins, D. G., A. J. Bleasby, and R. Fuchs. 1992. CLUSTAL V: improved software for multiple sequence alignment. *Comput. Appl. Biosci.* **8**:189–191.
29. Hulton, C. S., C. F. Higgins, and P. M. Sharp. 1991. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. *Mol. Microbiol.* **5**:825–834.
30. Jacoby, G. A., A. E. Jacob, and R. W. Hedges. 1976. Recombination between plasmids of incompatibility groups P-1 and P-2. *J. Bacteriol.* **127**:1278–1285.
31. Jones, J. D. G., and N. Gutterson. 1987. An efficient mobilizable cosmid vector, pRK7813, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* strain HV37a. *Gene* **61**:299–306.
32. Kuykendall, L. D., and G. H. Elkan. 1976. *Rhizobium japonicum* derivatives differing in nitrogen-fixing efficiency and carbohydrate utilization. *Appl. Environ. Microbiol.* **32**:511–519.
33. Laguerre, G., M.-R. Allard, F. Revoy, and N. Amarger. 1994. Rapid identification of rhizobia by restriction fragment length polymorphism analysis of PCR-amplified 16S rRNA genes. *Appl. Environ. Microbiol.* **60**:56–63.
34. Lennon, E., P. D. Gutman, H. Yao, and K. W. Minton. 1991. A highly conserved repeated chromosomal sequence in the radioresistant bacterium *Deinococcus radiodurans* SARK. *J. Bacteriol.* **173**:2137–2140.
35. Lupski, J. R., and G. M. Weinstock. 1992. Short, interspersed repetitive DNA sequences in prokaryotic genomes. *J. Bacteriol.* **174**:4525–4529.
36. Mantis, N. J., and S. C. Winans. 1993. The chromosomal response regulatory gene *chvI* of *Agrobacterium tumefaciens* complements an *Escherichia coli* *phoB* mutation and is required for virulence. *J. Bacteriol.* **175**:6626–6636.
37. Meade, H. M., S. R. Long, G. B. Ruvkun, S. E. Brown, and F. M. Ausubel. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J. Bacteriol.* **149**:114–122.
38. Miksch, G., W. Arnold, P. Lentzsch, U. B. Priefer, and A. Puhler. 1994. A 4.6kb DNA region of *Rhizobium meliloti* involved in determining urease and hydrogenase activities carries the structural genes for urease (*ureA*, *ureB*, *ureC*) interrupted by other open reading frames. *Mol. Gen. Genet.* **242**:539–550.
39. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
40. Morrison, N. A., C. Y. Hau, M. J. Trinick, J. Shine, and B. G. Rolfe. 1983. Heat curing of a Sym plasmid in a fast-growing *Rhizobium* sp. that is able to nodulate legumes and the nonlegume *Parasponia* sp. *J. Bacteriol.* **153**:527–531.
41. Newbury, S. F., N. H. Smith, and C. F. Higgins. 1987. Differential mRNA stability controls relative gene expression within a polycistronic operon. *Cell* **51**:1131–1143.
42. Østerås, M., B. T. Driscoll, and T. M. Finan. 1995. Molecular and expression analysis of the *Rhizobium meliloti* phosphoenolpyruvate carboxykinase (*pckA*) gene. *J. Bacteriol.* **177**:1452–1460.
43. Østerås, M., T. M. Finan, and J. Stanley. 1991. Site-directed mutagenesis and DNA sequence of *pckA* of *Rhizobium* NGR234, encoding phosphoenolpyruvate carboxykinase: gluconeogenesis and host-dependent symbiotic phenotype. *Mol. Gen. Genet.* **230**:257–269.
44. Parkinson, J. S., and E. C. Kofoid. 1992. Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* **26**:71–112.
45. Rastogi, V. K., and R. J. Watson. 1991. Aspartate aminotransferase activity is required for aspartate catabolism and symbiotic nitrogen fixation in *Rhizobium meliloti*. *J. Bacteriol.* **173**:2879–2887.
46. Ronson, C. W., B. T. Nixon, L. M. Albright, and F. M. Ausubel. 1987. *Rhizobium meliloti* *ntrA* (*rpoN*) gene is required for diverse metabolic functions. *J. Bacteriol.* **169**:2424–2431.
47. Rosenberg, C., and T. Huguet. 1984. The pAtC58 plasmid is not essential for tumor induction. *Mol. Gen. Genet.* **196**:533–536.
48. Ruvkun, G. B., and F. M. Ausubel. 1981. A general method for site-directed mutagenesis in prokaryotes. *Nature (London)* **289**:85–88.
49. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
50. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
51. Schmid, C. W., and W. R. Jelinek. 1982. The *Alu* family of dispersed repetitive sequences. *Science* **216**:1065–1070.
52. Sciaki, D., A. L. Montoya, and M.-D. Chilton. 1977. Fingerprints of *Agrobacterium Ti* plasmids. *Plasmid* **1**:238–253.
53. Sharples, G. J., and R. G. Lloyd. 1990. A novel repeated DNA sequence located in the intergenic regions of bacterial chromosomes. *Nucleic Acids Res.* **18**:6503–6508.
54. Shatters, R. G., J. E. Somerville, and M. L. Kahn. 1989. Regulation of glutamine synthetase II activity in *Rhizobium meliloti* 104A14. *J. Bacteriol.* **171**:5087–5094.
55. Stanley, J., and E. Cervantes. 1991. A review: biology and genetics of the broad host range *Rhizobium* sp. NGR234. *J. Appl. Bacteriol.* **70**:9–19.
56. Stanley, J., D. N. Dowling, and W. J. Broughton. 1988. Cloning of *hemA* from *Rhizobium* sp. NGR234 and symbiotic phenotype of a gene-directed mutant in diverse legume genera. *Mol. Gen. Genet.* **215**:32–37.
57. Steffan, R. J., and R. M. Atlas. 1991. Polymerase chain reaction: applications in environmental microbiology. *Annu. Rev. Microbiol.* **45**:137–161.
58. Stern, M. J., G. F.-L. Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins. 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. *Cell* **37**:1015–1026.
59. Stock, J. B., A. J. Ninfa, and A. M. Stock. 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* **53**:450–490.
60. Trinick, M. J. 1980. Relationships amongst the fast-growing rhizobia of *Lablab purpureus*, *Leucaena leucocephala*, *Mimosa* spp., *Acacia farnesiana* and *Sesbania grandiflora* and their affinities with other rhizobial groups. *J. Appl. Bacteriol.* **49**:39–53.
61. van Slooten, J. C., E. Cervantes, W. J. Broughton, C. H. Wong, and J. Stanley. 1990. Sequence and analysis of the *rpoN* sigma factor gene of *Rhizobium* sp. strain NGR234, a primary coregulator of symbiosis. *J. Bacteriol.* **172**:5563–5574.
62. Versalovic, J., T. Koeuth, and J. R. Lupski. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genome. *Nucleic Acids Res.* **19**:6823–6831.
63. Vieira, J., and J. Messing. 1987. Production of single-stranded plasmid DNA. *Methods Enzymol.* **153**:3–34.
64. Vincent, J. M. 1941. Serological properties of the root-nodule bacteria. I. Strains of *Rhizobium meliloti*. *Proc. Linn. Soc. N. S. W.* **66**:145–154.
65. Watson, R. J. Personal communication.
66. Wheatcroft, R., and R. J. Watson. 1988. Distribution of insertion sequence *IS*Rm*1* in *Rhizobium meliloti* and other gram-negative bacteria. *J. Gen. Microbiol.* **134**:113–121.
67. Yanagi, M., and K. Yamasato. 1993. Phylogenetic analysis of the family *Rhizobiaceae* and related bacteria by sequencing of 16S rRNA gene using PCR and DNA sequencer. *FEMS Microbiol. Lett.* **107**:115–120.
68. Yang, Y., and G. F.-L. Ames. 1988. DNA gyrase binds to the family of prokaryotic repetitive extragenic palindromic sequences. *Proc. Natl. Acad. Sci. USA* **85**:8850–8854.