APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Jan. 1993, p. 150–155
0099-2240/93/010150-06$02.00/0

Vol. 59, No. 1

Characterization of RFRS9, a Second Member of the *Rhizobium fredii* Repetitive Sequence Family from the Nitrogen-Fixing Symbiont *R. fredii* USDA257†

HARI B. KRISHNAN AND STEVEN G. PUEPPKE*

Department of Plant Pathology, 108 Waters Hall, University of Missouri, Columbia, Missouri 65211

Received 4 August 1992/Accepted 4 November 1992

The genome of the nitrogen-fixing symbiont, *Rhizobium fredii* USDA257, contains nine copies of repetitive sequences known as the *R. fredii* repetitive sequence (RFRS) family. We previously sequenced RFRS3, which is linked to symbiosis plasmid-borne nodulation genes of this organism and has substantial homology to the T-DNA of *Agrobacterium rhizogenes* and lesser homology to reiterated sequences of *Bradyrhizobium japonicum*. Here we characterize a second family member, RFRS9. The *Eco*RI fragment containing RFRS9 is 1,248 bp in length and contains a single 666-bp open reading frame that is flanked by perfect 8-bp inverted repeats. Nucleic and amino acid sequences corresponding to the C terminus of the putative RFRS9 protein are nearly identical to those of RFRS3, and they retain homology to DNA from *A. rhizogenes*. The central portion of the RFRS9 protein also appears to be related to the S locus-specific glycoprotein family of pollen stigma incompatibility glycoproteins from *Brassica oleracea*, which are involved in signal perception. Sequences that define the RFRS family are restricted to the open reading frame of RFRS9 and associated upstream sequences. These regions also contain a second group of repetitive sequences, which is present in four copies within the genome of USDA257. Both families of repetitive sequences are ubiquitous in *R. fredii*, and they are preferentially localized on symbiosis plasmids. Southern hybridization confirms that sequences homologous to RFRS9 are present in broad-host-range *Rhizobium* sp. strain NGR234, in *A. rhizogenes*, and in two biotype 3 strains of *Agrobacterium tumefaciens*.

The genomes of nitrogen-fixing rhizobia are characterized by the presence of various sorts of repetitive DNA sequences (3, 5). These elements include both gene and promoter duplications (2, 40), as well as insertion sequences and other reiterated regions. Although the biological significance of most of these DNA duplications is obscure (1), some multicopy genes in rhizobia are recognized as playing pivotal roles in root-nodule symbiosis. *nodD*, for example, which is present in two or three copies per *Rhizobium* genome, is known to encode a family of host recognition proteins. There is evidence that the interplay between these related sensors serves to fine tune the bacterial host range with legumes (7). *nodH*, which encodes nitrogenase reductase (24), and *nodPQ*, which encode sulfation enzymes (29, 34), also are reiterated.

Insertion sequences compose another subclass of repetitive sequences from rhizobia. These elements have two diagnostic features. First, they tend to be restricted in their occurrence to just one or a few species of rhizobia. IS*Rm1*, which was first discovered in the alfalfa and sweet clover symbiont, *Rhizobium meliloti* (31), is an example. This insertion sequence is present in 80% of the strains from a North American collection of *R. meliloti* (42) and 97% of the strains from a similar European collection (10). Only one strain of *Rhizobium leguminosarum* bv. phaseoli and two taxonomically intermediate isolates from alfalfa also harbor this sequence (42). IS*Rm2* (4) and IS*Rm3* (41) are similarly distributed in *R. meliloti*, but both also have been detected in biovars of *R. leguminosarum*. A second characteristic of *Rhizobium* insertion sequences is their strong association with and often preferential transposition into nodulation (*nod*) and nitrogen fixation (*nif*) genes. Indeed, both IS*Rm1* (31) and IS*Rm2* (4) were discovered as serendipitous mutagens of *nod* genes in *R. meliloti*, and IS*Rm3*, too, has been detected as a natural inactivator of the *nod* gene *nodG* (19). Other repetitive sequences have been localized in *nod* and *nif* regions of *R. meliloti* (27) and *Bradyrhizobium japonicum* (8, 9, 13, 26, 30) and on the symbiosis plasmid of *Rhizobium fredii* USDA193 (17). Although such sequences have been proposed to play roles in nodulation (30) and genome instability (9), convincing evidence for these functions in nature is still not available.

As part of an ongoing analysis of the *nod* genes of the broad-host-range, nitrogen-fixing symbiont *R. fredii* USDA257, we recently discovered that *nodABC* of this strain is adjacent to a DNA sequence that is repeated elsewhere in the genome (14, 15). This sequence, which lies about 500 bp upstream from *nodA*, contains many short inverted and direct repeats, as well as two overlapping open reading frames (ORFs) of opposite polarity. These ORFs have significant homology to a pair of ORFs from the left transferred DNA (T-DNA) of *Agrobacterium rhizogenes* (35), as well as weaker homology to the RSα family of repetitive sequences from *B. japonicum* (8, 9, 13, 26). The repetitive sequences are present in 9 different *Eco*RI fragments within the genome of USDA257 and in 6 to 11 such fragments within the genomes of other *R. fredii* strains, but they are not detectable by hybridization in other species of rhizobia. Cosmids containing all nine copies of the sequences have been isolated from a DNA library of USDA257, and with the exception of the two smallest restriction fragments, which are present on a single cosmid, the elements do not appear to be closely linked. We designate the repetitive sequences of USDA257 as RFRS1 to RFRS9, in the order of descending size of the *Eco*RI fragments that contain them.

* Corresponding author.

† Journal series No. 11793 of the Missouri Agricultural Experiment Station.

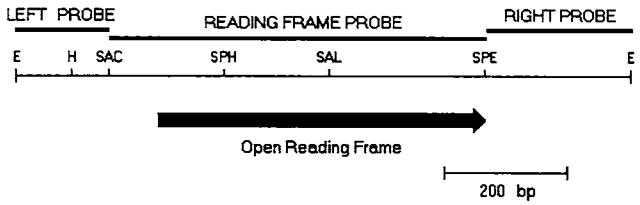

FIG. 1. Coordinated physical and genetic map of the 1.25-kb *Eco*RI fragment containing RFRS9 of *R. fredii* USDA257. Restriction sites used for mapping the fragment are indicated as E, *Eco*RI; H, *Hin*dIII; Sac, *Sac*I; Sal, *Sal*I; Spe, *Spe*I; and Sph, *Sph*I. The thick arrow below the map indicates the position of the large ORF and its direction of transcription. The thin bars above the map indicate the three hybridization probes that were prepared from the fragment.

Here we present data on the characterization of the 1.25-kb *Eco*RI fragment that contains RFRS9. We show that this region retains extensive homology to the T-DNA of *A. rhizogenes* and to RFRS3, the *nodABC*-linked repetitive sequence, but contains considerably more-limited homology to RSα9. We also provide evidence that RFRS9 has some features of an insertion sequence, that it is linked to a second family of repetitive sequences, and that related sequences are present in *Agrobacterium tumefaciens*.

MATERIALS AND METHODS

Bacterial strains. Strains of *R. fredii* were obtained from the culture collection of the U.S. Department of Agriculture, Beltsville, Md. *Rhizobium* sp. strain NGR234 was a gift of W. J. Broughton, University of Geneva. S. Shantharam of Iowa State University provided IA728, a Sym plasmid-cured derivative of *R. fredii* USDA193 (25). *A. tumefaciens* strains used and their sources were as follows: AT181 and A178 from V. K. Anand, University of Missouri-St. Louis; ATCC15955 and S1005 from W. B. Gurley, University of Florida; A281 from L. Owens, U.S. Department of Agriculture, Beltsville, Md.; A6, AT181, C58, TM4, Ag162, Ag63, and FACH from R. N. Goodman of our department; Chry5 from R. Stall, University of Florida. *A. rhizogenes* ATCC15834 was from J. A. Lippincott, Northwestern University. All of these bacteria were maintained in 7.5% glycerol at −70°C and cultured routinely in liquid yeast extract-mannitol broth (38) at room temperature. *Escherichia coli* DH5α was cultured in Luria broth at 37°C.

Molecular and genetic methods. pRSRF9, which contains a 1.25-kb *Eco*RI fragment with homology to the *nodABC*-linked repetitive sequence of USDA257, has been described previously (14). The 1.25-kb *Eco*RI fragment from this cosmid was subcloned into pGEM7Zf(+), from Promega Biotech, as pHBK314 (Fig. 1). The entire 1,248-bp fragment was sequenced by the dideoxy-chain termination method of Sanger et al. (33). Sequencing was initiated at the flanking *Eco*RI sites and extended with the aid of a series of oligonucleotide primers. The entire sequence was verified by complete analysis of both strands; computerized manipulation was done with the GenBank software package. General procedures for DNA isolation and cloning, restriction, and Southern hybridizations, etc., were done essentially as described by Sambrook et al. (32). Hybridizations were done in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) containing 0.5% sodium dodecyl sulfate (SDS), 5× Denhardt's solution (32), and 100 μg of denatured salmon sperm DNA per ml. Filters were washed at 68°C in 0.3× SSC containing 0.1% SDS. *Hin*dIII digests of bacteriophage lambda served as molecular size markers on gels.

```
GAATTCCGCATCGCGATGGACGAGACGACTGTCGGCCGTGAGTTGAAGGC   50
GCTCGGCTTCGCCAAGCTGTCGGCCCGCCCGCGCCACTACGCCCAGAACG  100
AGCTGGAGGCAGAAGCTTTTAAAAAGACTTCCCCGCCGCTCTGGCAAGAG  150
ATCCGAGGCCGGCTCCCGCGCGGCACCGACCTCGAGCTCTGGTGGGCCGA  200
CGAAGCGCGCATAGGCCAGAAGAACAAGATCACGCGGCGGTGGGCGTCGC  250

GGAACCAGACCCTCCGCACCCTTGGATCAGCGCACCATGTGGGCCTACAT  300
                                        M  W  A  Y  I
CTTCGGGGCCGTCTGCCCGCGGAAGGGAAAGGGTGCGGGCCTTGTCCTGC  350
  F  G  A  V  C  P  R  K  G  K  G  A  G  L  V  L
CCTATTGTGATACCGAAGCCATGCAGCAGCATCTCGCCGAGATCAGCCAG  400
P  Y  C  D  T  E  A  M  Q  Q  H  L  A  E  I  S  Q
GCCGTCGATGAGGGAGCGCATGCCGTGCTCATCCTCGATCAGGCCGGATC  450
  A  V  D  E  G  A  H  A  V  L  I  L  D  Q  A  G  S
GGACGTCACGCCGAAGCTCAAAGTGCCGGACTTCATCACCCTGATGTTCC  500
  D  V  T  P  K  L  K  V  P  D  F  I  T  L  M  F
TGCCGCCTCGTTCGCCGGAACTGAACCCCGTGGAAAACGTCTGGCAGTTC  550
L  P  P  R  S  P  E  L  N  P  V  E  N  V  W  Q  F
CTGCGCGACAACTGCCTCTCGAACCGCATCTTCAAAGACTACGACGACAT  600
  L  R  D  N  C  L  S  N  R  I  F  K  D  Y  D  D  I
CGTCGCCCATTCGTGCGCCGCGTGGAACAAGCTCGTCGACCAGCCGTGGA  650
  V  A  H  S  C  A  A  W  N  K  L  V  D  Q  P  W
AGATCATGTCCATCGGACTCCGCGAATGGGCGCATCGGTCTCGATCACCG  700
K  I  M  S  I  G  L  R  E  W  A  H  R  S  R  S  P
CTCGTTGGTATTACCTTCCACTTCACGCCGACGTCCTGCTCGTGGCTCAA  750
  L  V  G  I  T  F  H  F  T  P  T  S  C  S  W  L  N
CGCCGTCGAGGGCTTCTTCGCCAAACTGTCGAAGCGTCGCCTCAAGCGCG  800
  A  V  E  G  F  F  A  K  L  S  K  R  R  L  K  R
GCGTCTTTCATCTGGTCGTTGACCTCCAGGCCGCCATCAACCGCTTCCTT  850
G  V  F  H  L  V  V  D  L  Q  A  A  I  N  R  F  L
ACAGAGCACAACCAACAACCCAAGCCCTTCACCTGGACCGCCGATCCCGA  900
  T  E  H  N  Q  Q  P  K  P  F  T  W  T  A  D  P  D
CAAAATCATCGCTGCCGTCAAACGGGGGCACCAAGTGTTAGATTCCATCC  950
  K  I  I  A  A  V  K  R  G  H  Q  V  L  D  S  I
ACTAGTTGCCACGGTGCGCGCAATTCTCGACCGACTGGCGAAGGCCGCAT 1000
H  *
AACAGGGTATGGATGGTTCTGGCCAAAGAACCGCCACCGAGCTTTCGACA 1050

TGGCTCCGTCGCTAGACCTAACTCAAGAATTGGTCCGGGAGCGTCTACCA 1100
CTGTTGCCCATACCTGAAAGGGCATCCACCTTGCCAAGGCAGGACCGTCG 1150
GGAGCCGCGTACCATTGACATGATGTCAGCCATTGCGCCGGTACGCGCAG 1200
GGCAAACCGATCATCTTCTTCGACCATCGGATACAAACCAAGGAATTC
```

FIG. 2. Sequence of the RFRS9 region. The sequence covers 1,248 nucleotides and is flanked by *Eco*RI sites. The translation of the 666-bp ORF is given. The pair of 8-bp, insertion sequence-like inverted repeats is underlined. The amino acid region with strong homology to ORF1 of RFRS3 is accented in boldface type.

RESULTS AND DISCUSSION

Nucleotide sequence of RFRS9. Cosmid pRFRS9 contains a 1.25-kb *Eco*RI fragment that hybridizes to the *nodABC*-linked RFRS3 repetitive sequence of *R. fredii* USDA257 (14). As a first step in understanding the relationship between these two elements, we have subcloned RFRS9 (Fig. 1) and sequenced it in its entirety (Fig. 2). The 1,248-bp *Eco*RI fragment contains a single large ORF that extends from nucleic acid position 287 to position 953 and encodes a putative 222-amino-acid protein with a size of 25,127 Da. The protein lacks a membrane transit sequence and thus is predicted to be cytosolic.

Although overall homology between RFRS3 and RFRS9 is only 61%, a 245-bp sequence between positions 710 and 955 of RFRS9 (Fig. 2) is virtually identical to that of RFRS3, differing by only four bases. This conserved region, which is sufficient to account for the observed hybridization between the elements, corresponds to the C terminus of the RFRS9 ORF. Homology ends precisely after the TAG termination codon. Although the N-terminal and central portions of the two proteins bear little relationship to one another, an 80-amino-acid region of close identity begins at Thr-143 and encompasses the C-terminal third of the protein (boldface letters, Fig. 2). This region corresponds to that of strong nucleic acid homology between the two sequences.

Like RFRS3, RFRS9 has extensive homology to nucleic acid sequences from two other members of the family *Rhizobiaceae*: a segment of the T-DNA of *A. rhizogenes* A4 (35) and the RSα family of repetitive sequences from *B. japonicum* USDA110 (13). Significant homology is restricted to the leftmost two-thirds of the RFRS9 restriction fragment, including the ORF and its associated upstream nontranslated sequences. Thus, sequences between positions 625 and 954 of RFRS9 have 71% identity to T-DNA, and those between positions 1 and 994 of RFRS9 have 53% identity to RSα9. Amino acid homology between the RFRS9 ORF and the corresponding ORF from *B. japonicum* RSα9, however, is inconsequential. Although the same is true for the comparison between the N termini of the RFRS9 ORF and ORF6 of the *A. rhizogenes* T-DNA region, 56 of the 80 C-terminal amino acids of these proteins are identical.

Only one other group of proteins has significant homology to the deduced protein encoded by RFRS9. These are the S locus-specific glycoproteins (SLSGs) of *Brassica oleracea*, which control pollen stigma incompatibility during flowering (18). A central 101-amino-acid domain of the 222-amino-acid RFRS9 protein has 21% identity with 106 amino acid residues from the interior of the 414-amino-acid SLSG-s6 protein. The structural conservation of these aligned sequences (39) is 28%, yielding an overall similarity of 49%. A second eukaryotic protein, ZmPK1 from maize, also has significant homology to the SLSGs (39) and to RFRS9. Although we do not know whether these relationships are important, Walker and Zhang (39) have proposed that the SLSGs and ZmPK1 both mediate cellular responses to external signals.

RFRS9 has properties of an insertion element. Ramsmeier and Göttfert (26) recently have shown that two of the sequences with homology to RFRS9, the RSα elements of *B. japonicum* and a portion of the T-DNA of *A. rhizogenes*, have characteristics of insertion sequences. These include 9- to 15-bp inverted repeats that flank the ORFs and the presence of potential target site duplications. On the basis of these data, they have argued that these two elements may represent related, novel prokaryotic insertion elements (26). RFRS3 lacks these structural characteristics (14), but they are evident in RFRS9, in which the ORF is flanked by two perfect 8-bp inverted repeats (underlined in Fig. 2), the second of which lies 12 bp beyond the termination codon. Thus, RFRS9, but not RFRS3, has structural characteristics of an *R. fredii* insertion element.

RFRS9 is linked to a second family of repetitive sequences. We separated the 1.25-kb *Eco*RI fragment into three contiguous hybridization probes, which were used to localize regions of homology between RFRS9 and other parts of the USDA257 genome. The left probe extends from the first *Eco*RI site to a *Sac*I site located 100 bp upstream from the ORF (Fig. 1). The ORF probe spans the region from this *Sac*I site to a *Spe*I site within the final codon of the ORF, and the right probe extends from this site to the end of the restriction fragment (Fig. 1). These probes were hybridized separately to *Eco*RI digests of genomic DNA from USDA257.

Figure 3 shows that the right probe hybridizes only to the 1.25-kb fragment and thus lacks repetitive sequences entirely. The reading frame probe hybridizes to all members of the RFRS family, except for the 1.6-kb RFRS8 fragment. Unexpectedly, this probe also hybridizes to three other *Eco*RI fragments with lengths of 7.0, 2.8, and 2.5 kb. These fragments lack homology to RFRS3 (14), indicating that they represent a second family of repetitive sequences, members of which are reiterated four times in the USDA257 genome.

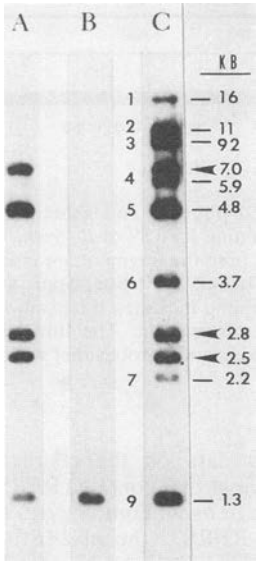

FIG. 3. Localization of repetitive sequences within the RFRS9 fragment of *R. fredii* USDA257. *Eco*RI-digested genomic DNA was hybridized to an *Eco*RI-*Sac*I fragment that lies left of the ORF (lane A), an *Sph*I-*Eco*RI fragment that lies right of the ORF (lane B), and *Sac*I-*Sph*I fragment that includes the entire ORF (lane C). The eight RFRS elements are identified by numbers to the left of lane C. Fragments belonging to the second family of repetitive sequences are indicated by arrowheads on the right. The sizes (in kilobases) of the elements are given at the extreme right.

The 7.0-, 2.8-, and 2.5-kb fragments are also apparent when the left probe from RFRS9 is used, but most members of the RFRS family are absent. Thus, the majority of the RFRS sequences are associated with the RFRS9 reading frame, most likely the highly conserved C-terminal region, but the second family of repetitive sequences must span the *Sac*I site that demarcates the ORF from upstream sequences.

RFRS elements are characteristic of *R. fredii* symbiosis plasmids and of some agrobacteria. Using a 0.95-kb *Hin*dIII-*Sal*I fragment of RFRS3 as the hybridization probe, we previously showed that RFRS sequences have homology to 6 to 11 *Eco*RI fragments in each of 10 strains of *R. fredii* (14). We have probed DNA from each of these strains with the 1.25-kb RFRS9 fragment and extended the study by including a Sym plasmid-cured strain, a mutant with a Sym plasmid deletion, and *Rhizobium* sp. strain NGR234, a broad-host-range strain believed to be related to *R. fredii* (12). These hybridization data (Fig. 4A) allow three major conclusions to be drawn. First, the numbers of hybridizing bands and polymorphisms are even more complex than those with RFRS3. We attribute this to the second family of repetitive sequences, which is present in all wild-type *R. fredii* strains. In fact, the 2.5-kb fragment is present in all strains, and the 2.8-kb fragment is evident in all strains but USDA217. Conversely, the only RFRS3-homologous band that is present in all strains, RFRS8, is uniformly absent in all strains probed with RFRS9. This complexity approaches that recently observed for another soybean symbiont, *B. japonicum*, in which conserved hyperreiterated sequences are duplicated from 18 to 21 times in strains of serocluster 123 (30).

Comparison of USDA193 to its Sym plasmid-cured mutant IA728 allows a second conclusion to be drawn: both the

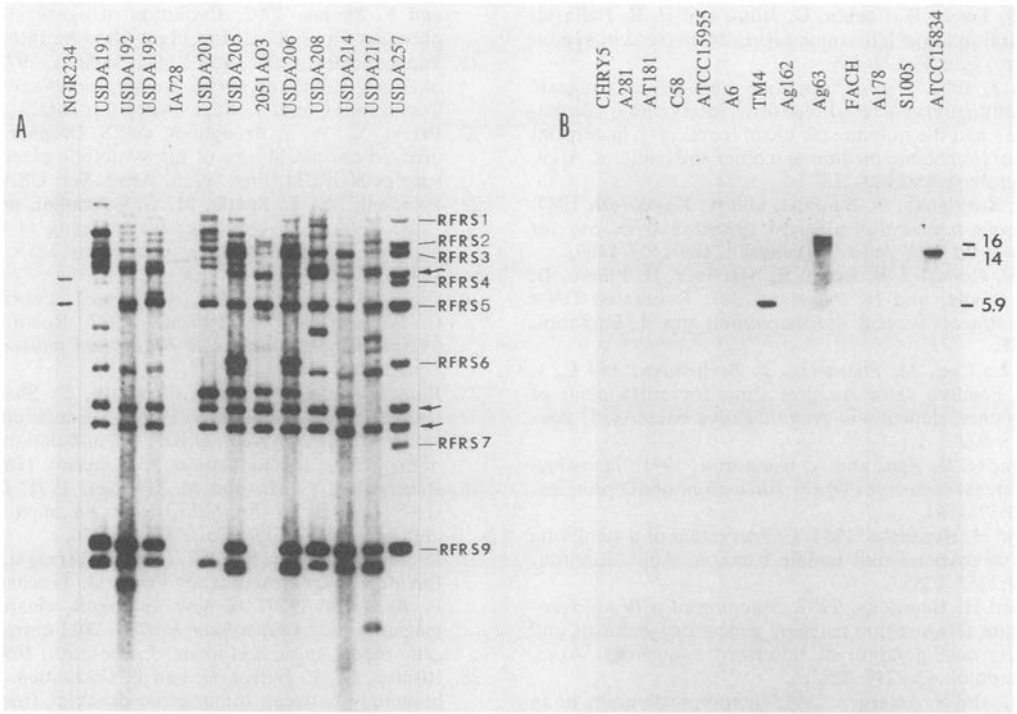

FIG. 4. Repetitive sequences with homology to RFRS9 are widespread in *R. fredii* and detectable in *A. rhizogenes* and *A. tumefaciens*. (A) The 1.25-kb RFRS9 fragment was hybridized to *Eco*RI digests of genomic DNA from *R. fredii* strains and from *Rhizobium* sp. strain NGR234. The RFRS fragments are identified on the right, and arrows mark the positions of fragments belonging to the second family of repetitive sequences; (B) *A. rhizogenes* and *A. tumefaciens*. Sizes of the hybridizing fragments from agrobacteria are indicated on the right.

RFRS elements and the second family of sequences have a strong Sym plasmid bias. Thus, 13 of the 14 hybridizing fragments of USDA193 are absent in IA728. Comparison of USDA205 with 2051A03, which lacks a portion of the Sym plasmid (13a), is consistent with this interpretation and suggests that the Sym plasmid association may be generalized. Assuming a mean insert of 25 kb per cosmid and a Sym plasmid size of approximately 450 kb (11), the RFRS elements are likely to be interspersed at intervals of 20 to 30 kb or less on the Sym plasmids of *R. fredii*. The density of repetitive sequences becomes even greater when the second family of elements is considered. Although we have no ready functional explanation for the preferential localization of these elements on the plasmid, their highly selective association with plasmid-borne genes that control symbiotic phenotypes is consistent with some role in this process.

Figure 4A also provides evidence that the RFRS sequences are present in the genome of *Rhizobium* sp. strain NGR234, a strain that is closely allied with and perhaps a member of the species *R. fredii* (12). The two hybridizing bands from this organism are indistinguishable from those of RFRS4 and RFRS5, and we are currently attempting to align them on the map of contiguous sequences of the NGR234 Sym plasmid (22).

Sequences homologous to RFRS9 are present in the genomes of 3 of 13 diverse *Agrobacterium* strains (Fig. 4B). One of these, *A. rhizogenes* ATCC15834, is essentially identical to strain A4 (20, 43), the source of the T-DNA sequences with strong nucleic acid homology to RFRS9 (35). The conserved sequences lie on a 13.4-kb *Eco*RI fragment from strain A4 (23), and within experimental error, the size of this fragment corresponds to that of the 14-kb hybridizing fragment from ATCC15834.

The observation of high-stringency hybridization between RFRS9 and single *Eco*RI fragments from *A. tumefaciens* TM4 and Ag63 (Fig. 4B) is more difficult to rationalize. Although both these strains are pathologically specialized biotype 3 isolates from grapevine (21, 36), other such strains, including FACH and Ag162, lack homology, and thus, the relationship is not biotype specific. Moreover, the T-DNA of *A. rhizogenes* and that of *A. tumefaciens* are not well conserved, particularly in the region of sequence homology with RFRS9 (16, 28, 43). Consequently, the observed hybridization must be based on some other similarity between *R. fredii* and certain strains of agrobacteria. One possible candidate is IS*426* (also called IS*136*), an insertion element that has been detected in several strains of *A. tumefaciens* (6, 20, 37) and that has 53% sequence identity with 205 nucleotides between positions 721 and 926 of RFRS9, in the region encoding the C terminus of the putative RFRS9 protein.

ACKNOWLEDGMENT

This work was supported by the Food for the 21st Century Program, University of Missouri.

REFERENCES

1. **Barran, L. R., E. S. P. Bromfield, V. Rastogi, S. T. Whitwill, and R. Wheatcroft.** 1990. Transposition and copy number of insertion sequence IS*Rm*1 are not correlated with symbiotic performance of *Rhizobium meliloti* from two field sites. Can. J. Microbiol. **37**:576–579.

2. **Better, M., B. Lewis, D. Corbin, G. Ditta, and D. R. Helinski.** 1983. Structural relationships among *Rhizobium meliloti* symbiotic promoters. Cell **35:**479–485.
3. **DeBruijn, 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* and other soil bacteria. Appl. Environ. Microbiol. **58:**2180–2187.
4. **Dusha, I., S. Kovalenko, Z. Banfalvi, and A. Kondorosi.** 1987. *Rhizobium meliloti* insertion element IS*Rm2* and its use for identification of the *fixX* gene. J. Bacteriol. **169:**1403–1409.
5. **Flores, M., V. González, S. Brom, E. Martinez, D. Piñero, D. Romero, G. Dávila, and R. Palacios.** 1987. Reiterated DNA sequences in *Rhizobium* and *Agrobacterium* spp. J. Bacteriol. **169:**5782–5788.
6. **Gay, P., D. Le Coq, M. Steinmetz, T. Berkelman, and C. I. Kado.** 1985. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. J. Bacteriol. **164:**918–921.
7. **Györgypal, Z., G. B. Kiss, and A. Kondorosi.** 1991. Transduction of plant signal molecules by the *Rhizobium* nodD proteins. BioEssays **13:**575–581.
8. **Hahn, M., and H. Hennecke.** 1987. Conservation of a symbiotic DNA region in soybean root nodule bacteria. Appl. Environ. Microbiol. **53:**2253–2255.
9. **Hahn, M., and H. Hennecke.** 1987. Mapping of a *Bradyrhizobium japonicum* DNA region carrying genes for symbiosis and an asymmetric accumulation of reiterated sequences. Appl. Environ. Microbiol. **53:**2247–2252.
10. **Hartmann, A., and N. Amarger.** 1991. Genotypic diversity of an indigenous *Rhizobium meliloti* field population assessed by plasmid profiles, DNA fingerprinting, and insertion sequence typing. Can. J. Microbiol. **37:**600–608.
11. **Heron, D. A., and S. G. Pueppke.** 1984. Mode of infection, nodulation specificity, and indigenous plasmids of 11 fast-growing *Rhizobium japonicum* strains. J. Bacteriol. **160:**1061–1066.
12. **Jarvis, B. D. W., H. L. Downer, and J. P. W. Young.** 1992. Phylogeny of fast-growing soybean-nodulating rhizobia supports synonymy of *Sinorhizobium* and *Rhizobium* and assignment to *Rhizobium fredii*. Int. J. Syst. Bacteriol. **42:**93–96.
13. **Kaluza, K., M. Hahn, and H. Hennecke.** 1985. Repeated sequences similar to insertion elements clustered around the *nif* region of the *Rhizobium japonicum* genome. J. Bacteriol. **162:**535–542.

13a. **Krishnan, H. B.** Unpublished observation.

14. **Krishnan, H. B., and S. G. Pueppke.** 1991. Repetitive sequences with homology to *Bradyrhizobium japonicum* DNA and the T-DNA of *Agrobacterium rhizogenes* are closely linked to *nodABC* of *Rhizobium fredii* USDA257. Mol. Plant-Microbe Interact. **4:**521–529.
15. **Krishnan, H. B., and S. G. Pueppke.** 1991. Sequence and analysis of the *nodABC* region of *Rhizobium fredii* USDA257, a nitrogen-fixing symbiont of soybean and other legumes. Mol. Plant-Microbe Interact. **4:**512–520.
16. **Levesque, H., P. Delepelaire, P. Rouzé, J. Slightom, and D. Tepfer.** 1988. Common evolutionary origin of the central portions of the Ri TL-DNA of *Agrobacterium rhizogenes* and the Ti T-DNAs of *Agrobacterium tumefaciens*. Plant Mol. Biol. **11:**731–744.
17. **Masterson, R. V., and A. G. Atherly.** 1986. The presence of repeated DNA sequences and a partial restriction map of the pSym of *Rhizobium fredii* USDA 193. Plasmid **16:**37–44.
18. **Nasrallah, J. B., T.-H. Kao, C.-H. Chen, M. L. Goldberg, and M. E. Nasrallah.** 1987. Amino-acid sequence of glycoproteins encoded by three alleles of the S locus of *Brassica oleracea*. Nature (London) **326:**617–619.
19. **Ogawa, J., H. L. Brierley, and S. R. Long.** 1991. Analysis of *Rhizobium meliloti* nodulation mutant WL131: novel insertion sequence IS*Rm3* in *nodG* and altered *nodH* protein product. J. Bacteriol. **173:**3060–3065.
20. **Otten, L., J. Canaday, J.-C. Gérard, P. Fournier, P. Crouzet, and F. Paulus.** 1992. Evolution of agrobacteria and their Ti plasmids—a review. Mol. Plant-Microbe Interact. **5:**279–287.
21. **Panagopoulos, C. G., and P. G. Psallidas.** 1973. Characteristics of Greek isolates of *Agrobacterium tumefaciens* (E. F. Smith & Townsend) Conn. J. Appl. Bacteriol. **36:**233–240.
22. **Perret, X., W. J. Broughton, and S. Brenner.** 1991. Canonical ordered cosmid library of the symbiotic plasmid of *Rhizobium* species NGR234. Proc. Natl. Acad. Sci. USA **88:**1923–1927.
23. **Pomponi, M., L. Spanò, M. G. Sabradini, and P. Costantino.** 1983. Restriction endonuclease mapping of the root-inducing plasmid of *Agrobacterium rhizogenes* 1855. Plasmid **10:**119–129.
24. **Quinto, C., H. de la Vega, M. Flores, L. Fernández, T. Ballado, G. Soberón, and R. Palacios.** 1982. Reiteration of nitrogen fixation gene sequences in *Rhizobium phaseoli*. Nature (London) **299:**724–726.
25. **Ramakrishnan, N., R. K. Prakash, S. Shantharam, N. M. DuTeau, and A. G. Atherly.** 1986. Molecular cloning and expression of *Rhizobium fredii* USDA 193 nodulation genes: extension of host range for nodulation. J. Bacteriol. **168:**1087–1095.
26. **Ramsmeier, T. M., and M. Göttfert.** 1991. Codon usage and G+C content in *Bradyrhizobium japonicum* genes are not uniform. Arch. Microbiol. **156:**270–276.
27. **Renalier, M.-H., J. Batut, J. Ghai, B. Terzaghi, M. Gherardi, M. David, A.-M. Garnerone, J. Vasse, G. Truchet, T. Huguet, and P. Boistard.** 1987. A new symbiotic cluster on the pSym megaplasmid of *Rhizobium meliloti* 2011 carries a functional *fix* gene repeat and a *nod* locus. J. Bacteriol. **169:**2231–2238.
28. **Risuleo, G., P. Battistoni, and P. Costantino.** 1982. Regions of homology between tumorigenic plasmids from *Agrobacterium rhizogenes* and *Agrobacterium tumefaciens*. Plasmid **7:**45–51.
29. **Roche, P., F. Debellé, F. Maillet, P. Lerouge, C. Faucher, G. Truchet, J. Dénarié, and J.-C. Promé.** Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. Cell **67:**1131–1143.
30. **Rodriguez-Quiñones, F., A. K. Judd, M. J. Sadowsky, R.-L. Liu, and P. B. Cregan.** 1992. Hyperreiterated DNA regions are conserved among *Bradyrhizobium japonicum* serocluster 123 strains. Appl. Environ. Microbiol. **58:**1878–1885.
31. **Ruvkun, G. B., S. R. Long, H. M. Meade, R. C. van den Bos, and F. M. Ausubel.** 1982. ISRm1: a *Rhizobium meliloti* insertion sequence that transposes preferentially into nitrogen fixation genes. J. Mol. Appl. Genet. **1:**405–418.
32. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Sanger, F., A. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
34. **Schwedock, J., and S. R. Long.** 1989. Nucleotide sequence and protein products of two new nodulation genes of *Rhizobium meliloti*, *nodP* and *nodQ*. Mol. Plant-Microbe Interact. **2:**181–194.
35. **Slightom, J. L., M. Durand-Tardif, L. Jouanin, and D. Tepfer.** 1986. Nucleotide sequence analysis of TL-DNA of *Agrobacterium rhizogenes* agropine type plasmid. J. Biol. Chem. **261:**108–121.
36. **Szegedi, E., M. Czakó, L. Otten, and C. S. Koncz.** 1988. Opines in crown gall tumors induced by biotype 3 isolates of *Agrobacterium tumefaciens*. Physiol. Mol. Plant Pathol. **32:**237–247.
37. **Vanderleyden, J., J. Desair, C. DeMeirsman, K. Michiels, A. P. VanGool, M.-D. Chilton, and G. C. Jen.** 1986. Nucleotide sequence of an insertion sequence (IS) element identified in the T-DNA region of a spontaneous variant of the Ti-plasmid pTiT37. Nucleic Acids Res. **14:**6699–6709.
38. **Vincent, J. M.** 1970. A manual for the practical study of root-nodule bacteria. Blackwell Scientific Publications, Oxford.
39. **Walker, J. C., and R. Zhang.** 1990. Relationship of a putative receptor protein kinase from maize to the S-locus glycoproteins of *Brassica*. Nature (London) **345:**743–746.
40. **Watson, J. M., and P. R. Schofield.** 1985. Species-specific, symbiotic plasmid-located repeated DNA sequences in *Rhizo-*

bium trifolii. Mol. Gen. Genet. **199:**279–289.
41. **Wheatcroft, R., and S. Laberge.** 1991. Identification and nucleotide sequence of *Rhizobium meliloti* insertion sequence IS*Rm3*: similarity between the putative transposase encoded by IS*Rm3* and those encoded by *Staphylococcus aureus* IS*256* and *Thiobacillus ferrooxidans* IS*T2*. J. Bacteriol. **173:**2530–2538.
42. **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.
43. **White, F. F., and E. W. Nester.** 1980. Relationship of plasmids responsible for hairy root and crown gall tumorigenicity. J. Bacteriol. **144:**710–720.