CHAPTER 3

CHAPTER 3
MATERIALS AND METHODS

3.1 STRAINS USED

Nineteen indigenous rhizobia isolates previously characterised by Dagutat (1995), using SDS-PAGE of WCP, were included in the study together with nine strains isolated by Kruger (1998) and two isolates received from the Agricultural Research Council (Table 2).

TABLE 2 **List of indigenous rhizobia isolated by Dagutat (1995), isolates received from the Agricultural Research Council and rhizobia isolated by Kruger (1998) used in 16S rDNA analysis.**

Isolate	Host legume	Subfamily of Fabaceae
13b	*Lotononis bainesii*	Papilionoideae
13c sterre	*Lotononis bainesii*	Papilionoideae
15c	*Desmodium tortuosum*	Papilionoideae
24 slym	*Desmodium tortuosum*	Papilionoideae
26c	*Strongylodon macrobotrys*	Papilionoideae
27d	*Teramnus labialis*	Papilionoideae
46c2	*Acacia sieberana* var. *woodii*	Mimosoideae
47c3a	*Cassia floribunda*	Caesalpinoideae
48a	*Tephrosia purpurea*	Papilionoideae
48b	*Tephrosia purpurea*	Papilionoideae
49b	*Indigofera melanadenia*	Papilionoideae
68a onseker	*Mucuna coriacea*	Papilionoideae
68d	*Mucuna coriacea*	Papilionoideae
70a	*Crotalaria brachycarpa*	Papilionoideae
79c	*Acacia caffra*	Mimosoideae
85a1t1 onseker	*Acacia xanthophloea*	Mimosoideae
94	*Vigna subterranea*	Papilionoideae
102a	*Chamaecrista biensis*	Caesalpinoideae
103b	*Indigofera hilaris*	Papilionoideae
M3b[#]	*Lotononis falcata*	Papilionoideae
M3c[#]	*Lotononis falcata*	Papilionoideae
PL10 groot knop[*]	*Argyrolobium tomentosum*	Papilionoideae
PL18b[#]	*Medicago sativa*	Papilionoideae
PL19b[#]	*Indigofera verrucosa*	Papilionoideae
PL20a[#]	*Lessertia annularis*	Papilionoideae
PL20b onseker[#]	*Lessertia annularis*	Papilionoideae
PL27a[#]	*Chamaecrista mimosoides*	Caesalpinoideae
PL3[*]	*Tephrosia grandiflora*	Papilionoideae
TK1[#]	*Senna petersiana*	Caesalpinoideae
W1c[#]	*Vigna unguiculata*	Papilionoideae

* Received from the Plant Protection Research Institute, Agricultural Research Council, Roodeplaat.

\# Putative rhizobia isolated by Kruger (1998)

3.2 MAINTENANCE OF CULTURES

The isolates were maintained on yeast mannitol agar (YMA) [1% (m/v) mannitol (UniVar), 0,5% (m/v) K_2HPO_4 (Merck), 0,02% (m/v) $MgSO_4.7H_2O$ (Merck), 0,01% (m/v) NaCl (NT Chemicals), 0,04% (m/v) yeast extract (Biolab) and 1,5% (m/v) bacteriological agar (Biolab)] slants and the long-term storage of the isolates was done in glycerol. The isolates were grown in yeast mannitol broth (YMB) for 5 to 7 d at 25 - 28°C with vigorous shaking. The broth cultures were mixed 1:1 with sterile 50% (v/v) glycerol (Merck) in sterile cryotubes and stored in duplicate at -20°C and -70°C.

3.3 EXTRACTION OF GENOMIC DNA

Each of the strains was used to inoculate 5 ml YMB in a screw-cap tube. The broth cultures were incubated on a rotary shaker for 5 to 7 d at 25 - 28°C. After the growth period the broth appeared turbid; 2 ml of the turbid broth culture was used to inoculate 15 ml sterile tryptone yeast (TY) broth [0,5% (m/v) tryptone (Difco), 0,3% (m/v) yeast extract (Biolab), 0,13% (m/v) $CaCl_2.6H_2O$ (UniLab)] in a screw-cap tube. TY broth reduces slime formation by the rhizobia. The broth cultures were incubated for 2 d at 25 - 28°C with vigorous shaking.

A modified method described by Ehlers (1995) was used to extract genomic DNA. The method is useful for organisms such as rhizobia, which produce large amounts of slime. Before harvesting the cells, the broth culture was cooled to 4°C for 30 min on ice to reduce the activity of the cells' enzymes. The cells were then harvested by centrifugation at 8 000 rpm for 15 min. After centrifugation, the supernatant was discarded. The pellet was resuspended in 1 ml of STE buffer (10 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, pH 8,0) (4°C) with gentle swirling motions, while keeping it on ice to prevent the cells from bursting. After resuspension of the pellet, 150 μl of 20% (m/v) SDS (UniVar) and 50 μl of proteinase K (20 mg/ml) (Roche Molecular Biochemicals) were added to the cell suspension. The mixture was incubated overnight at 50°C. If, after the incubation period, the mixture was still not clear because undigested proteins were left, the incubation was repeated. Two gentle extractions with equal volumes of mixed phenol:chloroform:isoamylalcohol [25:24:1 (v/v) phenol to chloroform to isoamylalcohol] were carried out. The extraction mixture was mixed with gentle movements of the wrist. The extraction mixture was centrifuged for 10 min at 8 000

rpm to separate the phases. The aqueous DNA-containing top layer was removed with a wide bore pipet without disturbing the white protein-rich interface. A final extraction with an equal volume of chloroform:isoamylalcohol [24:1 (v/v) chloroform to isoamylalcohol] removed all phenol from the DNA suspension. The mixture was centrifuged for 10 min at 8 000 rpm and the aqueous DNA layer removed with a wide bore pipet. The aqueous DNA top layer was adjusted to 0,3 M sodium acetate with a 3 M sodium acetate stock solution (pH 5,2) (SAARchem). Two volumes of absolute ethanol (Merck) (-20°C) were added to the mixture, mixed gently and DNA precipitated overnight at -20°C, or at -70°C for two hours. The suspension was centrifuged for 10 min at 8 000 rpm to pellet the DNA. The supernatant was discarded and the DNA pellet was washed with 300 μl 70% (v/v) ethanol (Merck) (-20°C) to remove the salt from the DNA. After centrifugation for 6 min at 8 000 rpm, the supernatant was discarded. The washing step was repeated. After the second washing step the pellet was vacuum dried. The pellet was then dissolved in 100 μl sterile double-distilled water and stored at -20°C until used.

Aliquots (1 μl) of genomic extractions were examined by horizontal agarose gel electrophoresis (Sambrook, Fritsch and Maniatis, 1989) using 0,9% (m/v) agarose gels (Promega) in 1x TAE buffer (40 mM Tris-HCl, 20 mM NaOAc and 1 mM EDTA pH 8,5), stained with ethidium bromide (10 mg/ml) to determine the success of the extraction. The genomic DNA was visualised by UV fluorescence. Any contaminating RNA was removed from the extract by incubation with RNaseI (Epicentre Technologies) for 90 min at 37°C.

3.4 AMPLIFICATION OF THE 16S rDNA GENE

The universal primers fD1 and rP2 (Weisburg *et al.*, 1991) (Table 3) were used to amplify the 16S rDNA gene. The primers were synthesized by Roche Molecular Biochemicals. The PCR reaction was carried out in a volume of 50 μl with approximately 50 ng of genomic DNA, 50 pmole each of the universal primers fD1 and rP2, 1,5 mM $MgCl_2$, 10x Buffer [20 mM Tris-HCl (pH 8), 100 mM KCl, 0,1 mM EDTA, 1 mM DTT, 0,5% Tween®20, 0,5% Nonidet P-40® and 50% glycerol], 2,5 mM each of dATP, dCTP, dGTP, dTTP and 0,5 U *Takara Taq* polymerase (TaKaRa Biomedicals, Japan). The 50 μl reaction volume was overlaid with two drops of sterile mineral oil (Sigma). Amplification of the reactions was carried out in a Hybaid Omnigene Thermocycler with the temperature profile described by Laguerre *et al.*

(1994). This consisted of an initial "hot start" at 95°C for 3 min to denature the DNA, followed by 35 cycles consisting of denaturation at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 2 min. A final extension step at 72°C for 3 min was performed after the 35 cycles were completed.

Aliquots (5 μl) of the amplification reactions were examined by horizontal agarose gel electrophoresis (Sambrook *et al.*, 1989) using 0,9% (m/v) agarose gels (Promega) in 1x TAE buffer (40 mM Tris-HCl, 20 mM NaOAc and 1 mM EDTA pH 8,5) and stained with ethidium bromide (10 mg/ml). The amplified gene was visualised by UV fluorescence and the size of the product was estimated using DNA molecular weight standards (Molecular Weight Marker VI, Roche Molecular Biochemicals).

TABLE 3 Universal primers used for amplification of the 16S rDNA gene.

Universal Primer	Primer sequence (5' to 3')	Reference
Forward primer: fD1 (8 -27)*	AGAGTTTGATCCTGGCTCAG	Weisburg *et al.*, 1991
Reverse primer: rP2 (1491 - 1506)*	ACGGCTACCTTGTTACGACTT	Weisburg *et al.*, 1991

* The numbers indicate the positions relative to the *E. coli* 16S rRNA gene sequences (GenBank accession number J01859)

3.5 PURIFICATION OF THE AMPLIFIED 16S rDNA GENE

The amplified products were purified using the High Pure PCR Product Purification kit from Roche Molecular Biochemicals. The PCR product (40 μl) was transferred into a sterile Eppendorf tube and 400 μl of the binding buffer was added. The mixture was mixed and transferred to the assembled High Pure filter tube and collection tube and centrifuged for 30 sec at 13 000 rpm. After centrifugation, the filtrate (flow through) was discarded and the same filter tube and collection tube were reassembled. To the filter tube, 500 μl of the wash buffer was added. The tubes were centrifuged for 30 sec at 13 000 rpm, the filtrate discarded and 200 μl of the wash buffer added to the assembled tubes. Again the tubes were centrifuged at 13 000 rpm for 30 seconds, the filtrate discarded and the filter tube placed into a new sterile Eppendorf tube. The PCR product was eluted with 30 μl of sterile double-distilled water. The efficiency of the purification and the concentration of the DNA was determined by horizontal agarose gel electrophoresis (Sambrook *et al.*, 1989) using 0,9% (m/v) agarose gels (Promega) in 1x TAE buffer (40 mM Tris-HCl, 20 mM NaOAc and 1 mM EDTA pH 8,5) and stained with ethidium bromide (10 mg/ml).

3.6 SEQUENCING OF THE 16S rDNA GENE

The 16S rDNA gene was sequenced with the primers 16SRNAII-S and 16SRNAVI-S (Kuhnert, Capaul, Nicolet and Frey, 1996) (Table 4). The primers are internal. Primer 16SRNAII-S is a forward primer binding from position 682 to 702 and 16SRNAVI-S is a reverse primer starting from the same position. The sequencing reaction was performed with the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction kit (with AmpliTaq® DNA Polymerase, FS) (Perkin Elmer Applied Biosystems). The PCR sequencing reaction was performed in volumes of 5 µl in 0,2 ml MicroAmp PCR tubes (Perkin Elmer Applied Biosystems). The reaction mix consisted of 3,2 pmole primer (16SRNAII-S or 16SRNAVI-S), 2 µl terminator ready reaction mix [A-dye terminator, C-dye terminator, G-dye terminator, T-dye terminator, dITP, dATP, dCTP, dTTP, Tris-HCl (pH 9,0), $MgCl_2$, thermal stable pyrophosphatase and AmpliTaq DNA polymerase] and 50 ng template DNA. The PCR reactions were carried out in a Perkin Elmer GeneAmp PCR System 2400 thermocycler and consisted of 25 cycles of denaturation at 96°C for 10 sec, annealing at 50°C for 5 sec and extension for 4 min at 60°C. The sequencing products were precipitated with 20 µl of 60% (v/v) ethanol (Merck) for 15 min at room temperature. The tubes were centrifuged for 15 min at 13 000 rpm to pellet the sequencing product, supernatants discarded and the sequencing products washed with 25 µl of 70% (v/v) ethanol (Merck) (-20°C). The tubes were centrifuged for 10 min at 13 000 rpm and the supernatants discarded. To remove all traces of ethanol the sequencing products were vacuum dried and stored at -20°C. For analysis, the purified products were resuspended in 3,5 µl Blue dextran/EDTA loading buffer (Perkin Elmer Applied Biosystems). The loading buffer was prepared by combining deionized formamide and 25 mM EDTA (pH 8,0) containing 50 mg/ml Blue dextran in a ratio of 5:1 formamide to EDTA/Blue dextran. The resuspended products were denatured for 2 min at 90°C and loaded onto the ABI Prism model 377 DNA sequencer gel.

TABLE 4 Universal primers used for the partial sequencing of the 16S rDNA gene.

Universal primers	Primer sequence (5' to 3')	Reference
16SRNAII-S (682 - 702)*	GTGTAGCGGTGAAATGCGTAG	Kuhnert *et al.*, 1996
16SRNAVI-S (702 - 682)*	CTACGCATTTCACCGCTACAC	Kuhnert *et al.*, 1996

* The numbers indicate the positions relative to the *E. coli* 16S rRNA gene sequence (GenBank accession number J01859)

3.7 PHYLOGENETIC ANALYSIS OF SEQUENCE DATA

The sequencing gels were analysed and sequences edited with the ABI Prism Sequencing Analysis 3.1 and the ABI Prism Sequencing Navigator 1.0.1 computer programmes (Perkin Elmer Applied Biosystems). All the sequences (direction 5' to 3') and the reference sequences (Table 5) were aligned with *Bradyrhizobium japonicum* (GenBank accession number: X87272) and edited. The sequences obtained from primer 16SRNAII-S were edited to a total length of 450 bp. The sequences obtained from primer 16SRNAVI-S were edited to a total length of 440 bp. The ClustalX programme was used to analyse the sequences, the reverse and forward sequences were analysed separately. A distance matrix was constructed by the pairwise alignment of the sequences. Scores were calculated as the number of identical residues in the best alignment of two sequences, minus a fixed gap penalty of ten. All scores were converted to distance by dividing percent identity by 100 and subtracting from 1.0 to give the amount of difference between the sequences being compared (Thompson, Higgins and Gibson, 1994). The neighbour-joining method (Saitou and Nei, 1987) was used to construct a phylogenetic tree from the distance matrix. Branch lengths were proportional to the estimated divergence along each branch. Confidence levels of the phylogenies were estimated with the bootstrap method (Felsenstein, 1985). The bootstrap method also tests the topology of the tree. The data was randomly resampled, creating a new data table, which was then analysed. A record was kept of all the groups of species that form monophyletic subsets in the resulting estimated phylogeny. The resampling and estimation procedure was repeated several times and only groups appearing in 95% or more of the trees were considered statistically significant. Bootstrap-supported trees were constructed using a random seed generator of 111 and 100 bootstrap trials. The phylogenetic trees were displayed using NJPlot which is an interactive visualiser of phylogenetic data.

TABLE 5 **List of reference strains included in the analysis of the 16S rDNA sequence data obtained from the GenBank[1].**

Current classification	Culture collection no.	Host plant	GenBank accession no.[1]
The genus *Allorhizobium*			
Allorhizobium undicola	LMG 11875[T]	*Neptunia natans*	Y17047
The genus *Azorhizobium*			
Azorhizobium caulinodans	ORS 571[T]	*Sesbania rostrata*	X94200
The genus *Bradyrhizobium*			
Bradyrhizobium elkanii	USDA 76	*Glycine max*	U35000
Bradyrhizobium japonicum	DSM 30131[T]	NS	X87272
The genus *Mesorhizobium*			
Mesorhizobium amorphae	ACCC 19665	*Amorpha fruticosa*	AF041442
Mesorhizobium ciceri	UPM-Ca7[T]	*Cicer arietinum*	U07934
Mesorhizobium huakuii	IFO 15243	NS	D13431
Mesorhizobium loti	LMG 6123	*Lotus divaricatus*	Y14159
Mesorhizobium plurifarium	LMG 11892	NS	Y14158
Mesorhizobium tianshanense	USDA 3592	*Glycyrrhiza pallidiflora*	AF041447
The genus *Rhizobium*			
Rhizobium etli	USDA 9032	*Phaseolus vulgaris*	U28916
Rhizobium galegae	LMG 6214[T]	*Galegae orientalis*	X67226
Rhizobium gallicum	Strain R602sp[T]	*Phaseolus vulgaris*	U86343
Rhizobium giardinii	Strain H152[T]	*Phaseolus vulgaris*	U86344
Rhizobium huautlense	USDA 4900[T]	*Sesbania herbacea*	AF025852
Rhizobium leguminosarum	IAM 12609	NS	D14513/D01269
Rhizobium mongolense	USDA 1832	*Medicago ruthenica*	U89816
Rhizobium tropici	IFO 15247	NS	D11344
The genus *Sinorhizobium*			
Sinorhizobium fredii	LMG 6217[T]	*Glycine max*	X67231
Sinorhizobium meliloti	IAM 12611	NS	D14509/D01265
Sinorhizobium saheli	LMG 7837[T]	*Sesbania cannabina*	X68390
Sinorhizobium teranga	LMG 6463	*Sesbania rostrata*	X68387
Sinorhizobium xinjiangensis	IAM 14142	NS	D12796

1	The GenBank database website address: *www.ncbi.nlm.nih.gov/GenBank/*
T	Type strain
ACCC	Agricultural Culture Collection of China
DSM	Deutsche Sammlung von Mikroorganismen, Germany
IAM	Institute of Applied Microbiology, University of Tokyo, Tokyo, Japan
IFO	Institute for Fermentation, Osaka, Yodogawa-ku, Osaka 532, Japan
LMG	Laboratorium voor Mikrobiologie Gent Culture Collection, State University Gent, Belgium
ORS	ORSTOM Collection, Institut Français de Recherche Scientifique pour le Développement en Coopération, Dakar, Senegal
UPM	Universidad Politécnica Madrid, Spain
USDA	United States Department of Agriculture, Agriculture Research Service, Beltsville, USA
NS	not stated

CHAPTER 4

......the sure and definite determination of species of bacteria requires so much time, so much acumen of eye and judgement, so much perseverance and patience that there is hardly anything else so difficult. - Mueller

CHAPTER 4
RESULTS

4.1 EXTRACTION OF GENOMIC DNA

The genomic DNA obtained after extraction was of a high quality (results not shown). The purification method removed bacterial polysaccharides, which can have a negative effect on the amplification reaction. Each DNA preparation appeared intact as only one large fragment was observed on electrophoresis. The concentration of RNA was generally very low.

4.2 AMPLIFICATION OF 16S rDNA GENE

The primers amplified a fragment of approximately 1 500 bp from each DNA preparation (results not shown). The concentration of the unpurified PCR product was approximately 250 ng/µl as determined by the method of Sambrook *et al.* (1989). Amplification products of a lower concentration generally did not give acceptable results after purification.

4.3 PURIFICATION OF AMPLIFIED PRODUCT

After purification, the concentration of the amplified fragment was less than 250 ng/µl and in the region of 180 to 200 ng/µl (results not shown). Some of the PCR product was regularly lost during purification. Problems were also experienced with the purification kit. Some of the filter tubes did not bind the DNA in the initial purification step and often all DNA was lost in the subsequent steps. Some of the tubes also did not purify the product. The dNTP's and other components of the PCR reaction remained in the "purified product".

4.4 SEQUENCING OF 16S rDNA GENE

Sequenced regions of between 450 and 600 bp were obtained with both primers. The primary structures (5' to 3') of the sequenced products are shown in Fig. 2 (primer 16SRNAVI-S) and Fig. 3 (primer 16SRNAII-S).

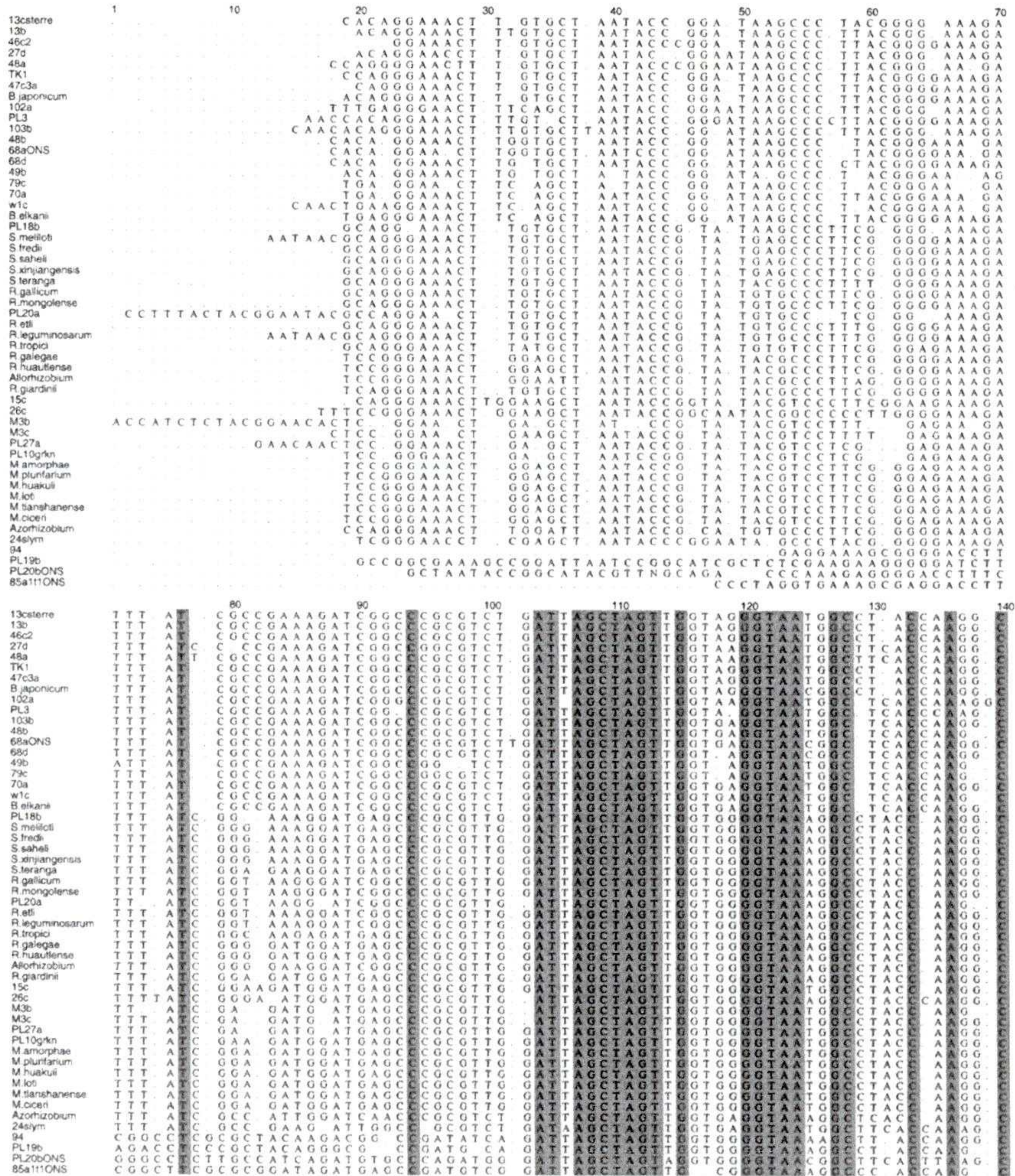

FIGURE 2 ALIGNMENT OF THE SEQUENCES OF THE 440 bp REGIONS OF THE 16S rDNA OF THE ISOLATES AND REFERENCE STRAINS INCLUDED IN THE STUDY. ALIGNMENTS WERE GENERATED BY ClustalX AND DISPLAYED WITH THE Shadybox COMPUTER PROGRAMME. THE REGIONS IN COLOUR ARE 100% HOMOLOGOUS.

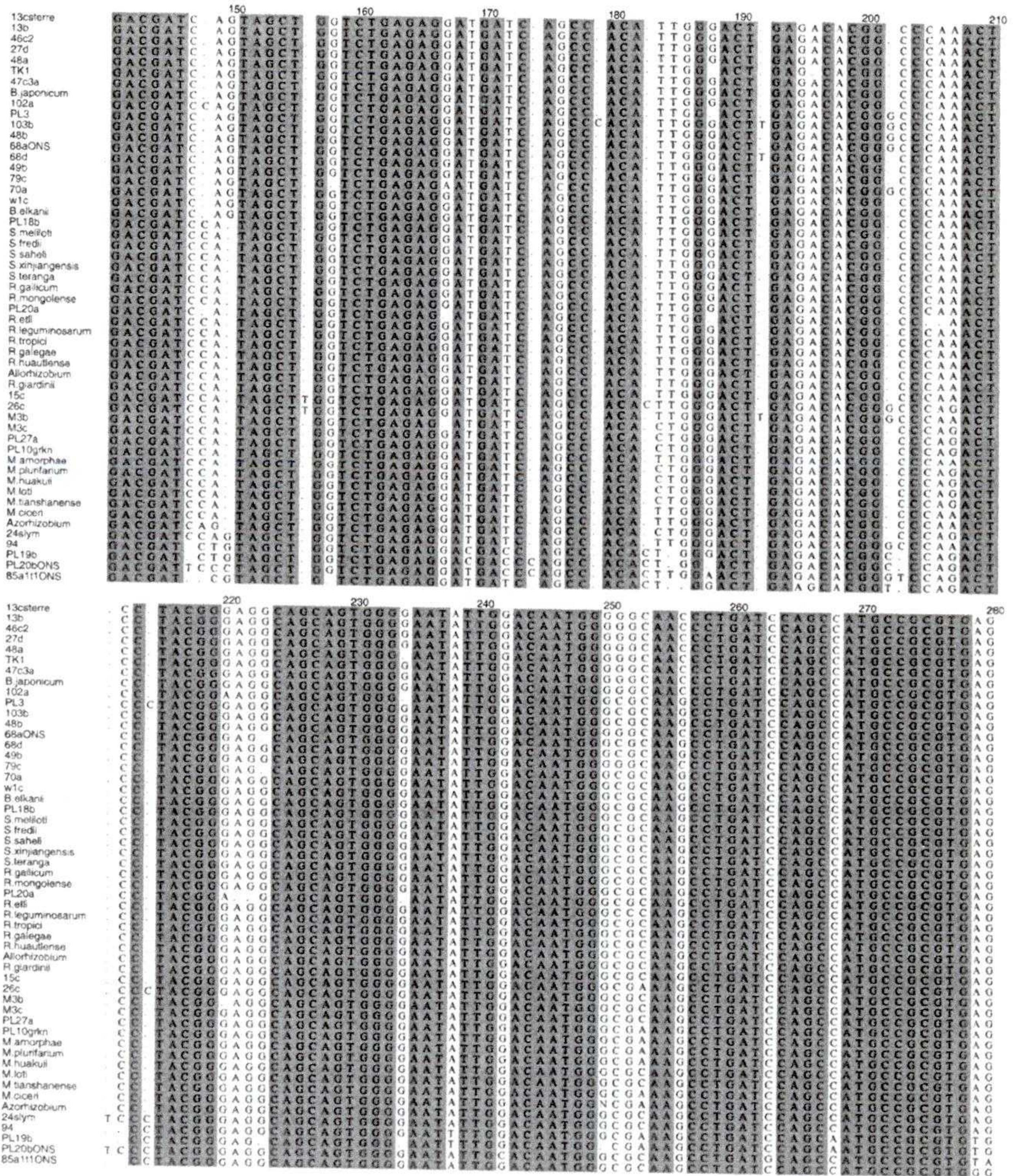

FIGURE 2 CONTINUED

FIGURE 2 CONTINUED

```
                430       440       450       460       470       480       490
13csterre       AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
13b             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
46c2            AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
27d             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
48a             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
TK1             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
47c3a           AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
B.japonicum     AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
102a            AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
PL3             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
103b            AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
48b             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
68aONS          AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
68d             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
49b             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
79c             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
70a             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
w1c             AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAACTCAGGGGTGAAATCCTGGA
B.elkanii       AGCGTTGCTCGGAATCACTGGGCGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGA
PL18b           AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATTGTTAAGTGAGGGGTGAAATCCCAGG
S.meliloti      AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATTGTTAAGTGAGGGGTGAAATCCCAGG
S.fredii        AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAGGGGTGAAATCCCGGG
S.saheli        AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAGGGGTGAAATCCCGGG
S.xinjiangensis AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAGGGGTGAAATCCCGGG
S.teranga       AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAGGGGTGAAATCCCGGG
R.gallicum      AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAGGGGTGAAATCCCAGA
R.mongolense    AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAGGGGTGAAATCCCAGA
PL20a           AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAGGGGTGAAATCCCAGA
R.etli          AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGATCAGTCAGGGGTGAAATCCCAGG
R.leguminosarum AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGATAAGTCAGGGGTGAAATCCCAGG
R.tropici       AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCCGATCGATCAGTCAGGGGTGAAATCCCAGG
R.galegae       AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATATTTAAGTCAGGGGTGAAATCCCAGA
R.huautlense    AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATATTTAAGTCAGGGGTGAAATCCCAGA
Allorhizobium   AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGGCATTTAAGTCAGGGGTGAAATCCCAGA
R.giardinii     AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGATCAGTCAGGGGTGAAATCCCAGA
15c             AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACTATTAAGTCAGGGGTGAAATCCCGGG
26c             AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATATTTAAGTCAGGGGTGAAATCCCGCA
M3b             AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATACTTAAGTCAGGGGTGAAATCCCGGG
M3c             AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATACTTAAGTCAGGGGTGAAATCCCGGG
PL27a           AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATACTTAAGTCAGGGGTGAAATCCCGGG
PL10grkn        AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATATTTAAGTCAGGGGT AAATCCCGGG
M.amorphae      AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATACTTAAGTCAGGGGTGAAATCCCGGG
M.plurifarium   AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATACTTAAGTCAGGGGTGAAATCCCGGG
M.huakuii       AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATACTTAAGTCAGGGGTGAAATCCCGGG
M.loti          AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATACTTAAGTCAGGGGTGAAATCCCGGG
M.tianshanense  AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATATTTAAGTCAGGGGTGAAATCCCGGG
M.ciceri        AGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATTGTTAAGTTAGGGGTGAAATCCCAGG
Azorhizobium    AGCGTTGCTCGGAATCACTGGGCGTAAAGCGCACGTAGGCGGATCGTTAAGTCAGGGGTGAAAGCCTGGA
24slym          AGCGTTGTTCGGATTTACTGGGCGTAAAGCGCACGTAGGCGGACTTTTAAGTCAGGGGTGAAATCCCAGA
94              AGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTTTGTAAGACGGATGTGAAATCCCCGG
PL19b           AGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTCCGCTAAGACAGATGTG
PL20bONS        AGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCG
85a1t1ONS       AGCGTTACTCGGAATTACTGGGCGTAAAGCGTGCGTAGGTGGTTGTTTAAGTCCGTTGTGAAAGCCCTGG

                         500       510
13csterre       GCTCAACTCCAGAACTG
13b             GCTCAACTCCAGAACTG
46c2            GCTCAACTCCAGAACTGCC
27d             GCTCAACTCCAGAACTG
48a             GCTCANCTCCAGAACT
TK1             GCTCAACTCCAGAACT
47c3a           GCTCAACTCCAGAACTG
B.japonicum     GCTCAACTCCAGAACT
102a            GCTCAACTCCAGA
PL3             GCTCAACT
103b            GCTCAACTCC
48b             GCTCAACTCCAGAACTG
68aONS          GCTCAACTCCAGAACTG
68d             GCTCAACTCCAGAACTG
49b             GCTCAACTCCAGAACTGCCTTTGATA
79c             GCTCAACTCCAGAACTGCCTTTGAT
70a             GCTCAACTCCAGAACTGC
w1c             GCTCAACTCCAGAAC
B.elkanii       GGTCAACTCCAGAACT
PL18b           GCTCAACCCTGGAACTGCC
S.meliloti      GCTCAACCCT
S.fredii        GCTCAACCCCGGAACT
S.saheli        GCTCAACCCCGGAACT
S.xinjiangensis GCTCAACCCCGGAACT
S.teranga       GCTCAACCCCGGAACT
R.gallicum      GCTCAACTCTGGAACT
R.mongolense    GCTCAACTCTGGAACT
PL20a           GCTCAACTCTGGAACT
R.etli          GCTCAACCCTGGAACT
R.leguminosarum GCTCAACCCT
R.tropici       GCTCAACCCTGGAACT
R.galegae       GCTCAACTCTGGAACT
R.huautlense    GCTCAACTCTGGAACT
Allorhizobium   GCTCAACTCTGGAACT
R.giardinii     GCTCAACTCTGGAACT
15c             GCTCAAC
26c             GCTCA
M3b             GCTCAACCC GGAACTGCTT
M3c             GCTCAACCC GGAACTGCCTTT
PL27a           GCTCAACCC GGAACT
PL10grkn         CTCAACCC GGAACTGCCTTTA
M.amorphae      GCTCAACCCCGGAACT
M.plurifarium   GCTCAACCCCGGAACT
M.huakuii       GCTCAACCCCGGAACT
M.loti          GCTCAACCCCGGAACT
M.tianshanense  GCTCAACCCCGGAACT
M.ciceri        NCTCAACCCTGGAACT
Azorhizobium    GCTCAACTCCAGAACT
24slym          GCTCAACTCTGGA
94              GCTTAACCTGGGAACTG
PL19b
PL20bONS
85a1t1ONS       GCTCAACCTGGAATTGC
```

FIGURE 2 CONTINUED

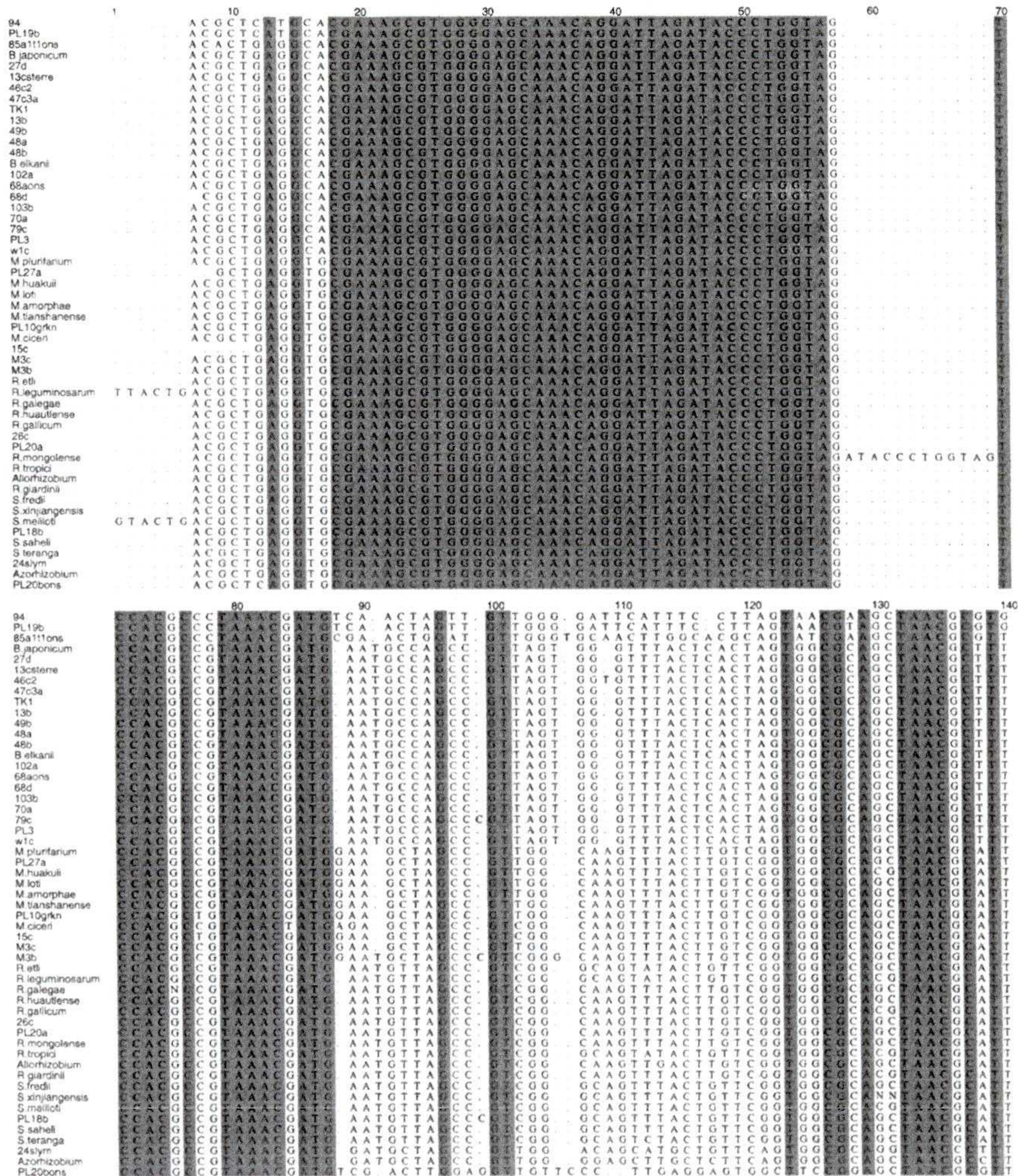

FIGURE 3 ALIGNMENT OF THE SEQUENCES OF THE 450 bp REGIONS OF THE 16S rDNA OF THE ISOLATES AND REFERENCE STRAINS INCLUDED IN THE STUDY. ALIGNMENTS WERE GENERATED BY ClustalX AND DISPLAYED WITH THE Shadybox COMPUTER PROGRAMME. THE REGIONS IN COLOUR ARE 100% HOMOLOGOUS.

FIGURE 3 CONTINUED

FIGURE 3 CONTINUED

FIGURE 3 CONTINUED

Variation in the success of the sequencing reactions can be attributed to several factors: the quality of the sequencing gel, the knowledge and expertise of the operator of the ABI 377 automatic sequencer, the quality of the purified product and correct concentrations of the different reagents of the sequencing reaction.

4.5 PHYLOGENETIC ANALYSIS OF DATA

The sequencing results obtained using primer 16SRNAVI-S were edited to a total length of 440 bp. The region corresponded to positions 131-570 (numbering corresponds to that of *E. coli*). The region included two of the variable regions (V2 and V3). Primary sequences of the primer 16SRNAII-S were edited to a total length of 450 bp, corresponding to positions 691-1140 (numbering corresponds to that of *E. coli*). Three variable regions (V5-V7) were included in the region sequenced by primer 16SRNAII-S. Only part of the variable region V7 was included in the primary sequence. The afore-mentioned primer sequenced more easily than primer 16SRNAVI-S.

The trees reconstructed after the multiple alignment of the sequences with the ClustalX programme are shown in Fig. 4 (tree reconstructed from sequences obtained with primer 16SRNAVI-S) and Fig. 5 (tree reconstructed from sequences obtained with primer 16SRNAII-S). The topologies of the trees differed, but the isolates clustered in the same clusters. The relationships of the isolates within each cluster differed between the two figures. The tree in Fig. 4 was selected to represent the main tree and the numbering of the clusters of Fig. 5 corresponds to that of Fig. 4.

4.6 TREE RECONSTRUCTED FROM PARTIAL SEQUENCE RESULTS OF PRIMER 16SRNAVI-S

Six distinct groups could be distinguished in the tree reconstructed from the data obtained using primer 16SRNAVI-S (Fig. 4). Each group is described separately, referring to the relevant positions of the groups and the isolates within each group.

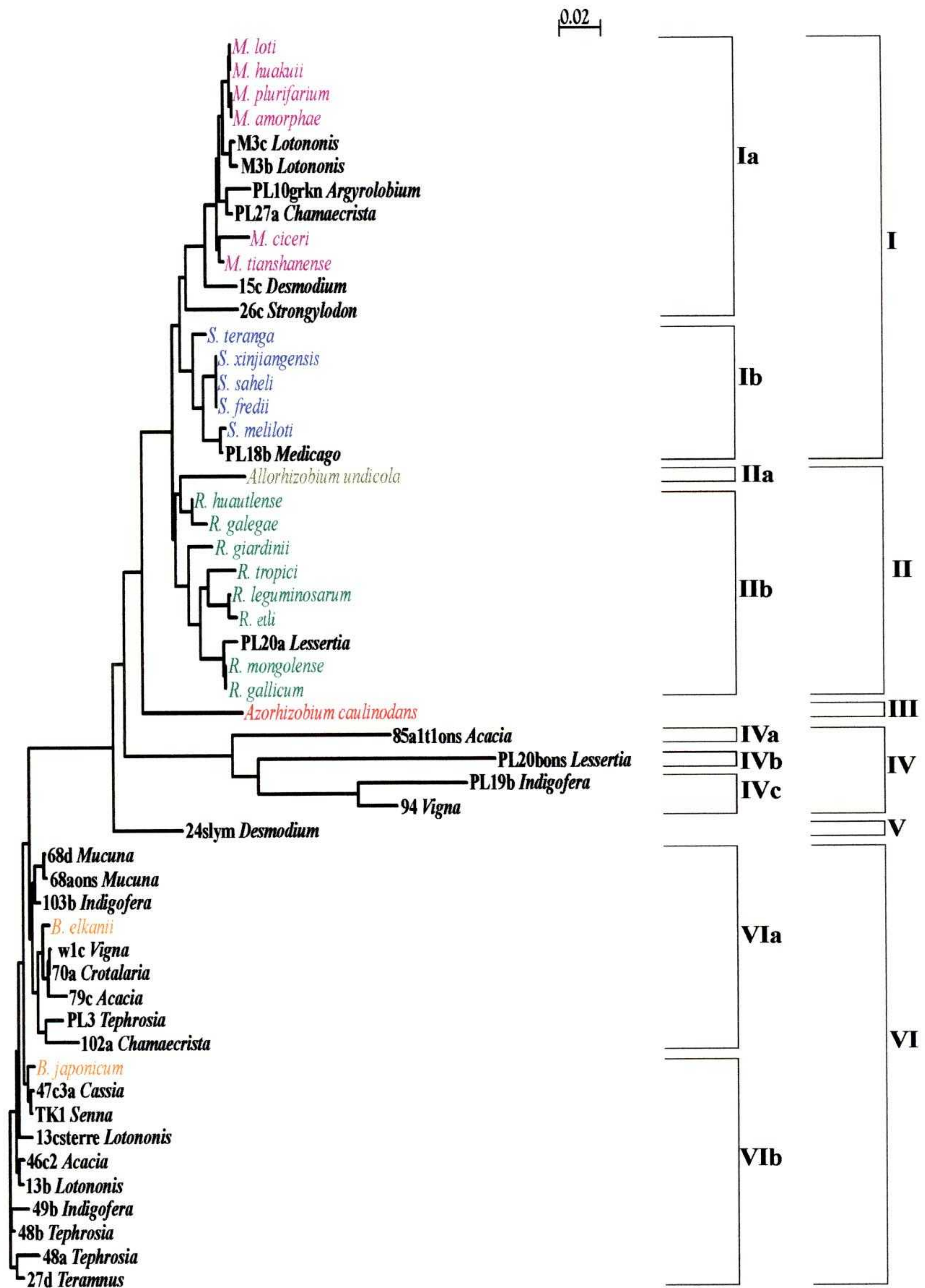

FIGURE 4 **An unrooted tree reconstructed from sequence data obtained from primer 16SRNAVI-S. Branch lengths are proportional to the estimated genetic distance between the strains. The scale represents 2% nucleotide difference. Vertical lengths are not significant and are set for clarity. Abbreviations: *M.*, *Mesorhizobium*; *S.*, *Sinorhizobium*; *R.*, *Rhizobium* and *B.*, *Bradyrhizobium*.**

Group I - *Mesorhizobium* and *Sinorhizobium* group

Group I could be divided into two subgroups, Ia and Ib, corresponding to the genera, *Mesorhizobium* (subgroup Ia) and *Sinorhizobium* (subgroup Ib). Reference strains of the different species within each genus clustered in the appropriate subgroup, as expected.

Subgroup Ia

The reference strains of *M. loti*, *M. huakuii*, *M. plurifarium* and *M. amorphae* clustered in a tight grouping. The two isolates from *Lotononis falcata* M3b and M3c clearly belonged to the genus *Mesorhizobium*, showing sequence similarity of 98,09% (M3b) and 99,09% (M3c) to the cluster of the above-mentioned reference strains. The reference strains *M. ciceri* and *M. tianshanense* formed a separate cluster in the group. The isolates PL10grkn (*Argyrolobium tomentosum*) and PL27a (*Chamaecrista mimosoides*) showed sequence similarities of 98,18% and 99,09% respectively to the *M. loti*, *M. huakuii*, *M. plurifarium* and *M. amorphae* cluster. The isolate from *Argyrolobium tomentosum* (PL10grkn) showed the closest relationship to the reference strain of *M. tianshanense* (98,09%) and the isolate from *Chamaecrista mimosoides* (PL27a) shared 98,36% sequence similarity with *M. tianshanense*. The isolate 15c (*Desmodium tortuosum*) clustered into a separate branch and shared sequence similarities of 97,82% with the previously discussed cluster and 97,64% with the *M. loti*, *M. huakuii*, *M. plurifarium* and *M. amorphae* cluster. The closest reference strain was *M. tianshanense* showing sequence similarity of 97,55% with 15c. The isolate from *Strongylodon macrobotrys* (26c) formed a separate branch in the group. The sequence similarity of isolate 26c with the other members of the group was relatively low and the position of the isolate in the *Mesorhizobium* subgroup was uncertain. *Mesorhizobium tianshanense* was the closest relative of the isolate from *Strongylodon macrobotrys* (26c) in the group, sharing 95,73% sequence similarity.

Subgroup Ib

All the reference strains of *Sinorhizobium* clustered into subgroup Ib (excluding *S. medicae*, which is not available on the GenBank database). Isolate PL18b from *Medicago sativa* displayed 99,45% sequence similarity with *S. meliloti*, 98,27% sequence similarity with the *S. xinjiangensis*, *S. saheli* and *S. fredii* cluster, and 97,82% sequence similarity with *S. teranga*.

Group II - *Allorhizobium* and *Rhizobium* group

The group displayed 99,27% sequence similarity with subgroup Ia and 98,91% with subgroup Ib. The group could be divided into two subgroups: IIa and IIb.

Subgroup IIa

The genus *Allorhizobium* was the only isolate in the subgroup and shared 96,18% sequence similarity with *Rhizobium huautlense*, the closest phylogenetic relative.

Subgroup IIb

All the reference species included in the genus clustered into the subgroup. Only one indigenous isolate from *Lessertia annularis* (PL20a) clustered into the *Rhizobium* group. The closest phylogenetic relatives (99,27% sequence similarity) were *R. mongolense* and *R. gallicum*.

Group III - *Azorhizobium* group

The reference strain *Azorhizobium caulinodans* was the only member of group III. As expected the closest phylogenetic relative was quite distant. The group shared 92,91% with group I, 93,27% sequence similarity with group II, 88,55% with group IV, 90,00% with group V and 89,27% with group VI.

Group IV

No rhizobial reference strains clustered in group IV. The group displayed 92,00% sequence similarity with group II, 91,64% similarity with group I, 88,55% similarity with group III, 90,55% with group V and 89,82% with group VI. The group could be divided into three subgroups: IVa, IVb and IVc.

Subgroup IVa

The isolate from *Acacia xanthophloea* (85a1t1ons) was the only member of this subgroup. It displayed 90,91% sequence similarity with subgroup IVb and IVc.

Subgroup IVb

Only one isolate from *Lessertia annularis* (PL20bons) clustered in this subgroup. The subgroup displayed 88,18% sequence similarity with subgroup IVc and 79,09% similarity with subgroup IVa.

Subgroup IVc

Two isolates, PL19b from *Indigofera verrucosa* and 94 from *Vigna subterranea* clustered in the subgroup. The two isolates displayed a sequence similarity of 92,55%.

Group V

The only member of the group was isolate 24slym from *Desmodium tortuosum*. The closest phylogenetic relative according to the tree was group II, the *Allorhizobium* and *Rhizobium* group, displaying a sequence similarity of 93,45%.

Group VI - *Bradyrhizobium* group

The group could be divided into two subgroups (VIa and VIb), each containing one of the *Bradyrhizobium* reference strains (no 16S rDNA sequence data for the species *B. liaoningense* exist in the GenBank database).

Subgroup VIa

The reference strain *B. elkanii* clustered into the subgroup. Other isolates falling in this group were phylogenetically closely related to the reference strain. The isolates displayed the following sequence similarities with *B. elkanii*: w1c (*Vigna unguiculata*) - 99,27%, 70a (*Crotalaria brachycarpa*) - 99,36%, 79c (*Acacia caffra*) - 98,45%, PL3 (*Tephrosia*

grandiflora) - 98,09%, 102a (*Chamaecrista biensis*) - 97,36%, 68d (*Mucuna coriacea*) - 98,64%, 68aons (*Mucuna coriacea*) - 98,55% and 103b (*Indigofera hilaris*) - 98,73%.

Subgroup VIb

The *Bradyrhizobium japonicum* reference strain clustered into this subgroup. The two isolates displaying the closest phylogenetic relationship to *B. japonicum* were the isolates from *Senna petersiana* (TK1) and *Cassia floribunda* (47c3a). The sequence similarity of both with *B. japonicum* was 99,27%. The remaining isolates in subgroup VIb displayed the following sequence similarities with *B. japonicum*: 13csterre (*Lotononis bainesii*) - 98,45%, 46c2 (*Acacia sieberana* var *woodii*) and 13b (*Lotononis bainesii*) - 98,82%, 49b (*Indigofera melanadenia*) - 97,91%, 48b (*Tephrosia purpurea*) - 98,64%, 48a (*Tephrosia purpurea*) - 97,36% and 27d (*Teramnus labialis*) - 98,09%.

4.7 TREE RECONSTUCTED FROM PARTIAL SEQUENCE RESULTS OF PRIMER 16SRNAII-S

Six groups could be distinguished in the tree reconstructed from the data obtained using primer 16SRNAII-S (Fig. 5). Their numbering corresponds to that of Fig. 4.

Subgroup Ia - *Mesorhizobium* group

Five rhizobia isolates clustered in the *Mesorhizobium* group. The following rhizobia isolates were included: PL10grkn (*Argyrolobium tomentosum*), PL27a (*Chamaecrista mimosoides*), M3c (*Lotononis falcata*), 15c (*Desmodium tortuosum*) and M3b (*Lotononis falcata*). The branching pattern of the isolates differed from that in Fig. 4. PL10grkn displayed a sequence similarity of 99,65% with the reference strains *M. tianshanense* and *M. amorphae*. The closest relative (nearly 100% sequence similarity) of PL27a was *M. plurifarium*. *Mesorhizobium loti* was the closest phylogenetic relative of the isolates M3c (96,29%), 15c (96,18%) and M3b (96,65%).

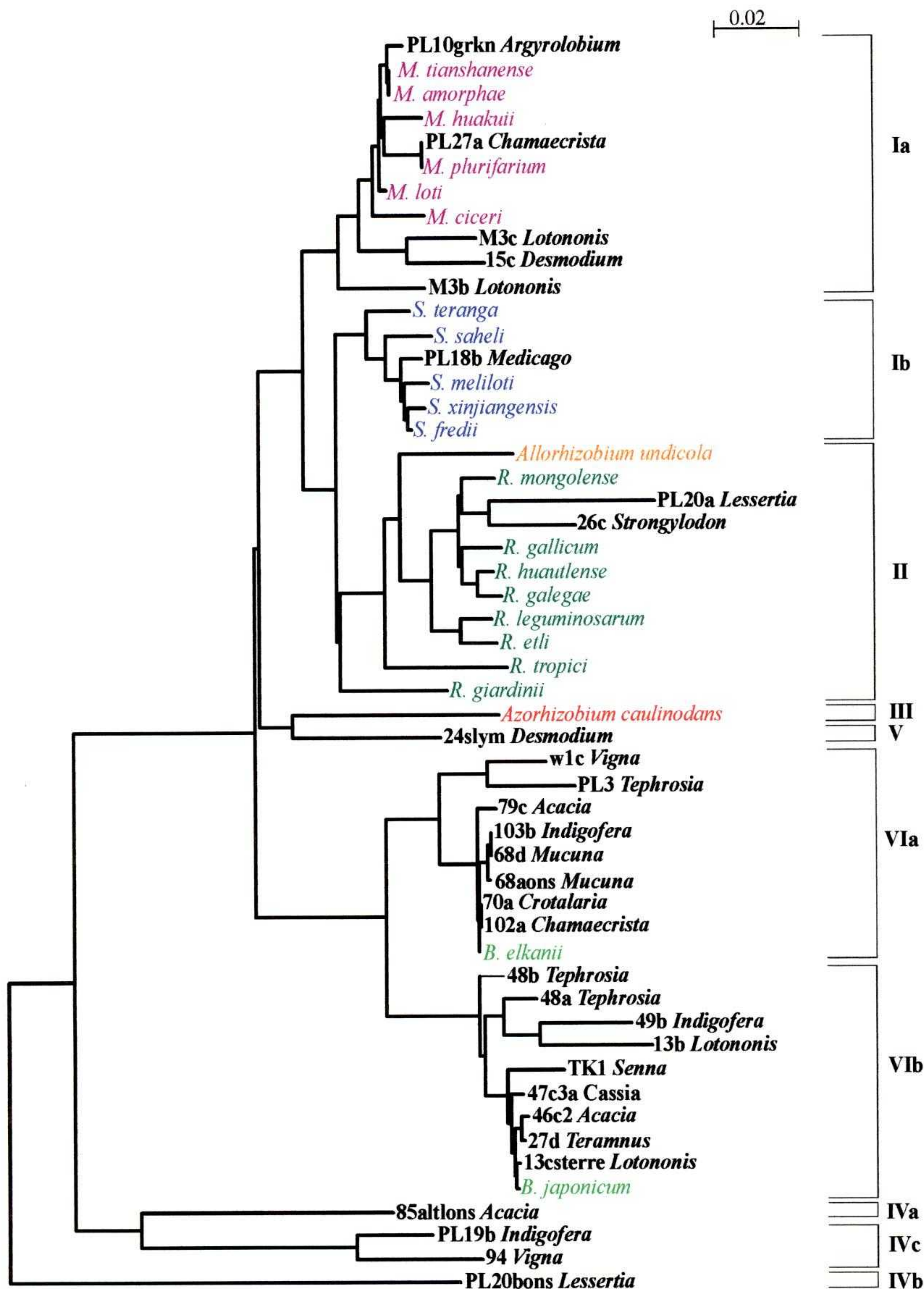

FIGURE 5 **An unrooted tree reconstructed from sequence data obtained from primer 16SRNAII-S. Branch lengths are proportional to the estimated genetic distance between the strains. The scale represents 2% nucleotide difference. Vertical lengths are not significant and are set for clarity. Abbreviations: *M.*, *Mesorhizobium*; *S.*, *Sinorhizobium*; *R.*, *Rhizobium* and *B.*, *Bradyrhizobium*.**

Subgroup Ib - *Sinorhizobium* group

All the species of *Sinorhizobium* clustered in the group together with the isolate PL18b from *Medicago sativa*. The isolate displayed different degrees of similarity with the reference strains: 98,94% (*S. meliloti*), 99,06% (*S. xinjiangensis*), 99,41% (*S. fredii*), 98,00% (*S. saheli*) and 97,77% (*S. teranga*).

Subgroup IIa - *Allorhizobium* group

The reference strain *A. undicola* was the only member of the group, no indigenous isolate clustered in the group.

Subgroup IIb - *Rhizobium* group

All the reference strains clustered in the group, except that the relative positions of the different species differed from Fig. 4, as well as the phylogenetic distance between the species. Two isolates clustered in the group, PL20a (*Lessertia annularis*) and 26c (*Strongylodon macrobotrys*). Isolate PL20a shared sequence similarity of 94,47% with *R. mongolense*, 94,35% with *R. gallicum* and *R. galegae* and 94,59% with *R. huautlense*. The isolate 26c was phylogenetically more related to the reference strains named above than PL20a. Sequence similarities of 96,35% (*R. mongolense*), 96,24% (*R. gallicum* and *R. galegae*) and 96,47% (*R. huautlense*) were displayed by the different reference strains.

Group III - *Azorhizobium* group

Azorhizobium caulinodans was the only member of group III, the closest phylogenetic relative being quite distant. The group displayed the following sequence similarities: 92,18% (subgroup Ia), 91,53% (subgroup Ib), 92,24% (group II), 81,41% (subgroup IVa), 77,18% (subgroup IVb), 82,94% (subgroup IVc), 91,29% (group V) and 90,82% (group VI). The group was the most similar to group II.

Subgroup IVa

The only isolate in the subgroup was from *Acacia xanthophloea* (85a1t1ons), with no reference strain grouping close to the isolate. The isolate was distantly related to subgroup IVb (79,29% sequence similarity) and to subgroup IVc (88,12% sequence similarity).

Subgroup IVb

Isolate PL20bons (*Lessertia annularis*) was the only member of the subgroup, differing significantly from any reference strain.

Subgroup IVc

Two isolates PL19b (*Indigofera verrucosa*) and 94 (*Vigna subterranea*) belonged to this subgroup. The two isolates shared 95,06% sequence similarity.

Group V

The isolate 24slym (*Desmodium tortuosum*) was the only member of this group. All the rhizobial reference strains in other subgroups were quite distant to the isolate displaying the following sequence similarities: 93,59% (subgroup Ia), 92,94% (subgroup Ib), 93,65% (group II), 91,29% (group III), 82,82% (subgroup IVa), 78,59% (subgroup IVb), 84,35% (subgroup IVc) and 92,24% (group VI). The group shared the highest similarity with group II, the *Allorhizobium* and *Rhizobium* cluster.

Subgroup VIa

Eight indigenous rhizobia isolates belonged to the subgroup, which also contained the *B. elkanii* reference strain. The indigenous isolates were w1c (*Vigna unguiculata*), PL3 (*Tephrosia grandiflora*), 79c (*Acacia caffra*), 103b (*Indigofera hilaris*), 68d (*Mucuna coriacea*), 68aons (*Mucuna coriacea*), 70a (*Crotalaria brachycarpa*) and 102a (*Chamaecrista biensis*). Two isolates w1c and PL3 were phylogenetically more distant from *B. elkanii* than the other isolates, sequence differences varying from 3,53% (w1c) to 4,35% (PL3). The isolates from *Crotalaria brachycarpa* (70a) and *Chamaecrista biensis* (102a) displayed nearly

100% sequence similarity with the reference strain. All the other isolates were more distantly related, but still shared more than 99% sequence similarity with the reference strain. The isolates displayed the following sequence similarities with the reference strain: 99,53% (79c), 99,76% (103b and 68d) and 99,65% (68aons).

Subgroup VIb

The reference strain *B. japonicum* clustered in the subgroup, together with nine indigenous rhizobia isolates. The isolates 13csterre (*Lotononis bainesii*), 27d (*Teramnus labialis*), 46c2 (*Acacia sieberana* var *woodii*), 47c3a (*Cassia floribunda*), TK1 (*Senna petersiana*), 13b (*Lotononis bainesii*), 49b (*Indigofera melanadenia*), 48a (*Tephrosia purpurea*) and 48b (*Tephrosia purpurea*) belonged to the subgroup. Phylogenetically, most of the isolates were closely related to *B. japonicum*, displaying the following sequence similarities: 99,76% (13csterre), 99,59% (27d), 99,47% (46c2), 99,53% (47c3a), 98,47% (TK1), 98,00% (48a) and 98,59% (48b). Two of the isolates displayed less than 97% similarity with the reference strain. Isolate 13b from *Lotononis bainesii* shared 96,41% sequence similarity with *B. japonicum* and isolate 49b from *Indigofera melanadenia* shared 95,71% sequence similarity with *B. japonicum*.