MOLECULAR PLANT-MICROBE INTERACTIONS
Vol. 2, No. 5, pp. 224-232, 1989

Isolation and Characterization of a *Rhizobium loti* Gene Required for Effective Nodulation of *Lotus pedunculatus*

Lawrence J. H. Ward,[1] Eva S. Rockman,[2] Philippa Ball,[2] Brion D. W. Jarvis,[1] and D. Barry Scott[2]

[1]Department of Microbiology and Genetics, [2]Molecular Genetics Unit, Massey University, Palmerston North, New Zealand.
Received 10 January 1989. Accepted 23 March 1989.

A *Rhizobium loti* gene required for effective invasion of the host *Lotus pedunculatus* has been identified by transposon Tn5 mutagenesis. Cosmids that complemented a previously isolated mutation (239) at this invasion (*inv*) locus were identified by *in planta* complementation and used to construct a physical map of the gene region. The insertion site of Tn5 in PN239 was mapped to a 7.5-kb *Eco*RI fragment, which complemented the mutation when subcloned into pLAFR1. Further Tn5 mutagenesis of the 7.5-kb fragment was carried out in *Escherichia coli* using bacteriophage λ467, and the mutations homogenotized into *R. loti* NZP2037. Three additional Fix⁻ mutations were isolated, and these were found to map adjacent to the position of the original mutation in strain PN239. All the other Tn5 insertions isolated in the 7.5-kb fragment gave a Fix⁺ phenotype on *L. pedunculatus*. Electron microscopic examination of the *L. pedunculatus* nodules induced by the isolated Fix⁻ mutants showed that bacteria were either blocked in release from the infection threads or were unable to undergo normal bacteroid development. The *inv* locus as defined by the Tn5 insertions was sequenced, and a single open-reading frame (ORF) of 576 bp, corresponding to a polypeptide of 21.3 kDa, was identified. The position and orientation of this ORF were consistent with those of the isolated Tn5 Fix⁻ insertions.

Additional keywords: DNA sequence, invasion (*inv*) gene, nitrogen fixation.

Establishing a nitrogen-fixing symbiosis between *Rhizobium* and leguminous plants is a complex developmental process involving the expression of symbiotic genes in both the host and the symbiont. Nodule ontogeny has been studied in a number of different *Rhizobium*-legume associations, and a number of common features are observed (Dart 1977; Goodchild 1978; Newcomb 1981; Robertson and Farnden 1980).

The common mechanism of bacterial invasion for temperate legumes is infection through newly emerged root hairs (Bauer 1981; Calvert *et al.* 1984). Attachment and colonization of the newly formed root hair result in root hair curling and enclosure of the bacterial colony within the "pocket" of the curled hair. Infection then proceeds by a plant-derived infection thread that develops within the root hair at the point of contact of the enclosed rhizobial colony. Degradation of the plant cell wall is evident at the origin of the infection thread (Callaham and Torrey 1981), and it is thought that the new cell wall layers of the thread are formed by apposition of cellulosic material by the host cytoplasm (Bauer 1981). The presence of rhizobia on the surface of the plant root hair cells induces a series of anticlinal divisions in the cells of the outer cortex (Newcomb 1981; Calvert *et al.* 1984). This dedifferentiation process spreads to the inner cortex resulting in the formation of nodule primordia as the infection thread penetrates the cells of the outer and inner cortex.

The bacteria are released from the infection thread(s) into meristematic cells by an endocytotic mechanism (Dart 1977; Goodchild 1978), resulting in enclosure of the bacterium within a plant membrane, the peribacteroid membrane (PBM) (Robertson *et al.* 1978). This release mechanism gives rise to cortical cells packed with rhizobia that always remain extracellular to the plant cell cytoplasm. The PBM is thought to be derived from membrane vesicles of the Golgi apparatus (Robertson *et al.* 1978), and recent work has shown that nodule-specific plant proteins (nodulins) are targeted to the PBM (Fortin *et al.* 1985).

Once enveloped within the PBM, the rhizobia differentiate into pleiomorphic forms called bacteroids that are capable of nitrogen fixation within the nodule. Various sets of *Rhizobium* genes are required at the different stages leading to the establishment of a nitrogen-fixing nodule. A set of plant-inducible nodulation (*nod*) genes is required for the early steps of host recognition and initiation of a nodule meristem (reviewed by Kondorosi and Kondorosi 1986; Rossen *et al.* 1987). Subsequent invasion of the host is dependent on the presence of specific polysaccharides on the surface of the *Rhizobium*. A range of Fix⁻ mutants has been isolated. They form empty nodules that have either no infection threads (Chua *et al.* 1985; Finan *et al.* 1985; Leigh *et al.* 1985; Vandenbosch *et al.* 1985; Dylan *et al.* 1986) or contain aborted infection threads (Noel *et al.* 1986), demonstrating that the process of nodulation can be uncoupled from bacterial invasion of the plant host. These Fix⁻ mutants have been shown to have alterations in the composition of the exopolysaccharide (Finan *et al.* 1985; Leigh *et al.* 1985), or lipopolysaccharide (Noel *et al.* 1986), or β-glucan synthesis (Dylan *et al.* 1986). The role of these different polysaccharide fractions in the infection process is still to be explained.

Genes required for bacterial release from the infection thread and the subsequent steps leading to differentiation of a bacteroid remain poorly described. There have been several reports of mutants that are blocked in bacterial

Address all correspondence to D. B. Scott.

Nucleotide and/or amino acid sequence data is to be submitted to GenBank as accession number JO3677.

release (Pankhurst *et al.* 1972; Pankhurst 1974; Truchet *et al.* 1980; Chua *et al.* 1985; Morrison and Verma 1987), and recently one such gene was cloned and sequenced from *Bradyrhizobium japonicum* (Buchanan) Jordan (Morrison and Verma 1987). We previously reported the isolation of a mutant of *R. loti* Jarvis *et al.* strain NZP2037 that appeared to be blocked in release from the infection thread in *Lotus pedunculatus* nodules (Chua *et al.* 1985). We describe here the cloning and further characterization of this gene.

MATERIALS AND METHODS

Bacterial strains and plasmids. The strains and plasmids used in this study are described in Table 1.

Growth of bacteria. *R. loti* cultures were grown at 28° C in either tryptone-yeast extract (TY) medium (Beringer 1974) or S10-defined medium (Scott and Ronson 1982) and supplemented where necessary with neomycin (Nm) (400 μg/ml), tetracycline (Tc) (2 μg/ml), gentamycin (Gm) (50 μg/ml), or streptomycin (Str) (200 μg/ml). *Escherichia coli* cultures were grown in either TY medium or Luria-Bertani (LB) medium (Miller 1972). Antibiotic concentrations used for *E. coli* were Tc, 15 μg/ml; kanamycin (Km), 25 μg/ml; nalidixic acid (Nal), 25 μg/ml; and Gm, 25 μg/ml .

Table 1. Bacterial strains, plasmids, and phages

Strain, plasmid or phage	Relevant characteristics[a]	Source or reference[b]
Strains		
Rhizobium loti		
NZP2037	Nod^+Fix^+ (*Lotus pedunculatus, L. corniculatus*)	DSIR Culture Collection
PN184	NZP2037 *str* 1	Chua *et al.* 1985
PN239	PN184 *inv* 239::Tn5	Chua *et al.* 1985
PN1093	PN184 *inv* 41::Tn5	This study
PN1094	PN184 *inv* 78::Tn5	This study
PN1095	PN184 *inv* 18::Tn5	This study
Escherichia coli		
JM101	*sup*E, *thi*, Δ(*lac-pro*AB), [F′, *tra*D36, *pro*AB, *lac*I^q ZΔM15]	Yanisch-Perron *et al.* 1985
HB101	*pro leu thi gal lac*Y *rec*A *str hsd*D *hsd*M	Boyer and Roulland-Dussoix 1969
C2110	*pol*A Nal^r	Leong *et al.* 1982
PN372	HB101/pPN318	This study
PN374	HB101/pPN320	This study
PN1000	HB101/pPN27	This study
PN1001	HB101/pPN28	This study
Plasmids		
pRK2073	Km^s derivative of pRK2013	Leong *et al.* 1982
pPH1JI	Gm^r IncP	Beringer *et al.* 1978
pBR328	$Ap^rTc^rCm^r$	Bolivar *et al.* 1977
pLAFR1	λcos derivative of pRK290	Friedman *et al.* 1982
pPN318	pLAFR1 cosmid containing *inv* gene from NZP2037	This study
pPN320	pLAFR1 cosmid containing *inv* gene from NZP2037	This study
pPN27	pBR328 containing 7.5-kb *Eco*RI *inv* fragment from pPN318	This study
pPN28	pLAFR1 containing 7.5-kb *Eco*RI *inv* fragment from pPN318	This study
Bacteriophages		
λ467		de Bruijn and Lupski 1984
M13mp8		Messing and Vieira 1982
M13mp9		Messing and Vieira 1982

[a]Abbreviations: Nal, nalidixic acid; Km, kanamycin; Gm, gentamycin; Ap, ampicillin; Tc, tetracycline; Cm, chloramphenicol; ^{r}, resistant and ^{s}, sensitive; and *inv*, invasion (gene).

[b]DSIR, Department of Scientific and Industrial Research, Palmerston North, New Zealand.

Crosses. Crosses were carried out on TY medium by the patch plate method (Dixon *et al.* 1976). Transfer of pLAFR1 cosmids into rhizobia or *E. coli* was carried out by the triparental mating system (Ditta *et al.* 1980), using pRK2073 as the helper plasmid.

Site-directed Tn5 mutagenesis of fix locus. This was carried out by the method of de Bruijn and Lupski (1984) using phage λ467 as the source of Tn5 and *E. coli* strain C2110 as the recipient to select for Tn5 insertions in pPN28 following conjugal transfer, using pRK2073 as the helper plasmid. The positions of the Tn5 insertions in the 7.5-kb (kilobase) fragment were determined by restriction enzyme digestion and hybridization, using plasmid DNA isolated from the C2110 transconjugants. Following mapping, plasmids containing Tn5 insertions in the 7.5-kb *Eco*RI fragment were transferred to *R. loti* PN184 in a triparental cross, and the Tn5 homogenotized into wild-type sequences by plasmid incompatibility (Ruvkun and Ausubel 1981) using pPH1JI.

Nodulation tests. Nodulation tests were carried out using *L. pedunculatus* Cav. 'Grasslands Maku' and *L. corniculatus* L. by methods previously described (Scott and Ronson 1982). After 6 wk, plants were examined for the presence (Nod^+) or absence (Nod^-) of nodules and for symptoms of nitrogen starvation by measuring the dry weights of the tops of both inoculated and uninoculated plants. No significant difference in dry weight indicated an ineffective (Fix^-) symbiosis. Bacteria were isolated from nodules as previously described (Scott and Ronson 1982).

Light and electron microscopy. Light and electron microscopy of nodule sections was carried out as previously described (Pankhurst *et al.* 1979), using nodules harvested from plants 3–4 wk after inoculation.

Preparation of DNA. Total DNA was isolated by a modification (Scott *et al.* 1984) of the method of Fischer and Lerman (1979). Amplifiable plasmids were isolated by the cleared lysate method (Clewell and Helinski 1969), and pLAFR1 cosmids were isolated from *E. coli* by the alkaline lysis method of Ish-Horowicz and Burke (1981). For rapid screening of plasmids, small preparations were obtained by the method of Holmes and Quigley (1981).

Restriction enzyme digestion of DNA and agarose gel electrophoresis. Preparation of DNA digests and electrophoresis of the samples were carried out as previously described (Chua *et al.* 1985).

DNA hybridization procedures. Restriction enzyme DNA digests separated on 0.7% (w/v) agarose gels (Chua *et al.* 1985) were transferred to nitrocellulose by the method of Southern (1975). DNA was labeled with [^{32}P]-dCTP (dCTP, deoxycytidine triphosphate) (3,000 Ci/mmol, New England Nuclear, Boston, MA) by primed synthesis, using DNA polymerase I (Klenow fragment) and denatured herring sperm DNA primers as previously described (Whitfeld *et al.* 1982). Hybridizations were carried out at

65° C for 16 hr as previously described (Whitfeld *et al.* 1982).

Cloning procedures. Plasmids pPN27 and pPN28 were constructed by subcloning the 7.5-kb *Eco*RI fragment from pPN318 into pBR328 and pLAFR1, respectively, using methods previously described (Scott *et al.* 1985).

DNA sequencing. The *inv* gene region identified between the boundaries of the Tn5 insertions 56 and 63 (Fig. 1C) was sequenced by the dideoxy chain termination method (Sanger *et al.* 1977) using M13mp8 and M13mp9 vectors (Messing and Vieira 1982). Both strands were sequenced using a series of overlapping clones generated by both directed cloning, using appropriate restriction enzymes, and by random cloning of *Hae*III and *Alu*I generated fragments. The primer used for most of these reactions was the univeral M13 17mer sequencing primer from New England Biolabs, Beverly, MA. Sequencing of Tn5 junctions was also carried out by subcloning *Xho*I-*Xho*I fragments from the appropriate Tn5-containing cosmids and by using a Tn5 sequencing primer (3′-TTCATCGCAGGACTTGC-5′) previously described (Schofield and Watson 1986). Samples were separated on standard 40 cm, 6% polyacrylamide, 8 M urea gels using TBE buffer (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA, pH 8.3). Gels were fixed in a mixture of 10% ethanol and 10% acetic acid, and dried before autoradiography. Several compressions were encountered in the sequence, and these were resolved by running 6% acrylamide gels containing 7 M urea and 40% formamide, as described in an article on overcoming DNA sequencing artifacts in the ninth volume of *Focus*, published by Bethesda Research Laboratories, Gaithersburg, MD.

Computer analysis. DNA sequences were entered, assembled, and analyzed on a VAX750 using the University of Wisconsin Genetics Computer Group (UWGCG), Madison, package (Devereux *et al.* 1984).

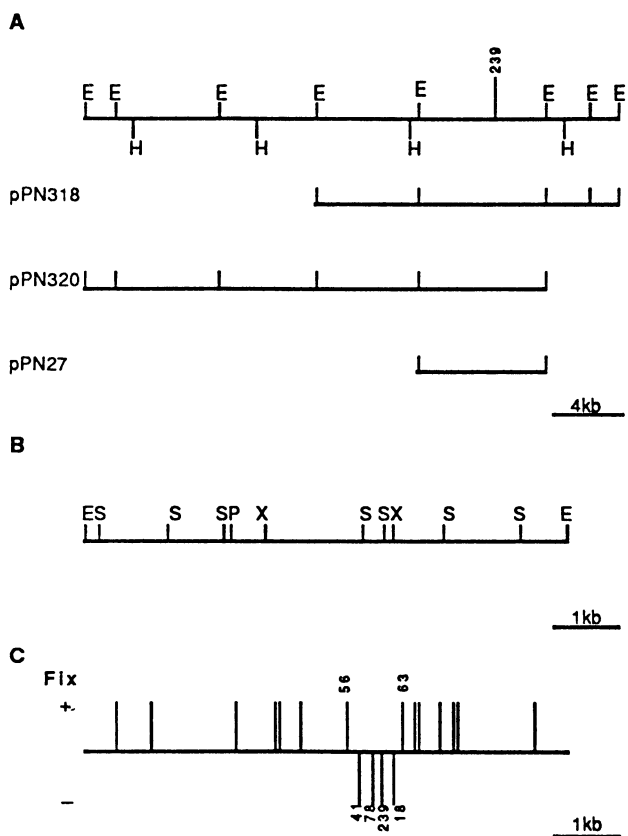

Fig. 1. A physical and genetic map of the *Rhizobium loti* invasion (*inv*) −239 locus. **A**, An *Eco*RI (E) and *Hin*dIII (H) restriction enzyme map of the *inv* −239 locus of *R. loti* strain NZP2037 and the cosmids used to construct the map. **B**, Restriction enzyme map of the 7.5-kb *Eco*RI fragment showing sites for *Sal*I (S), *Pst*I (P), and *Xho*I (X). **C**, Positions of Tn5 insertions in the 7.5-kb *Eco*RI fragment and the nodule phenotype of the corresponding *R. loti* mutants when inoculated onto *Lotus pedunculatus*.

RESULTS

Cloning and mapping of the *inv* gene. Cosmids previously isolated (Chua *et al.* 1985) by *in planta* complementation of the *R. loti* Fix⁻ mutant, strain PN239, and designated here as pPN318 and pPN320 were used to construct an *Eco*RI and *Hin*dIII restriction enzyme map of the *inv* gene locus (Fig. 1A). The map was constructed by comparing single and double digests of the cosmids using *Eco*RI and *Hin*dIII, and by hybridization of Southern blots of *Eco*RI and *Hin*dIII digests of the cosmids using individual *Hin*dIII fragments from the cosmids as probes. Both cosmids contained an *Eco*RI fragment of the size (7.5 kb) predicted to be the Tn5 insertion site in strain PN239 (Chua *et al.* 1985). This was confirmed by subcloning the 7.5-kb fragment from pPN318 into the *Eco*RI site of pBR328 and using this plasmid, pPN27, as a probe against a Southern blot of *Eco*RI cut total DNA from strains NZP2037 and PN239. Hybridization was observed to bands of 7.5 and 13.2 kb, respectively.

Complementation of strain PN239. To test whether the 7.5-kb fragment could complement strain PN239, a pLAFR1 recombinant plasmid containing this fragment, pPN28, was transferred to strain PN239 in a triparental cross. Six Tc^r-Nm^r-Str^r (^{r}, resistant) transconjugants from this cross were single-colony purified and inoculated onto *L. pedunculatus* seedlings. All tested transconjugants formed nodules within 2 wk of inoculation, and at 6 wk the plants were all effectively nodulated. Bacteria isolated from these nodules were Tc^r-Nm^r-Str^r.

Tn5 mutagenesis of the 7.5-kb *Eco*RI fragment containing the *inv* gene. To determine the boundaries of the *inv* locus identified here, further Tn5 mutagenesis of the 7.5-kb fragment was carried out in *E. coli* strain PN1001(HB101/pPN28) using phage λ467 as the Tn5 source. Transposon Tn5 insertions in pPN28 were selected by conjugating this plasmid into *E. coli* strain C2110, in a triparental cross, and selecting for Km^r-Nal^r-Tc^r transconjugants. Plasmid DNA was prepared from randomly selected colonies and cut with *Eco*RI to identify whether the Tn5 had inserted into the pLAFR1 vector or into the 7.5-kb *Eco*RI fragment. Inserts in the 7.5-kb fragment were mapped by preparing *Hin*dIII-*Eco*RI double digestions of each plasmid and then probing a Southern blot of these digests with the 2.8-kb *Eco*RI-*Xho*I fragment (Fig. 1B).

The positions of the isolated Tn5 insertions are shown in Figure 1C. Each of the pPN28::Tn5 cosmids was then transferred to *R. loti* strain PN184 in a further triparental cross, and the Tn5 homogenotized into the genome by plasmid incompatibility using pPH1JI. Total DNA was isolated from these Str^r-Gm^r-Tc^r colonies; this DNA was

cut with *Eco*RI and Southern blots of the digests were probed with the 7.5-kb *Eco*RI fragment to confirm that a double crossover had occurred. Each confirmed mutant was then inoculated onto *L. pedunculatus* to determine the symbiotic phenotype. Of the 17 Tn5-containing mutants tested, three (corresponding to insertions 18, 41, and 78) gave a Fix$^-$ phenotype and these all mapped close to the Tn5 insertion point in strain PN239 (Fig. 1C). The outermost boundaries of this *inv* locus were defined by insertions 56 and 63 (Fig. 1C). When strains PN1093, PN1094, and PN1095 (containing insertions 41, 78, and 18, respectively) were tested on *L. corniculatus*, they all formed fully effective nodules, as did strain PN239, indicating that the gene is host-specific. All mutants were prototrophic and formed mucoid colonies like the wild type.

Electron microscopy of *L. pedunculatus* nodules induced by Fix$^-$ mutants. To define the phenotype of the Fix$^-$ nodules induced by strains PN1093, PN1094, and PN1095 on *L. pedunculatus* and to compare them with nodules formed by the original mutant (PN239) isolated at this locus, sections of nodules were examined by light and electron microscopy. Nodule development on *L.*

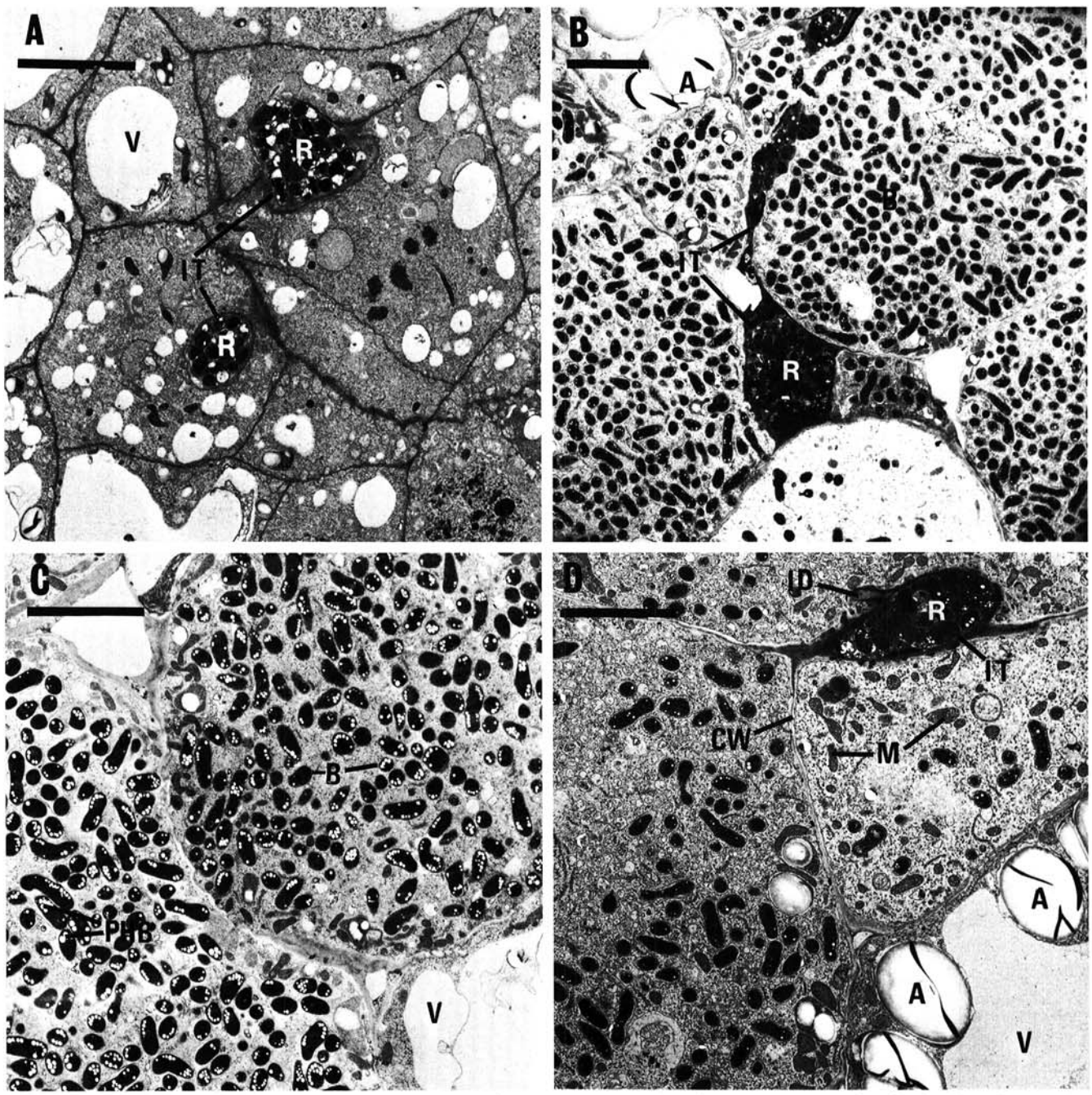

Fig. 2. Electron micrographs of sections from 28-day-old nodules formed on *Lotus pedunculatus* by mutant strains of *Rhizobium loti*. **A**, Strain PN239. The plant cells contain enlarged infection threads (IT) to which the rhizobia (R) are mainly confined. **B**, Strain PN1093. Intercellular infection threads are packed with rhizobia, and plant cells contain rhizobia that have been released to form bacteroidlike structures (B). **C**, Strain PN1094, released rhizobia containing deposits of polyhydroxybutyrate (PHB). **D**, Strain PN1095. Rhizobia are confined to infection threads and vacuolated (V) cells containing amyloplasts (A) are present. Other abbreviations: CW, cell wall; ID, infection droplet; and M, mitochondria. Scale bars = 5 μm.

pedunculatus induced by effective and ineffective *R. loti* strains has been previously described (Pankhurst *et al.* 1979; Chua *et al.* 1985), so only details relevant to a description of the mutant phenotypes are presented here. In analyzing the nodule phenotype of each mutant, nodules from at least four different plants were sectioned.

L. pedunculatus nodules induced by strain PN239 and the three other Fix$^-$ mutants described here were spherical and of normal external morphology, but when sections of the nodules were examined by light microscopy, considerable variation was found in the extent to which the rhizobia had invaded the plant cells. Where there was little evidence of invasion, electron micrographs (Fig. 2) showed that rhizobia were predominantly confined to infection threads (micrograph A) with only a few cells showing release of the bacteria into the plant cell. Examples of this phenotype were observed for nodules induced by strains PN239 and PN1095, but in both cases an examination of other nodules showed that release of rhizobia did occur. Where light micrographs revealed more extensive invasion of the central nodule cortex, electron micrographs (Fig. 2) of sections through this region showed the phenotype of the cells to be highly varied. Cells of varying cytoplasmic intensity containing released rhizobia (micrographs B and C) were interspersed with highly vacuolate cells containing starch granules (micrograph D).

Examples of rhizobia confined to infection threads (micrographs B and D) were also observed. Rhizobia released into plant cells showed the pleiomorphic forms characteristic of bacteroids; but when compared with wild-type nodules, there was extensive breakdown of the PBM, and the bacterial cells contained deposits of poly-hydroxybutyrate. The different intensities of the plant cytoplasm in these cells presumably reflected different stages of cellular senescence.

DNA sequence analysis. The strategy used to determine the sequence of the *inv* locus, with boundaries defined by Tn5 insertions 56 and 63 (Fig. 1), is shown in Figure 3. Both strands were sequenced by the chain termination method using a series of overlapping clones generated by subcloning various restriction enzyme fragments from the wild-type DNA and *Xho*I-*Xho*I fragments containing the junction of Tn5 and *Rhizobium* DNA into M13mp8 and M13mp9 vectors (Fig. 3). The DNA sequence of this region is shown in Figure 4. Six-phase translation of this sequence revealed two major (> 300 bp) open-reading frames (ORFs), one in each orientation. Analysis of this entire sequence for potential coding regions using the programs Testcode (Fig. 5A) (Fickett 1982), Codon Preference (Fig. 5C) (Gribskov *et al.* 1984), and Third Position GC Bias (Fig. 5B) (Bibb *et al.* 1984) predicted a coding sequence of 576 bp starting with the ATG codon at position 376 and ending with a TGA (stop) codon at position 952. Analysis of the complement of this sequence using the same programs suggested that the second identified ORF was noncoding. The position and orientation of the ORF shown in Figure 4 are consistent with the Tn5 insertional inactivation data (Fig. 1).

A potential ribosomal binding site GACGAGGG (Shine and Dalgarno 1974; Stormo *et al.* 1982) was observed at positions −15 to −8 with respect to the ATG start codon (361–368). This is consistent with the location of the predicted coding region. Upstream of this presumptive translational start site is a putative *Rhizobium* promoter sequence, −293 to −272 (84–105), identified by analyzing the entire sequence for the consensus sequence TTPuANN...17b....PuAPuPuPuPu..4-5b..CA (Ronson and Astwood 1985) using the UWGCG programs Bestfit (Smith and Waterman 1981) and Gap (Needleman and Wunsch 1970). No homology was found in this upstream region to the *ntr*A-activable consensus sequence CTGGYAYR-N_4-TTGCA (Gussin *et al.* 1986). Several inverted repeat sequences that have the potential to form stemloop structures were identified downstream of the stop codon (Fig. 4). The deduced amino acid sequence of the protein contained 192 residues with an unmodified molecular weight of 21.3 kDa. The amino acids tryptophan and cysteine were not represented in the sequence.

An examination of the charged residues would suggest that the protein has a net positive charge of eight (Kyte and Doolittle 1982) with 34 aspartate and glutamate residues, 31 arginine and lysine residues, and 7 histidines. No significant similarities to proteins in the National Biomedical Research Foundation (NBRF) protein resource data base (version 16) (Sidman *et al.* 1988) were seen. There was also no significant homology to any DNA sequence in the GenBank (version 56) (Bilofsky and Burks 1988) or European Molecular Biology Laboratory (EMBL) (version 15) (Cameron 1988) DNA sequence data bases or to the NBRF Nucleotide Sequence Data Library (version 33).

DISCUSSION

Rhizobium invasion of a plant cell is a tightly regulated process between host and symbiont. Any imbalance between the partners may result in a symbiotic defect such as infection thread abortion (Noel *et al.* 1986), a failure of the bacteria to be released (Pankhurst 1974; Pankhurst

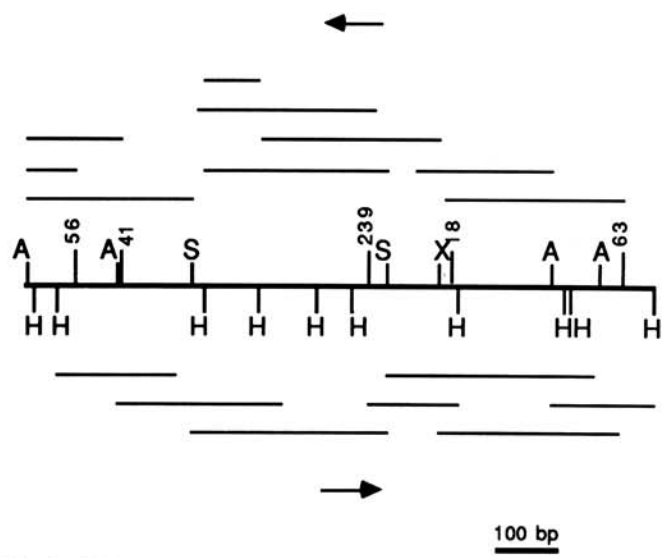

Fig. 3. DNA sequencing strategy for the invasion (*inv*) gene from *Rhizobium loti* strain NZP2037. Arrows show the direction in which each clone was sequenced, and bars show the length and position of each restriction fragment and Tn5 junction fragment sequenced. The sequenced position of each Tn5 insertion is indicated. Restriction sites have been abbreviated as follows: A, *Alu*I; H, *Hae*III; S, *Sal*I; and X, *Xho*I.

et al. 1972; Noel *et al.* 1982; Truchet *et al.* 1980; Morrison and Verma 1987) or failure of the rhizobia to differentiate into mature bacteroids (Ma *et al.* 1982; Chua *et al.* 1985; Hirsch and Smith 1987).

In this work we have identified a gene required for effective nodulation of *L. pedunculatus*. DNA sequence analysis identifies an ORF of 576 bp coding for an unmodified protein of 21.3 kDa. The position and orientation of this sequence are consistent with the transposon Tn5 insertion data shown in Figure 1. Although an earlier report (Chua *et al.* 1985) indicated that this locus was involved in bacterial release from the infection thread, the isolation here by allelic replacement of three further Fix$^-$ mutations within this locus and an analysis of their nodule phenotype by electron microscopy show that the phenotype is much more complex than just a block in release. While examples of rhizobia confined to infection threads were observed, many of the examined nodules showed release of rhizobia that had failed to develop into mature bacteroids (Fig 2). In some cases there was extensive degeneration of the PBM, a phenotype characteristic of a number of *nif* and *fix* mutations (Ma *et al.* 1982; Hirsch and Smith 1987). However, no significant homology was detected between the gene sequence presented here and

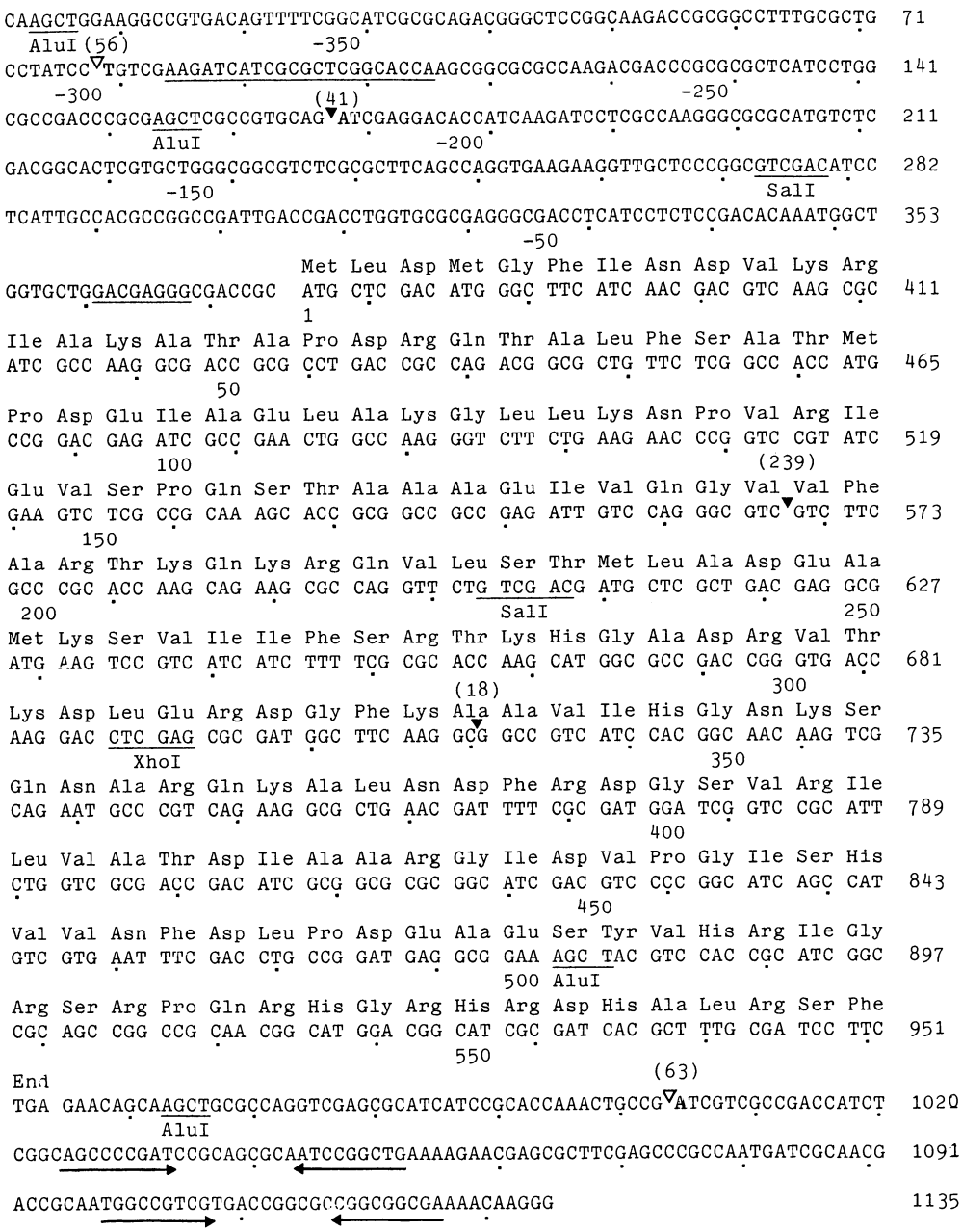

Fig. 4. DNA sequence and deduced amino acid sequence of the *Rhizobium loti* invasion (*inv*) gene. The numbers below the lines refer to the nucleotide positions in relation to the ATG initiation codon (1). The first digit is aligned with the numbered nucleotide. A putative promoter region (84 to 105) (−293 to −272) and ribosome binding site (361 to 368) (−15 to −8) are underlined. Possible stemloop structures in the terminal region are indicated by arrows. The positions of Tn5 insertions in the sequence are shown by ▽ (Fix$^+$) and ▼ (Fix$^-$). The number given in brackets above the symbol corresponds to the number assigned to the Tn5 insertion in Figure 3.

any other *Rhizobium* sequence in the current versions of the EMBL and GenBank data bases, including both *nif* and *fix* genes and a recently reported sequence of a *B. japonicum* gene involved in bacterial release (Morrison and Verma 1987).

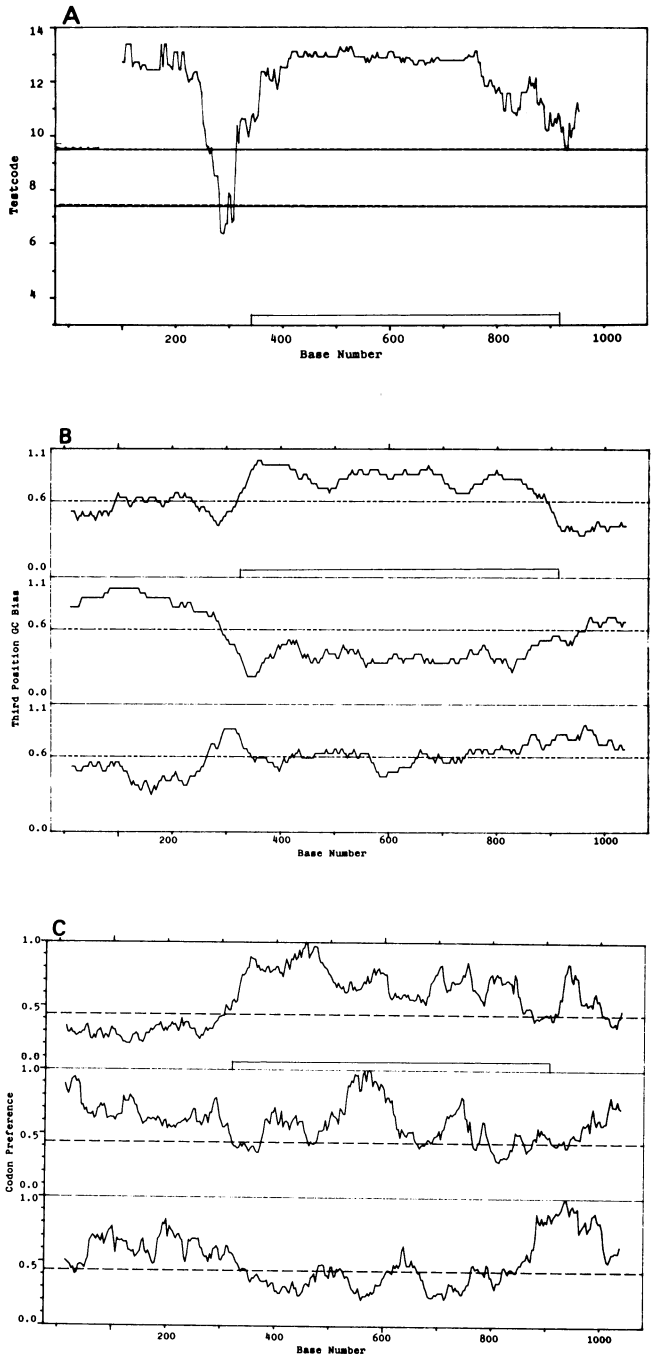

Fig. 5. Analysis of the sequence presented in Figure 4. The bar in each figure indicates the open-reading frame for the invasion (*inv*) gene. **A**, Testcode analysis (Fickett 1982). Two dotted horizontal lines divide the figure into three regions. Where the plot is above the two lines, the sequence is predicted to be coding; below the two lines, noncoding; and between the two lines, Testcode makes no prediction. **B**, Third Position GC Bias (Bibb *et al.* 1984) for each reading frame in the 5′ to 3′ direction. Nonrandom GC usage, indicative of possible coding regions, is indicated by the plot rising above the horizontal dotted line. **C**, Codon Preference (Gribskov *et al.* 1984). The codon usage in each reading frame in the 5′ to 3′ direction is compared with an *Escherichia coli* codon frequency table. Where the codon usage is more like the codon frequency table than a random sequence, the plot rises above the horizontal dotted line.

The overall capacity of the mutants to invade the cortical cells of *L. pedunculatus* was reduced, a situation observed for Fix$^-$ mutations of *Bradyrhizobium* sp. (*Arachis*) (Wilson *et al.* 1987). Even within those cells where infection and release occurred, cell-to-cell variation was observed, indicative of a "titration" effect by some host factor. Although the function of the gene product deduced here is not known, the phenotype of the nodules would suggest that the tight coupling between host and symbiont is broken down, resulting in a "resistance"-type response by the host. This could take the form of a cell surface change resulting in an increased sensitivity by the bacterium to toxic plant substances.

L. pedunculatus is known to produce high levels of condensed tannins, especially the flavolan prodelphinidin, and the presence of these compounds has been correlated with the inability of certain *R. loti* strains to form an effective symbiosis with this host (Pankhurst *et al.* 1979). However, strain NZP2037 is one of a small group of strains that is capable of forming an effective symbiosis with *L. pedunculatus*, and it is easy to envision how a slight perturbation of the cell surface of this strain could restrict its host range effectiveness. On the other hand, the mutation could result in a biochemical change that reduces the ability of the bacterium to elicit the appropriate response in the host, and thereby leads to a block in cell invasion.

In support of the latter hypothesis is the recent observation by Morrison and Verma (1987) that a mutation in *B. japonicum* which prevents release of the bacterium into the host cell differentially influences the induction of PBM nodulins. While the synthesis of one PBM nodulin (nodulin 26) was unaffected by the mutation, two others, nodulins 23 and 24, were not fully expressed. Morrison and Verma suggest that full induction of these specific PBM nodulins is required for efficient release of rhizobia and the development of mature bacteroids. This tight coupling between host and symbiont is also reflected by the fact that all the mutations isolated here are host-specific, that is they were Fix$^-$ on *L. pedunculatus* but Fix$^+$ on *L. corniculatus*. Wilson *et al.* (1987) have also isolated host-specific, nitrogen-fixation mutations of *Bradyrhizobium* sp. (*Arachis*) and a number of *Rhizobium* wild-type strains, including *R. loti* (Pankhurst 1977), and shown they are effective on one host and ineffective on another.

Our results demonstrate how tightly regulated the interaction is between host and symbiont and emphasize the need to identify the biochemical function of this gene product and those encoded by similar *inv* loci in order to explain the molecular basis of this interaction.

ACKNOWLEDGMENTS

We thank Doug Hopcroft and Raymond Bennett for assistance in the preparation of nodule samples for electron microscopy and for the electron micrographs. We also thank Mark Pritchard for help with computing.

This work was partly supported by the Riverside Farm Research Fund and the Massey University Research Fund. L. J. H. Ward was the recipient of a Leonard Condell Scholarship.

LITERATURE CITED

Bauer, W. D. 1981. Infection of legumes by rhizobia. Annu. Rev. Plant Physiol. 32:407-449.

Beringer, J. E. 1974. R-factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. 84:188-198.

Beringer, J. E., Beynon, J. L., Buchanan-Wollaston, A. V., and Johnston, A. W. B. 1978. Transfer of the drug resistant transposon Tn5 to *Rhizobium*. Nature 276:633-634.

Bibb, M. I., Findlay, P. R., and Johnson, M. W. 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene 30:157-166.

Bilofsky, H. S., and Burks, C. 1988. The GenBank® genetic sequence data bank. Nucleic Acids Res. 16:1861-1863.

Bolivar, F., Rodriguez, R., Greene, P. J., Betlach, M., Heyneker, H. L., Boyer, H. W., Crosa, I., and Falkow, S. 1977. Construction and characterisation of new cloning vehicles. II. A multi-purpose cloning system. Gene 2:95-113.

Boyer, H. W., and Roulland-Dussoix, D. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. 41:459-472.

Callaham, D. A., and Torrey, J. G. 1981. The structural basis for infection of root hairs of *Trifolium repens* by *Rhizobium*. Can. J. Bot. 59:1647-1664.

Calvert, H. E., Pence, M. K., Pierce, M., Malik, N. S. A., and Bauer, W. D. 1984. Anatomical analysis of the development and distribution of *Rhizobium* infections in soybean roots. Can. J. Bot. 62:2375-2384.

Cameron, G. N. 1988. The EMBL data library. Nucleic Acids Res. 16:1865-1867.

Chua, K. Y., Pankhurst, C. E., MacDonald, P. E., Hopcroft, D. H., Jarvis, B. D. W., and Scott, D. B. 1985. Isolation and characterization of transposon Tn5-induced symbiotic mutants of *Rhizobium loti*. J. Bacteriol. 162:335-343.

Clewell, D. B., and Helinski, D. R. 1969. Supercoiled DNA-protein complex in *Escherichia coli*: Purification and induced conversion to an open circular form. Proc. Natl. Acad. Sci. USA 62:1159-1166.

Dart, P. J. 1977. Infection and development of leguminous nodules. A Treatise on Dinitrogen Fixation. R. W. F. Hardy, ed. Pages 367-472 in: Section 3, Biology. W. S. Silver, ed. Wiley (Interscience), New York.

de Bruijn, F. J., and Lupski, J. R. 1984. The use of transposon Tn5 mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids - A review. Gene 27:131-149.

Devereux, J., Haeberli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12:387-395.

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 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.

Dixon, R., Cannon, F., and Kondorosi, A. 1976. Construction of a P plasmid carrying nitrogen fixation genes from *Klebsiella pneumoniae*. Nature 260:268-271.

Dylan, T., Ielpi, L., Stanfield, S., Kashyap, L., Douglas, C., Yanofsky, M., Nester, E., Helinski, D. R., and Ditta, G. 1986. *Rhizobium meliloti* genes required for nodule development are related to chromosomal virulence genes in *Agrobacterium tumefaciens*. Proc. Natl. Acad. Sci. USA 83:4403-4407.

Fickett, J. W. 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. 10:5303-5318.

Finan, T. M., Hirsch, A. M., Leigh, J. A., Johansen, E., Kuldan, G. A., Deegan, S., Walker, G. C., and Signer, E. R. 1985. Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. Cell 40:869-877.

Fischer, S. G., and Lerman, L. S. 1979. Length-independent separation of DNA restriction fragments in two dimensional gel electrophoresis. Cell 16:191-200.

Fortin, M. G., Zelechowska, M., and Verma, D. P. S. 1985. Specific targeting of membrane nodulins to the bacteroid-enclosing compartment in soybean nodules. EMBO J. 4:3041-3046.

Friedman, A. M., Long, S. R., Brown, S. E., Buikema, W. J., and Ausubel, F. M. 1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. Gene 18:289-296.

Goodchild, D. J. 1978. The ultrastructure of root nodules in relation to nitrogen fixation. Int. Rev. Cytol. 41:235-288.

Gribskov, M., Devereux, I., and Burgess, R. R. 1984. The codon preference plot: Graphic analysis of protein coding sequences and prediction of gene expression. Nucleic Acids Res. 12:539-549.

Gussin, G. N., Ronson, C. W., and Ausubel, F. M. 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. 20:567-591.

Hirsch, A. M., and Smith, C. A. 1987. Effects of *Rhizobium meliloti* *nif* and *fix* mutants on alfalfa root nodule development. J. Bacteriol. 169:1137-1146.

Holmes, D. S., and Quigley, M. 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114:193-197.

Ish-Horowicz, D., and Burke, J. F. 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. 9:2989-2998.

Kondorosi, E., and Kondorosi, A. 1986. Nodule induction on plant roots by *Rhizobium*. TIBS 11:296-298.

Kyte, J., and Doolittle, R. F. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

Leigh, J. A., Signer, E. R., and Walker, G. C. 1985. Exopolysaccharide-deficient mutants of *Rhizobium meliloti* that form ineffective nodules. Proc. Natl. Acad. Sci. USA 82:6231-6235.

Leong, S. A., Ditta, G. S., and Helinski, D. R. 1982. Heme biosynthesis in *Rhizobium*. J. Biol. Chem. 257:8724-8730.

Ma, Q.-S., Johnston, A. W. B., Hombrecher, G., and Downie, J. A. 1982. Molecular genetics of mutants of *Rhizobium leguminosarum* which fail to fix nitrogen. Mol. Gen. Genet. 187:166-171.

Messing, J., and Vieira, J. 1982. A new pair of M13 vectors for selecting either DNA strand or double-digest restriction fragments. Gene 19:269-276.

Miller, J. H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Morrison, N., and Verma, D. P. S. 1987. A block in the endocytosis of *Rhizobium* allows cellular differentiation in nodules but affects the expression of some peribacteroid membrane nodulins. Plant Mol. Biol. 9:185-196.

Needleman, S. B., and Wunsch, C. D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48:443-453

Newcomb, W. 1981. Nodule morphogenesis and differentiation. Int. Rev. Cytol. Suppl. 13:247-298.

Noel, K. D., Stacey, G., Tandon, S. R., Silver, L. E., and Brill, W. J. 1982. *Rhizobium japonicum* mutants defective in symbiotic nitrogen fixation. J. Bacteriol. 152:485-494.

Noel, K. D., Vandenbosch, K. A., and Kulpaca, B. 1986. Mutations in *Rhizobium phaseoli* that lead to arrested development of infection threads. J. Bacteriol. 168:1392-1401.

Pankhurst, C. E. 1974. Ineffective *Rhizobium trifolii* mutants examined by immune-diffusion, gel-electrophoresis and electron microscopy. J. Gen. Microbiol. 82:405-413.

Pankhurst, C. E. 1977. Symbiotic effectiveness of antibiotic resistant mutants of fast- and slow-growing strains of *Rhizobium* nodulating *Lotus* species. Can. J. Microbiol. 23:1026-1033.

Pankhurst, C. E., Craig, A. S., and Jones, W. T. 1979. Effectiveness of *Lotus* root nodules. I. Morphology and flavolan content of nodules formed on *Lotus pedunculatus* by fast-growing *Lotus* rhizobia. J. Exp. Bot. 30:1085-1093.

Pankhurst, C. E., Schwinghamer, E. A., and Bergersen, F. J. 1972. The structure and acetylene-reducing activity of root nodules formed by a riboflavin requiring mutant of *Rhizobium trifolii*. J. Gen. Microbiol. 70:161-177.

Robertson, J. G., Lyttleton, P., Bullivant, S., and Grayston, G. F. 1978. Membranes in lupin root nodules. I. The role of Golgi bodies in the biogenesis of infection threads and peribacteroid membranes. J. Cell. Sci. 30:129-149.

Robertson, J. G., and Farnden, K. J. F. 1980. Ultrastructure and metabolism of the developing legume root nodule. Pages 65-113 in: The Biochemistry of Plants, Vol. 5. E. Cohn, ed. Academic Press, New York.

Ronson, G. W., and Astwood, P. M. 1985. Genes involved in the carbon metabolism of bacteroids. Pages 201-207 in: Nitrogen Fixation Research Progress. H. J. Evans, P. J. Bottomley, and W. E. Newton, eds. Martinus Nijhoff, Boston.

Rossen, L., Davis, E. O., and Johnston, A. W. B. 1987. Plant-induced expression of *Rhizobium* genes involved in host specificity and early stages of nodulation. TIBS 12:430-433.

Ruvkun, G. B., and Ausubel, F. M. 1981. A general method for site-directed mutagenesis in prokaryotes. Nature 289:85-88.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.

Schofield, P. R., and Watson, J. M. 1986. DNA sequence of *Rhizobium trifolii* nodulation genes reveals a reiterated and potentially regulatory sequence preceding *nod*ABC and *nod*FE. Nucleic Acids Res. 14:2891-2903.

Scott, D. B., Court, C. B., Ronson, C. W., Scott, K. F., Watson, J. M., Schofield, P. R., and Shine, J. 1984. Organization of nodulation

and nitrogen fixation genes on a *Rhizobium trifolii* symbiotic plasmid. Arch. Microbiol. 139:151-157.

Scott, D. B., and Ronson, C. W. 1982. Identification and mobilization by cointegrate formation of a nodulation plasmid in *Rhizobium trifolii*. J. Bacteriol. 151:36-43.

Scott, D. B., Chua, K. Y., Jarvis, B. D. W., and Pankhurst, C. E. 1985. Molecular cloning of a nodulation gene from fast- and slow-growing strains of *Lotus* rhizobia. Mol. Gen. Genet. 201:43-50.

Shine, J., and Dalgarno, L. 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA 71:1342-1346.

Sidman, K. E., George, D. G., Barker, W. C., and Hunt, L. T. 1988. The protein identification resource (PIR). Nucleic Acids Res. 16:1869-1871.

Smith, T. F., and Waterman, M. S. 1981. Comparison of biosequences. Adv. Appl. Math. 2:482-489.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Stormo, G. D., Schneider, T. D., and Gold, L. M. 1982. Characterization of translational initiation sites in *E. coli*. Nucleic Acids Res. 10:2971-2996.

Truchet, G., Michel, M., and Dénarié, J. 1980. Sequential analysis of the organogenesis of lucerne (*Medicago sativa*) root nodules using symbiotically defective mutants of *Rhizobium meliloti*. Differentiation 16:163-172.

Vandenbosch, K. A., Noel, K. D., Kaneko, Y., and Newcomb, E. H. 1985. Nodule initiation elicited by noninfective mutants of *Rhizobium phaseoli*. J. Bacteriol. 162:950-959.

Whitfeld, P. L., Seeburg, P. H., and Shine, J. 1982. The human proopiomelanocortin gene: Organisation, sequence and interspersion with repetitive DNA. DNA 1:133-143.

Wilson, K. J., Anjaiah, V., Nambiar, P. T. C., and Ausubel, F. M. 1987. Isolation and characterisation of symbiotic mutants of *Bradyrhizobium* sp. (*Arachis*) strain NC92: Mutants with host-specific defects in nodulation and nitrogen fixation. J. Bacteriol. 169:2177-2186.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119.