CHAPTER 5

Characterisation of the indigenous rhizobia by partial 16S rDNA Sequencing and PCR-Restriction Fragment Length Polymorphism of the 16S-23S rDNA Intergenic Spacer region.

ABSTRACT

A selection of rhizobial isolates associated with legumes indigenous to South Africa was characterised by partial 16S rDNA sequencing and RFLP of the 16S-23S IGS region. Although the nodule isolates were obtained from a wide range of leguminous hosts, 16S rDNA sequence analysis, showed most of them to be related to the genera *Bradyrhizobium* and *Mesorhizobium*. A few isolates were shown to be heterogeneous since no clear affiliation to a specific genus could be established. Previous research (Willems *et al.*, 2001b) has shown that 16S rDNA sequences contain insufficient discriminatory ability to describe the highly diverse bradyrhizobia. IGS-RFLP analysis was therefore performed on isolates related to the genus *Bradyrhizobium*, as well as a group of isolates showing high homology to the *Mesorhizobium* genus. Using this approach it was possible to indicate the presence of additional genotypes, not clearly resolved by the highly conserved 16S rRNA gene.

Keywords: *Rhizobium*, 16S rDNA sequencing, 16S-23S IGS RFLP

INTRODUCTION

Rhizobia are considered bacteria of agricultural and environmental significance since they are able to fix atmospheric nitrogen during their symbiotic association with leguminous plants of the family Fabaceae and the nonleguminous Parasponia. This advantageous association with these plants, many of which are important food crops, has led to numerous investigations into the diversity of these symbionts. These efforts relied heavily on modern molecular biological techniques and have led to the description of the genera *Rhizobium* (Frank, 1889), *Bradyrhizobium* (Jordan, 1982), *Sinorhizobium* (Chen *et al.*, 1988; de Lajudie *et al.*, 1994), *Mesorhizobium* (Jarvis *et al.*, 1997), *Allorhizobium* (de Lajudie *et al.*, 1998a) and *Azorhizobium* (Dreyfus *et al.*, 1988).

The taxonomy of root- and stem nodulating bacteria is directed by a set of minimal standards (Graham *et al.*, 1991), which requires phenotypic, genetic and phylogenetic traits to be considered in the description of these bacteria. These requirements are in line with the polyphasic approach (Colwell, 1970). Within this approach characteristics of various discriminatory powers are targeted to resolve the complex intra-and intergenic relationships of the bacteria (Murray *et al.*, 1990). Polyphasic taxonomy has consequently contributed to an increased stability of the rhizobial taxonomy.

The nodulation status of legumes growing in South Africa remained unknown for a long period of time until Grobbelaar and co-workers embarked on a systematic survey of approximately 1400 legume species (Strijdom, 1998). This survey stretched over two decades and included almost 1000 species growing under diverse geographical and climatological conditions. Unfortunately, the associated bacterial symbionts were not isolated. More recently, Dagutat (1995) and Dagutat & Steyn (1995) reported on the diversity of rhizobia associated with a diverse range of leguminous hosts. Their results showed most of the indigenous strains to be related to the genus *Bradyrhizobium* (Jordan, 1982), while others showed no association or correspondence with the known rhizobial genera. However, this research was limited to the analysis of growth rate characteristics, colony morphology and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) of total soluble cellular proteins and did not comply with the minimal standards as

discussed earlier. Kruger (1998) therefore undertook an expansive study to include more diverse leguminous plants and analyses of the 16S rDNA gene by RFLP using a set of four different tetrameric restriction endonucleases. Following this approach the author was able to elucidate the identity of most of the indigenous rhizobia. According to Moyer *et al.* (1996), using a combination of three or more tetrameric restriction endonucleases, more than 99% of the operational taxonomic units may be identified. However, it is not meant to replace the 16S rDNA sequencing, but rather a means to more efficiently and effectively describe microbial diversity prior to the sequencing step. Furthermore, phylogenetic relationships based on 16S rDNA sequence data show that rhizobia are polyphyletic: some being closer related to non-symbiotic, non-nitrogen fixing genera within the α-*Proteobacteria* than to each other (Young, 1996; van Berkum & Eardly, 1998). Thus, the aim of this work was to determine the partial 16S rDNA sequence of a selection of rhizobia from indigenous legumes in an effort to either confirm or establish their phylogenetic position. Furthermore, the 16S-23S intergenic spacer (IGS) region is reported to be a useful target region to characterise bacterial strains at a higher discriminating power than that sometimes achieved with 16S rDNA sequences, (Laguerre *et al.*, 1996; Selensky-Pobell *et al.*, 1996, Terefework *et al.*, 1998) and therefore the IGS regions of the indigenous strains was also characterised in an effort to show additional diversity present amongst the indigenous rhizobial isolates.

MATERIALS AND METHODS

Bacterial strains

Bacterial strains used in this study are indicated in Table 5.1. These isolates were selected from a rhizobial culture collection established in previous studies by Dagutat (1995) and Kruger (1998). Additional rhizobial strains (Table 5.1), isolated from the root nodules of *Aspalathus linearis* were obtained from the ARC-Plant Protection Research Institute (Private Bag X134, Pretoria 0001, South Africa). Reference cultures of the different rhizobial genera were obtained from the bacterial culture collections of the Laboratorium voor Microbiologie, State University Gent, Belgium and the United States Department of Agriculture (USDA), Soybean and Alfalfa Research Laboratory, Maryland, USA. The 16S rDNA sequences of the rhizobial reference strains and other members of the α-*Proteobacteria* were obtained from

GenBank database of the National Centre for Biotechnology (NCBI) [website address: www.ncbi.nlm.nih.gov/Genbank/]

Maintenance of bacterial cultures

Strains were maintained on yeast extract mannitol (YM) agar, containing 1% (m/v) mannitol, 0.5% (m/v) K_2HPO_4, 0.02% (m/v) $MgSO_4.7H_2O$, 0.01% (m/v) NaCl, 0.04% (m/v) yeast extract and 1.5% (m/v) bacteriological agar. For long term storage cultures were grown (with rigorous shaking) in yeast mannitol broth (YMB) at 28 °C for approximately five days. The turbid culture broth was subsequently mixed at a 1:1 ratio with sterile 50% glycerol in sterile cryotubes and copies of each stored at both -20 °C and -70 °C.

Genomic DNA extraction

Bacterial cultures were grown in YMB for four to seven days and genomic DNA extracted by the hexadecyltrimethyl ammonium bromide (CTAB) method as described by Wilson (1990). Briefly, approximately 1.5 ml cell suspension was centrifuged and the cell pellet resuspended in 567 µl TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8.0) followed by the addition of 30 µl of 10% SDS and 3 µl of 20 mg/ml proteinase K. Sufficient cell lyses was achieved by incubating this mixture overnight at 55 °C. Nucleic acids were separated from cell debris, denatured proteins and polysaccharides by the addition of 100 µl 5 M NaCl and 100 µl CTAB/ NaCl solution (10% CTAB in 0.7 M NaCl). This was followed by incubation at 65 °C for 10 min. The CTAB-protein/polysaccharide complexes were then separated from the nucleic acids using phenol-chloroform extraction as described by Saambrook *et al.* (1989). Nucleic acids were precipitated from the aqueous phase by the addition of 0.6 volume isopropanol. The precipitated DNA was pelleted by centrifugation and excess salt removed by washing with 70% ethanol. The genomic DNA was dried in a vacuum dryer and dissolved in double distilled sterilised water. The integrity and concentration of the purified DNA samples were determined by agarose gel electrophoresis (Saambrook *et al.*, 1989).

Table 5.1 List of indigenous rhizobia analysed in this study.

Subfamilies of the Fabaceae are indicated by **P:** Papilionoideae, **C:** Caesalpinioideae, **M:** Mimosoideae

Isolate	Host legume	Subfamily	Isolate	Host legume	Subfamily
2	*Rhynchosia nervosa*	P	**104a1**	*Chamaecrista sp.*	P
3a	*Trifolium sp.*	P	**114d**	*Pseudarthria hookeri*	P
7b	*Alysicarpus rugosus*	P	**123b**	*Desmodium repandum*	P
13b	*Lotononis bainesii*	P	**125e**	*Indigofera woodii*	P
15b	*Desmodium tortuosum*	P	**128a**	*Melolobium obcordatum*	P
28c	*Desmodium tortuosum*	P	**PL10a**	*Argyrolobium tomentosum*	P
29d	*Alysicarpus rugosus*	P	**UP20b**	*Lotus hispidus*	P
33b	*Crotalaria sp.*	P	**UP26b**	*Crotalaria damrensis*	P
36b	*Neonotonia wightii*	P	**UP27b**	*Crotalaria damrensis*	P
40a	*Sesbania bispinosa*	P	**SA3**	*Trifolium africanum*	P
42bs	*Neonotonia wightii*	P	**xhm5**	*Aspalathus cordata*	P
46c2	*Acacia sieberana var woodii*	M	**xhj7**	*Aspalathus linearis*	P
48a	*Tephrosia purpurea*	P	**xhj8**	*Aspalathus linearis*	P
49b	*Indigofera melanadenia*	P	**xhj12s**	*Aspalathus linearis*	P
60	*Rhynchosia monophylla*	P	**xhj12fr**	*Aspalathus linearis*	P
66a1t1	*Acacia robusta*	M	**xhj13**	*Aspalathus linearis*	P
70a	*Crotalaria brachycarpa*	P	**xhj15**	*Aspalathus linearis*	P
70b2	*Crotalaria brachycarpa*	P	**xhj18**	*Aspalathus linearis*	P
74a	*Indigofera arrecta*	P	**xhj20**	*Aspalathus linearis*	P
82a	*Tephrosia purpurea*	P	**xhj23**	*Aspalathus linearis*	P
98d2	*Bolusanthus speciosus*	P	**xhj26**	*Aspalathus linearis*	P
103b	*Indigofera hilaris*	P	**xhj27**	*Aspalathus linearis*	P

PCR amplification of the 16S rDNA gene and the 16S-23S IGS region

The 16S rRNA gene of each isolate was amplified using universal primers fD1 and rP2 as described by Weisburg *et al.* (1991). However, since no PCR product cloning procedures were anticipated, linker sequences containing the restriction enzyme recognition sites were not included during the synthesis of these oligonucleotides. These primers were therefore designated fD1SHRT and rP2SHRT (Table 5.3). The PCR reaction was carried out in a volume of 100 µl and contained 50 pmol of each primer, 250 µM of each dNTP, 1.5 mM $MgCl_2$, approximately 50 ng genomic DNA and 0.5 U Taq DNA polymerase (Southern Cross Biotechnologies). Additionally, the reaction also contained 50 mM KCl, 10 mM Tris-HCl pH 9.0 and 0.1% Triton X-100 as supplied in the reaction buffer. Amplification was carried out on a Perkin Elmer GeneAmp PCR System 2400 thermocycler using the following thermal

Table 5.2: Primers used for amplification and/ or sequencing of the different genomic regions analysed in this study.

*Primer	Sequence	Target Region	Reference
fD1SHRT	5' AGAGTTTGATCCTGGCTCAG 3'	16 rRNA gene	Weisburg *et al.* (1991)
rP2SHRT	5' ACGGCTACCTTGTTACGACTT 3'	16 rRNA gene	Weisburg *et al.* (1991)
16SRNAII-S	5' GTGTAGCGGTGAAATGCGTAG 3'	16 rRNA gene	Kuhnert *et al.* (1996)
FGPS1490	5' TGCGGCTGGATCACCTCCTT 3'	16S-23S IGS	Laguerre *et al.* (1996)
FGPS132	5' CCGGGTTTCCCCATTCGG 3'	16S-23S IGS	Laguerre *et al.* (1996)

* All primers synthesised by Roche Molecular Diagnostics (Germany)

profile: initial denaturation step at 95 °C for 3 min, 30 cycles of denaturation (94 °C for 1 min), annealing (55 °C for 30 sec), and extension (72 °C for 1 min). An additional extension step (72 °C for 10 min) was performed after completion of the 30 cycles. To determine the size, purity and concentration of the amplified products, aliquots (usually 5 µl) were run on a 1% agarose gel (results not shown). DNA molecular weight marker VI (Roche Molecular Diagnostics, South Africa) was included on each gel.

Since residual reaction components, such as unincorporated dNTPs, primers, etc., can interfere with subsequent DNA sequencing methodologies, 16S rDNA PCR products were purified using the High Pure PCR Product Purification kit (Roche Molecular Diagnostics, South Africa) or Qiagen PCR Purification Kit (Southern Cross Biotechnologies, South Africa) according to the manufacturer's instructions. To assess the purity and concentration of the purified product, 1 µl was subjected to electrophoresis on a 1% agarose gel [Promega] (results not shown).

DNA sequence determination of the 16S rRNA gene

The purified PCR products were sequenced directly, without any additional cloning procedures, using the ABI Prism BigDye^{TM} Terminator Cycle Sequencing Ready Reaction kit (with AmpliTaq^{R} DNA Polymerase, FS) (PE Applied Biosystems). Each sequencing reaction was carried out in a 5 µl volume containing approximately 100 ng template DNA, 12.5 pmol primer, and 2 µl ready reaction premix (supplied with the sequencing kit, containing the dye terminators, dNTP's, AmpliTaq DNA Polymerase, $MgCl_2$, and Tris-HCl

buffer pH 9.0). For each isolate two different reactions were performed, each respectively containing primers 16SRNAII-S, an internal forward primer (Kuhnert *et al.*, 1996), and rP2SHRT (Weisburg *et al.*, 1991). The reactions were carried out in a Perkin Elmer GeneAmp PCR System 2400 thermocycler and consisted of 25 cycles of denaturation (96 °C for 10 sec), annealing (50 °C for 5 sec), and extension (60 °C for 4 min). Products of the sequencing reactions were precipitated with 60% (v/v) ethanol at room temperature for 15 min, washed with 70% (v/v) ethanol, vacuum dried and stored at minus 20 °C until needed. Prior to electrophoresis, the purified products were resuspended in 3.5 µl Blue dextran/EDTA loading buffer (PE Applied Biosystems), 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 reactions were denatured for 2 min at 90 °C and loaded onto the ABI Prism model 377 automated DNA sequencer, which uses a high resolution polyacrylamide gel.

Phylogenetic analysis of the 16S rDNA sequences

The sequencing gels were analysed with the ABI Prism Sequencing Analysis 3.1 and the ABI Prism Sequencing Navigator 1.0.1 computer programmes (PE Applied Biosystems). For each isolate, the sequence of both strands was determined using the rP2SHRT and 16SRNAII-S primer pair. These sequences were aligned, producing a layout from which ambiguous sites could be resolved. The ClustalX programme (Thompson *et al.*, 1997) was used to analyse the nucleic acid sequences. A distance matrix was constructed by pairwise alignment of the sequences. The phylogenetic tree was constructed from the distance matrix using the neighbour-joining method of Saitou & Nei (1987). All branch lengths were proportional to the estimated divergence along each branch. The bootstrap method (Felsenstein, 1985) was used in combination with the neighbour-joining method to estimate confidence levels of the phylogenies. The phylogenetic trees were displayed using NJplot (Perrière & Gouy, 1996).

16S-23S rDNA IGS RFLP analysis

The 16S-23S rDNA IGS region was amplified using the primer set FGPS1490 and FGPL132 (Laguerre *et al.*, 1996) and using the same set of conditions as those described for the amplification of the 16S rRNA gene. The PCR products were used directly in a restriction

reaction, without any prior purification. The digestion reactions were performed using 10 U of each of the following restriction enzymes *Hae*III, *Cfo*I, *Taq*I and *Nde*II (Roche Molecular Diagnostics, South Africa) in optimal buffers as prescribed by the supplier. Electrophoresis of the restricted DNA was carried out on a 3% agarose gel in a Hybaid Maxi Gel System at 80 V for 90 min. DNA molecular weight marker VI (Roche Molecular Diagnostics, South Africa) was included at multiple positions on gels as a standard. Restriction patterns were identified visually for each of the restriction enzymes used. A matrix, describing the presence (indicated by 1) or absence (indicated by zero), of a particular pattern within an isolate was generated. In this way a specific profile for each isolate was compiled. The Dice coefficient (Nei & Li, 1979) within the Bionum computer programme (Applied Maths, Kortrijk, Belgium) was used to construct a distance matrix from this one (1)-zero (0) profiles. These distance values were subsequently analysed to generate a tree using the unweighted pair group method of arithmetic mean algorithm (UPGMA) in GelCompar 4.0 (Applied Maths, Kortrijk, Belgium).

RESULTS

16S rDNA amplification and sequence determination

PCR amplification of the 16S rRNA gene using primers fD1SHRT and rP2SHRT resulted in the detection of an amplified fragment of about 1500 bp for all rhizobial isolates tested. This corresponds to the expected size of the 16S rRNA genes as previously determined for the *rrn* operon of *E. coli* (Brosius *et al.*, 1978).

A partial sequence (about 700 bp) of the 16S rRNA gene was determined using the primer set 16SRNAII-S (an internal forward primer) and rP2SHRT. These sequences correspond approximately to positions 750 and 1450 (numbering according to the *E. coli* 16S rDNA gene sequence). Sequence data, especially at the extremities of the fragment, usually contained ambiguous site, which the automated sequencer was unable to resolve and therefore DNA sequences from both strands was particularly useful to accurately determine the relevant nucleotide residue at such sites. The sequences were submitted to GenBank and their respective accession numbers are shown in Figure 5.1.

Phylogenetic position of the indigenous rhizobia as indicated by partial 16S rDNA sequence analyses

Phylogenetic relationships of the indigenous rhizobia were investigated by comparative sequence analyses of a 700 bp fragment of the 16S rRNA gene. Nucleotide sequences of the rhizobial reference strains were obtained from the GenBank database, suitably edited to correspond to this 700 bp region and included for comparative purposes. A phylogenetic tree (Figure 5.1) was constructed by the neighbour-joining algorithm (Saitou & Nei, 1987) using the ClustalX computer programme.

The phylogenetic tree was composed of two main lineages: the first containing the genera *Mesorhizobium*, *Sinorhizobium*, *Rhizobium*, *Agrobacterium* and *Allorhizobium*, while the second indicated the phylogenetic position of the slow-growing *Bradyrhizobium*. *Azorhizobium undicola* had an intermediate position between these two lineages. The indigenous rhizobia were related to all of the rhizobial genera except *Allorhizobium* and *Azorhizobium*. Two isolates, PL10a and 36b, showed no clear generic affiliation.

Isolates obtained from the root nodules of *Aspalathus linearis,* a shrublike legume, part of the fynbos ecosystem in South Africa, showed high homology to members of the genus *Mesorhizobium*. Isolates xhj26, xhj20, xhj13, xhj23, xhj18, xhj7, xhj15, xhj8, xhj27 (Cluster A) shared almost identical sequence with *M. loti*. Isolate xhm5, obtained from *Aspalathus cordata*, was also closely related to *M. loti*. The other *A. linearis* isolate, xhj12fr, was clearly separated from the "mesoloti-group", sharing 97% sequence homology with *M. ciceri*. Isolates 98d2 and 128a were the only other symbionts showing relatedness to the genus *Mesorhizobium*, showing sequence correspondence to *M. plurifarium* and *M. chacoense,* respectively.

Species of the *Sinorhizobium* genus formed a phylogenetic coherent group, with two of the indigenous isolates, xhj12s and UP27b, related to this group. Both xhj12fr and xhj12s were obtained from the same nodule, but were shown here to belong to different genera.

The phylogenetic branch carrying the genera *Rhizobium, Allorhizobium* and *Agrobacterium* contained the isolates UP26b, 66a1t1, SA3, 28c and 7b. Isolates UP26b and 66a1t1 were

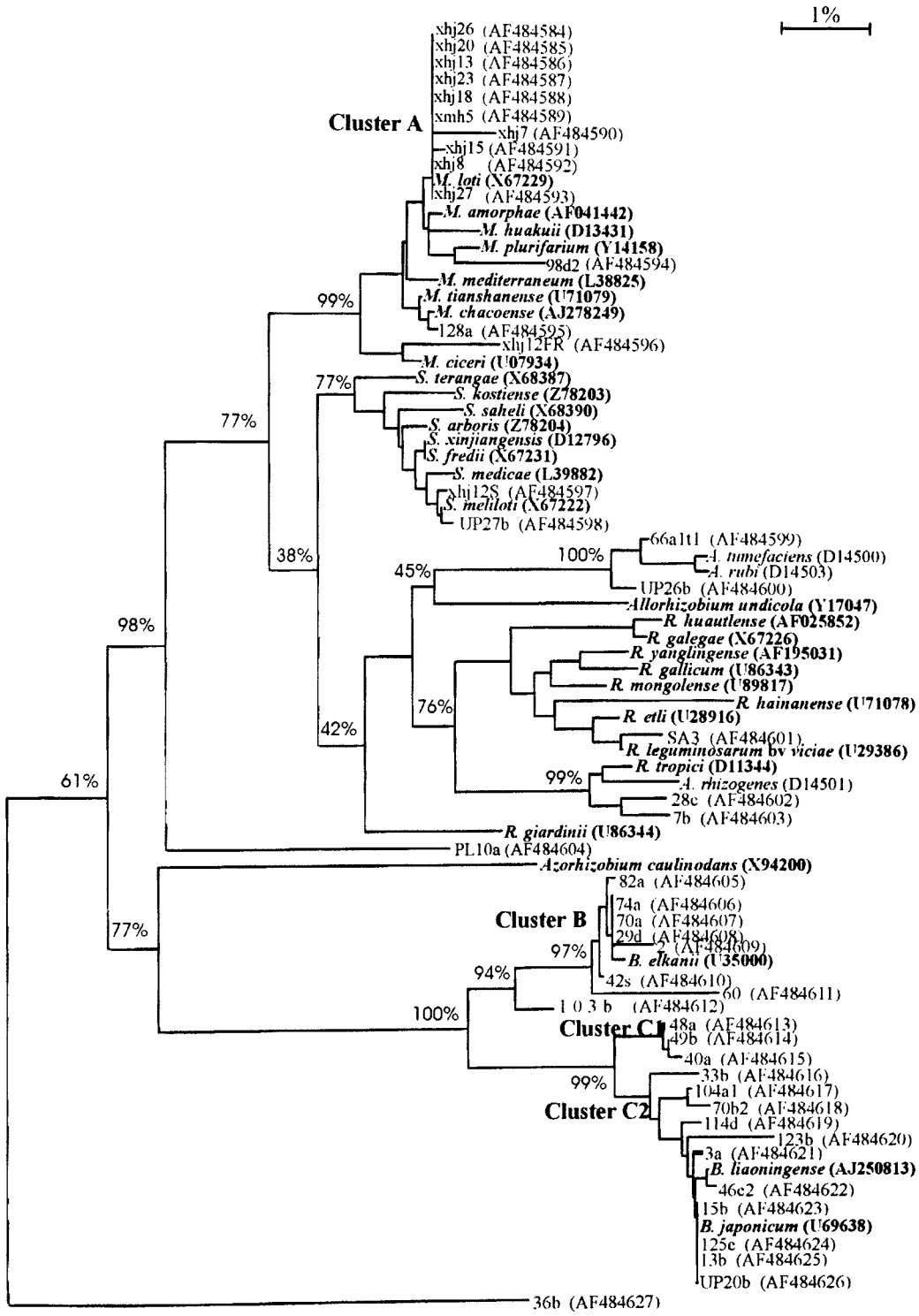


Figure 5.1 (Previous *page*) Phylogenetic relationships of the indigenous rhizobial isolates and other rhizobial reference strains. This analysis was based on comparative sequence analysis of approximately 700 bp fragment of the 16S rRNA gene. The tree was constructed using the Neighbour-joining method of Saitou & Nei (1987). Horisontal branch lengths reflect the phylogenetic distances, while vertical branch lengths are non-informative and set for clarity only. The scale bar indicates 1% nucleotide difference and bootstrap values of some of the major branching points are shown. GenBank accession numbers are indicated in brackets. *M: Mesorhizobium, S: Sinorhizobium, R: Rhizobium, A: Agrobacterium & B: Bradyrhizobium*.

related to *Agrobacterium tumefaciens* and *Agrobacterium rubi*, while 28c and 7b were present in the group containing *R. tropici* and *Agrobacterium rhizogenes*. In this analysis *Rhizobium giardinii* was also recovered as lineage, clearly separable from the one carrying *Rhizobium*, *Agrobacterium* and *Allorhizobium* reference strains.

The second main phylogenetic lineage consisted of the bradyrhizobia. Within this lineage two subgroups (Clusters B & C) were evident and most of the indigenous isolates were distributed between these subgroups. The first subgroup (B) contained *Bradyrhizobium elkanii* and isolates 42bs, 82a, 74a, 70a, 2, and 29d. Also related to this subgroup was isolates 103b and 60, which appeared to be part of a separate group that diverged prior to the branching of *B. elkanii*. Bootstrap values of 94% and 97%, respectively, supported this separate branching of 103b and 60. The second subgroup contained the closely related *B. japonicum* and *B. liaoningense* reference strains together with 48a, 49b, 40a, 33b, 104a1, 70b2, 114d, 123b, 3a, 46c2, 15b, 125e, 13b and UP20b. The branches within this group were poorly resolved which is due to the small sequence divergence (no more than 1.5%) among isolates within this group.

Phylogenetic position of the indigenous rhizobia within the α-*Proteobacteria*

The genera of the family *Rhizobiaceae* do not form a cluster distinctly separated from other genera in the α-*Proteobacteria* (Young, 1996; van Berkum & Eardly, 1998). This means that the rhizobia are closer related to other non-symbiotic bacteria than they are to each other. The 16S rDNA sequence analysis was therefore extended (Fig. 5.2) to include representative members of the α-*Proteobacteria* in an effort to establish an improved understanding of the phylogenetic position of the indigenous isolates.

The *Rhizobium-Agrobacterium-Allorhizobium* (RAA) cluster was related to *Brucella neotomae*, *Ochrobactrum anthropi* and *Bartonella hensellae*, while the *Phyllobacterium* species showed high relatedness to the *Mesorhizobium*, *Sinorhizobium* and RAA cluster. According to several authors (Young & Haukka, 1996; de Lajudie *et al.*, 1998a,b) strains representing the genera *Bartonella*, *Brucella*, *Phyllobacterium* are sometimes interspersed between members of the RAA-cluster. Additionally, the chosen algorithm and selection of

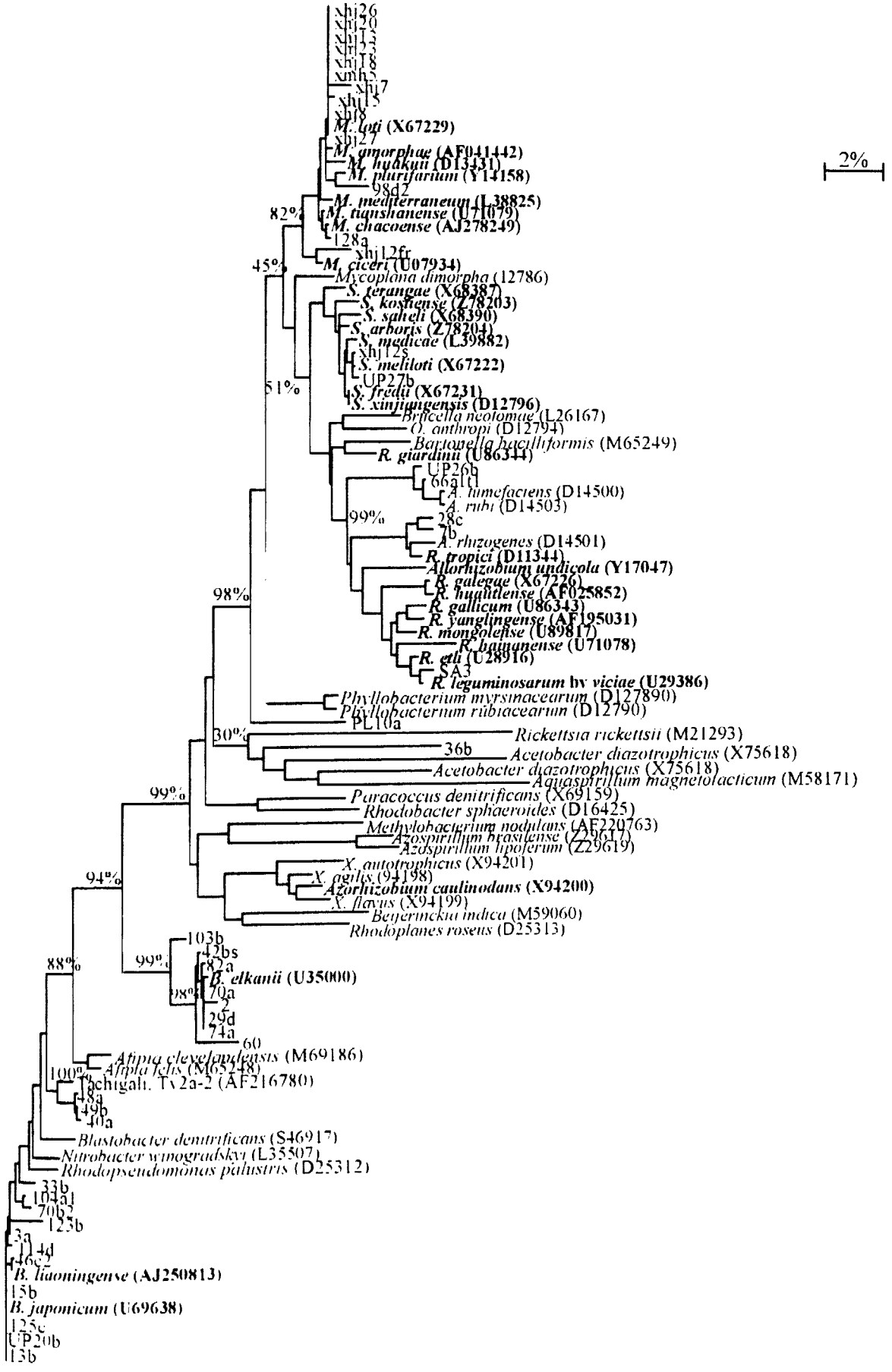


Figure 5.2. (Previous page) Phylogenetic relationships of the indigenous rhizobial isolates and other alpha-*Proteobacteria*. This analysis was based on comparative sequence analysis of approximately 700 bp fragment of the 16S rRNA gene. The tree was constructed using the Neighbour-joining method of Saitou & Nei (1987). Horisontal branch lengths reflect the phylogenetic distances, while vertical branch lengths are non-informative and set for clarity only. The scale bar indicates 2% nucleotide difference and bootstrap values of some of the major branching points are shown. GenBank accession numbers of the different reference strains are indicated in brackets, while those of the indigenous isolates are indicated in Fig 5.1. *Methylobacterium nodulans* was included, although it is not yet validly accepted as a new species. *M: Mesorhizobium, S: Sinorhizobium, O: Ochrobactrum, R: Rhizobium, A: Agrobacterium, X: Xanthobacter, B: Bradyrhizobium.*

included sequences may also influence the phylogenetic groupings (Amarger *et al.*, 1997; Chen *et al.*, 1997; Jarvis *et al.*, 1997).

The genus *Bradyrhizobium* showed a high relatedness to species of the genera *Afipia, Blastobacter, Nitrobacter, Rhodopseudomonas.* The indigenous isolates, previously resolved into the *Bradyrhizobium japonicum-Bradyrhizobium liaoningense* subcluster was again recovered within this group, except isolates 48a, 49b and 40a. These isolates shared high sequence identity with the divergent *Bradyrhizobium* symbiont, Tv2a-2, as described by Parker (2000). Tv2a-2 was isolated from the root nodule of a tree, *Tachigali versicolor* of the Caesalpinioideae subfamily. Additionally, these isolates shared a recent common ancestor with *Blastobacter denitrificans.*

No clear generic affiliation could be established for PL10a, while 36b was associated with the *Rhodospirillum*, *Acetobacter* and *Aquaspirillum* subgroup.

Analyses of combined 16S-23S IGS RFLP analyses of closely related 16S phylogenetic clusters.

The IGS-PCR products of different rhizobial genera are indicated in Figure 5.3. In general the length of the intergenic region of the different genera corresponded to approximately 1000 bp. A significant smaller product (~500 bp) was obtained for members of the genus *Agrobacterium*, while the IGS for *Mesorhizobium* was in the region of 1500 bp. Due to the significant variation in length and number of amplified products, only rhizobial reference strains, related to the observed clusters A, B, C1and C2 (see Figure 5.2), were included in the analyses. Each cluster was also analysed separately and dendrograms, describing the heterogeneity within each cluster, were generated using the UPGMA algorithm with the GelCompar software. The observed restriction fragment patterns for each isolate are shown in Table 5.3. The IGS-RFLP profiles of isolates not present in closely related clusters, such as 128a, 98d2, xhj12fr, UP26b, 66a1t1, SA3, 28c, 7b, xhj12S, UP27b, 36b and PL10a were not analysed. Isolates 60 and UP20b consistently gave amplified products outside the size range obtained with the other isolates and were also not included.

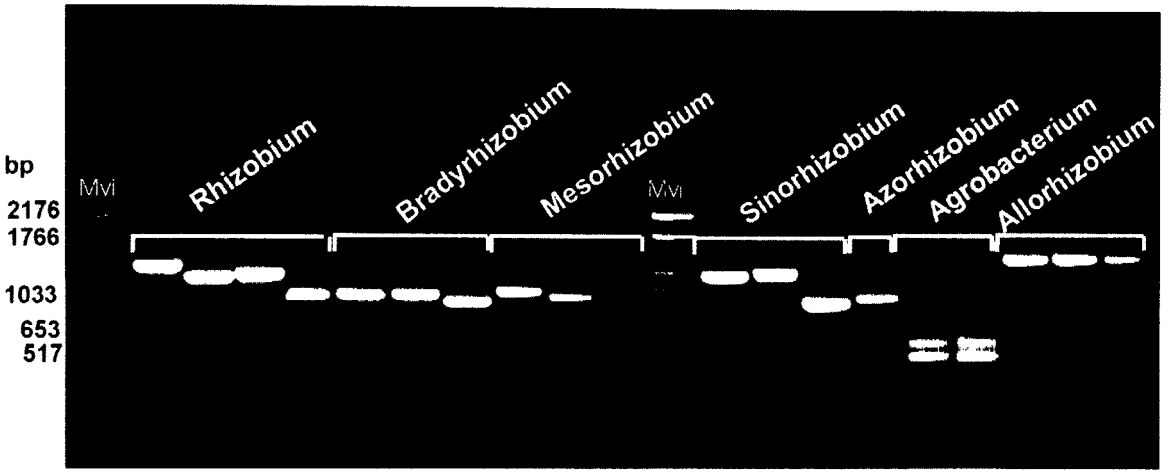


Figure 5.3. Agarose gel electrophoresis of 16S-23S IGS PCR products of representatives of the different rhizobial genera. Lane1 & 12: DNA molecular weight marker VI (Roche Molecular Biochemicals (SA). Lane2: *R. mongolense* (LMG 19141), Lane 3: *R. hainanense* (USDA 3588), Lane 4: *R. tropici* (LMG 9503), Lane 5: *R. leguminosarum* bv. *trifoli* (LMG 8820), Lane 6: *B. elkanii* (LMG 6134), Lane 7: *B. japonicum* (LMG 6138), Lane 8: *Bradyrhizobium* sp. (LMG 8319), Lane 9: *M. ciceri* (LMG 14989), Lane 10: *M. huakuii* (14107), Lane 11: *M. tianshanense* (LMG 18976), Lane 13: *S. saheli* (LMG 7837), Lane 14: *S. fredii* (LMG 6217), Lane 15: *S. terangae* (LMG 7834), Lane 16: *Azorhizobium caulinodans* (USDA 4892), Lane 17: *Agrobacterium rhizogenes* (LMG 150), Lane 18: *Agrobacterium radiobacter* (LMG 140), Lane 19: *Allorhizobium undicola* (USDA 4902), Lane 20: *Allorhizobium undicola* (USDA 4903), Lane 21: *Allorhizobium undicola* (USDA 4904).

An amplified product of about 970 bp was observed for most of the indigenous rhizobia related to clusters B, C1 and C2. In some instances a smaller, 500 bp product was observed. These bands were, however, less prominent in comparison to the 970 bp band. For the isolates present in cluster A, an amplified product of 1000 bp was obtained, as well as an additional less prominent band of about 650 bp.

Cluster A

This cluster comprised most of the isolates obtained from *Aspalathus linearis* and one from *Aspalathus cordata* (xhm5). The four restriction enzymes had equal discriminatory ability within this group, each generating, on average, 5 IGS RFLP types. Only reference strains *M. huakuii*, *M. plurifarium* and *M. loti* grouped close to the *Aspalathus* isolates. Isolates xhj7 and xhj8 (subgroup 1A) had identical *Hae*III, *Cfo*I, and *Taq*I profiles. However, *Nde*II digestion showed that the isolates had different genotypes. Such difference could not be resolved by 16S rDNA sequencing since the observed difference was only 1%. The *Hae*III

Table 5.3 : Restriction patterns obtained of *Mesorhizobium*, *Bradyrhizobium* spp. and indigenous rhizobial strains after digestion of IGS PCR products with the indicated enzymes. Clusters refer to those obtained in 16S rDNA phylogenetic tree (See Fig 5.1). NS indicates the absence of a restriction site. Numerical and alphabetical designations are used to differentiate between the different RFLP patterns. Alphabetical subscript annotations (a, b, c1 and c2) are used to indicate the RFLP type of a particular cluster. Distinctive patterns obtained from rhizobial reference strains are indicated by the following subscripts: *loti: M. loti; h: M. huakuii; tia: M. tianshanense, am: M. amorphae; plu: M. plurifarium; c: M. ciceri; med: M. mediterraneum, ml: M. mediterraneum & M. loti; h/am: M. huakuii & M. amorphae, Be: B. elkanii; Bjap: B. japonicum, Bliaon: B. liaoningense.*

Strain	Patterns obtained following restriction of 16S-23S IGS PCR products with:			
	*Hae*III	*Cfo*I	*Nde*II	*Taq*I
Cluster A				
xhj7	$H1_a$	$C1_a$	$N1_a$	$T1_a$
xhj8	$H1_a$	$C1_a$	N_h	$T1_a$
xhj26	$H4_a$	$C1_a$	N_h	$T1_a$
M. huakuii (LMG 14107)	H_h	$C_{h/am}$	N_h	T_h
xhj15	$H3_a$	C_{loti}	$N4_a$	$T3_a$
xhj18	$H3_a$	C_{loti}	$N4_a$	$T3_a$
xhj23	$H3_a$	C_{loti}	$N4_a$	$T3_a$
xhj20	H_{loti}	$C_{h/am}$	$N6_a$	$T4_a$
xhj27	$H5_a$	$C_{h/am}$	N_{tia}	$T5_a$
M.plurifarium (LMG 11892)	H_{plu}	C_{plu}	N_{plu}	T_{plu}
xhm5	H_{loti}	$C_{h/am}$	$N7_a$	$T4_a$
M. loti (LMG 6125)	H_{loti}	C_{loti}	N_{loti}	T_{loti}
M. ciceri (LMG 14989)	H_c	C_c	N_c	T_c
M. mediterraneum (LMG 17148)	H_{med}	C_{med}	N_{med}	T_{med}
M. tianshanense (LMG 18976)	H_{tia}	C_{tia}	N_{tia}	T_{tia}
M. amorphae (LMG 18977)	H_{am}	$C_{h/am}$	N_{am}	T_{am}
xhj13	H10(NS)	$C_{h/am}$	N_{tia}	T2 (NS)
Cluster B				
103b	$H1_b$	$C1_b$	$N1_b$	$T1_b$
42bs	$H2_b$	$C2_b$	$N2_b$	$T2_b$
82a	H_{Be}	C_{Bliaon}	N_{Be}	$T3_b$
70a	H_{Be}	C_{Bliaon}	N_{Be}	$T3_b$
29d	H_{Be}	C_{Bliaon}	N_{Be}	$T3_b$
2	H_{Be}	C_{Bliaon}	N_{Be}	$T4_b$
Cluster C1				
49b	$H1_{c1}$	$C1_{c1}$	$N1_{c1}$	$T1_{c1}$
48a	$H1_{c1}$	$C1_{c1}$	$N1_{c1}$	$T2_{c1}$
40a	$H1_{c1}$	$C1_{c1}$	$N1_{c1}$	$T2_{c1}$
Cluster C2				
104a1	$H1_{c2}$	$C1_{c2}$	$N1_{c2}$	$T1_{c2}$
70b2	$H1_{c2}$	$C1_{c2}$	$N1_{c2}$	$T1_{c2}$
123b	$H2_{c2}$	$C2_{c2}$	N_{Bjap}	$T3_{c2}$
46c2	$H2_{c2}$	$C2_{c2}$	$N2_{c2}$	$T4_{c2}$
13b	$H3_{c2}$	$C4_{c2}$	N_{Bjap}	$T2_{c2}$
15b	$H3_{c2}$	$C4_{c2}$	N_{Bjap}	$T2_{c2}$
3a	$H3_{c2}$	$C4_{c2}$	N_{Bjap}	$T2_{c2}$
114d	$H2_{c2}$	$C3_{c2}$	N_{Bliaon}	$T5_{c2}$
33b	$H5_{c2}$	$C3_{c2}$	N_{Bliaon}	$T6_{c2}$
125e	$H4_{c2}$	C_{Bjap}	N_{Bliaon}	$T1_{c2}$

and *Nde*II patterns also separated xhj26 from the closely related xhj7 and xhj8. *M. huakuii* is related to this cluster due to its common *Nde*II profile with xhj8 and xhj26. Isolates xhj15, xhj18 and xhj23 (subgroup 1B) had identical enzyme profiles, with their *cfo*I profile being similar to that of *M. loti*. Although xhj20 and xhj27 were also present in this subgroup, they were clearly divergent from the other members of this subgroup and each other. The *Cfo*I profile of *M. huakuii* and *M. amorphae* was the only common RFLP pattern between these isolates. The clustering of xhm5 and *M. loti* (subgroup 1C) is due to a common *Hae*III profile, while the position of xhj13 (in group 2), almost as an "outgroup" is owed to the absence of *Hae*III and *Taq*I. *M. ciceri, M. mediterraneum, M. tianshanense, M. amorphae* were also present in group 2, showing no relatedness to any of the symbionts of *Aspalathus*. This in contrast to the results obtained by partial 16S rDNA sequencing, since no more than 2% sequence divergence was observed among the *Mesorhizobium* reference strains and the *Aspalathus* isolates

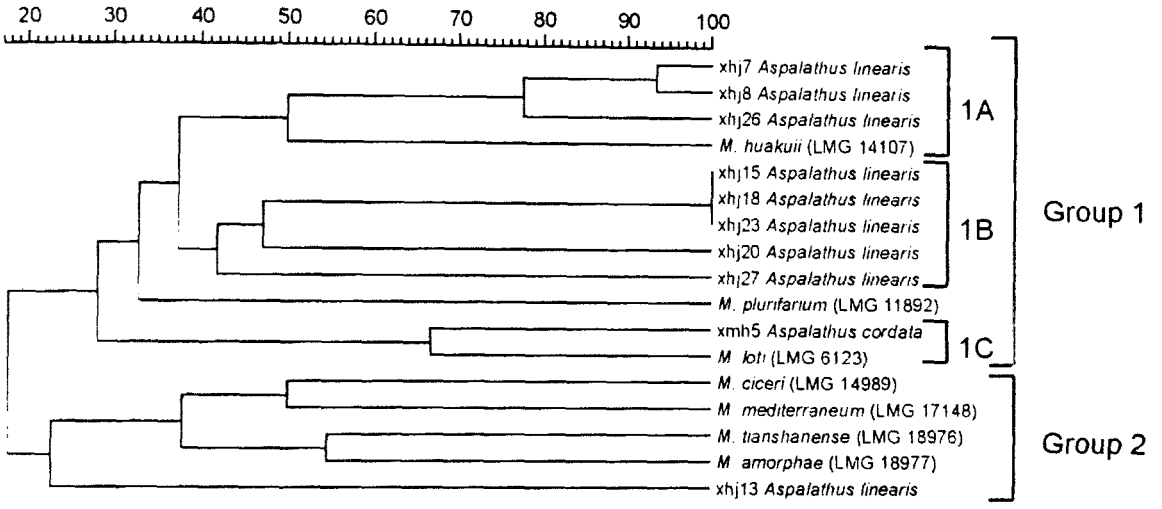
Cluster B

This cluster comprised isolates (82a, 70a, 29d and 2) showing high homology to *B. elkanii*. Isolate 2 differed only from the rest of the group due to a distinct *Taq*I pattern, while the *Hae*III and *Nde*II profiles were that of *B. elkanii*. It is however, noteworthy that these isolates had a characteristic *B. liaoningense-Cfo*I profile, while results of the 16S rRNA gene indicated that these isolates were related to *B. elkanii*. No correspondence was noted between the patterns of isolates 103b and 42bs among themselves or any of the bradyrhizobial reference strains.

Cluster C1

The different IGS patterns of isolates within this cluster do not correspond to any of the bradyrhizobial patterns. However, *B. liaoningense* appears to be the closest neighbour to these isolates. A closer examination of all the restriction types within this group indicates that restriction of the IGS-PCR products yielded on average four fragments of which two

Cluster A



Cluster B

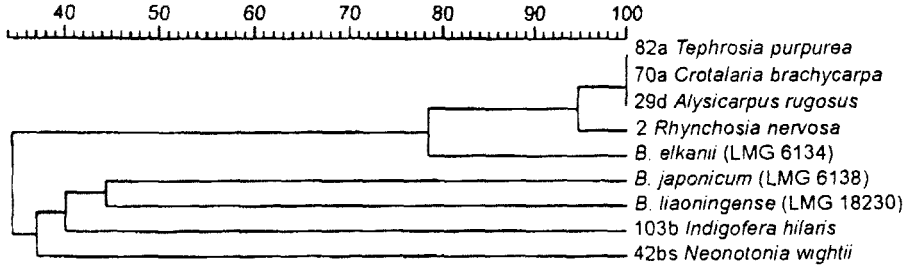
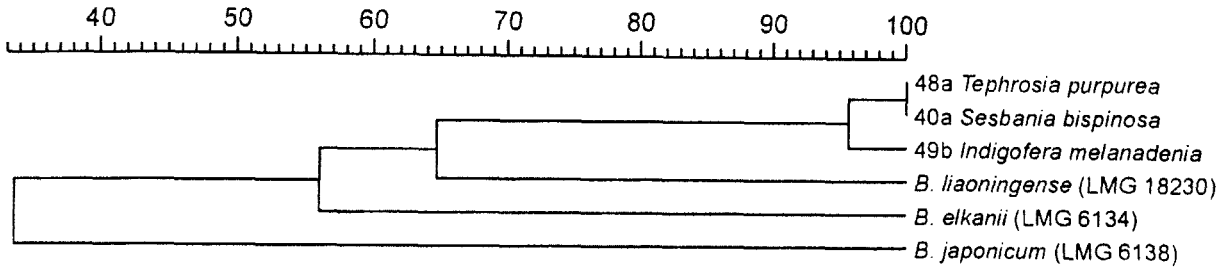


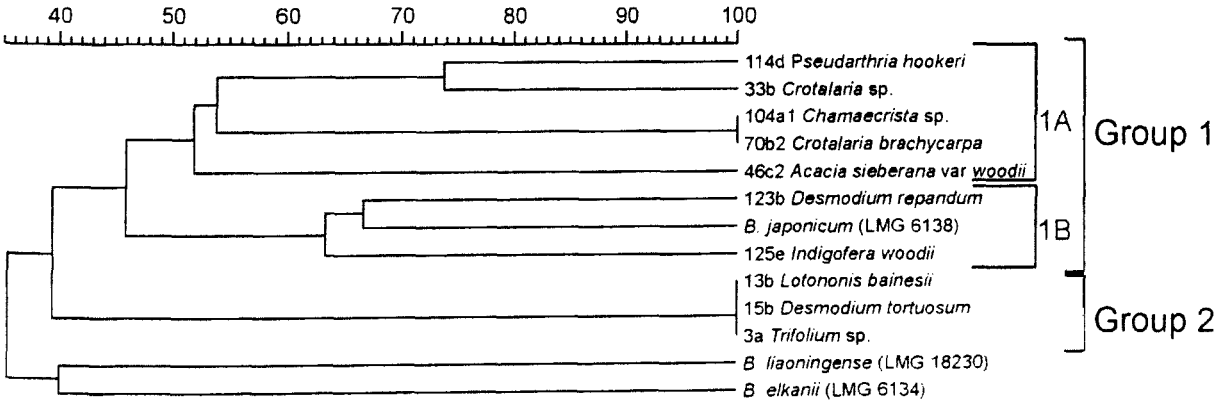
Figure 5.4: Dendrograms based on the combined *Hae*III, *Cfo*I, *Nde*II and *Taq*I restrictions of amplified 16S-23S IGS PCR products. The dendrograms of clusters (A, B, C1 & C2) are shown separately as described in text. Clusters were identified by 16S rDNA sequencing. Dendrograms were constructed using UPGMA. The *x*-axes represent the correlation between the isolates.

Figure 5.4. *continues*

Cluster C1



Cluster C2



were common in *B. liaoningense* patterns. A distinct *Taq*I profile for 49b is responsible for its separate placement within this group.

Cluster C2

Two main groups were identified within this cluster; group 1was related to *B. japonicum* and group 2 showing no clear generic affiliation. The genomic diversity within this group is evident from the five *Hae*III and six *Taq*I patterns, none of which corresponded to any bradyrhizobial reference pattern. Equally, of the five different *Cfo*I patterns, only one corresponded to the *B. japonicum* pattern, while one of each *B. japonicum* and *B. liaoningense* patterns were part of the *Nde*II profile for this cluster. Isolates 114d and 33b share common *Cfo*I and *Nde*II profiles, while isolate 46c2 only shares a common *Hae*III pattern with members of this subgroup. Although, there appears to be little correspondence between IGS-RFLP patterns of members of the subgroup 1A, individual common bands are present in some of the *Cfo*I, *Nde*II and *Taq*I patterns. Isolates in subgroup 1B, have a common *B. japonicum-Nde*II pattern and similar, yet unique patterns for the other enzymes.

DISCUSSION

A broad range of rhizobial species have been isolated from different leguminous species occurring in South Africa. These included mostly isolates from nodules of legumes of the subfamilies Papilionoideae and Mimosoideae, while only isolate, 104a1, was obtained from the root nodules of *Chamaechrista*, which belongs to the Caesalpinioideae subfamily. Except for *Azorhizobium* and *Allorhizobium*, the indigenous strains were related to all of the described genera. Most of these were however, related to the *Bradyrhizobium* genus. Previous morphological and preliminary phylogenetic investigations (Dagutat, 1995; Kruger, 1998) indicated the diversity of the indigenous rhizobial strains, with some isolates showing no apparent affiliation to the known rhizobial genera. Since no single method can produce a perfect classification, the polyphasic approach (Vandamme *et al.*, 1991) to bacterial taxonomy has been emphasized and used for the description of both generic and specific taxa. The aim of this study was therefore to analyse this apparent diversity of the indigenous strains, targeting both conserved (16S rRNA) and variable (16S-23S IGS) genomic regions.

Sequence analyses of the 16S rRNA gene remain one of the most reliable indicators of organismal phylogeny (Woese, 2000) allowing the rapid identification of a large number of strains. The 16S rRNA phylogeny of rhizobia has indicated them to be heterogeneous since some genera are closer related to other non-rhizobial genera of the α-*Proteobacteria* than to each other (Willems and Collins, 1993). Thus, the absence of such genera in any analyses could obscure the true phylogenetic identity of any putative rhizobial symbiont isolated from the indigenous legumes. Our comparative 16S rDNA sequence analyses were therefore subsequently extended to include the other genera of the α-*Proteobacteria*.

On the basis of 16S rDNA sequence analyses, the indigenous strains had two main generic affiliations: *Mesorhizobium* and *Bradyrhizobium*. Thirteen isolates, most of which were obtained from *Aspalathus* spp., were related to the *Mesorhizobium* genus, showing almost 98% 16S rDNA sequence identity among themselves and with *Mesorhizobium* spp.. Analysis of the more variable IGS region (Gűrtler & Stanisich, 1996), provided better resolution of this group, indicating the existence of at least seven genotypes which showed correspondence to only *M. huakuii, M. loti* and *M. plurifarium*. However, isolates xhj15, xhj18 and xhj23 exhibited the same RFLP pattern with all four enzymes used, indicating their similar genotype.

The *Bradyrhizobium* genus is clearly very diverse, with *B. japonicum* and *B. liaoningense* being closer related to *Afipia, Blastobacter, Nitrobacter* and *Rhodopseudomonas* than to *B. elkanii* (Willems *et al.*, 2001a). According to several authors (Barrera *et al.*, 1997; Willems *et al.*, 2001b), the highly homologous 16S rDNA sequences of bradyrhizobia contain insufficient discriminatory ability to describe this diverse group of symbionts. Alternatively, a recent report by Willems *et al.* (2001c) on sequence analyses of the IGS region of bradyrhizobia, indicated that this genomic region has a discriminatory ability comparable to that of DNA-DNA hybridisation analysis. The bradyrhizobial isolates identified in this study were obtained from a diverse group of host species. Based on their 16S rDNA sequences, three clusters (B, C1 and C2) were identified. The IGS-RFLP results indicate isolates 82a, 70a, 29d of cluster B, to be the same genotype. Although isolate 2 was related to the former isolates, it was clearly separable from them. Cluster C2, which was closely related to *B.*

japonicum and *B. liaoningense*, was a heterogeneous group containing seven genotypes among 10 isolates.

The grouping of 48a, 49b, 40a based on IGS-RFLP data suggest the existence of two genotypes. Their high homology with the divergent *Tachigali versicolor* symbiont, Tv2a-2 further suggests them to represent a separate phylogenetic lineage. According to Doyle *et al.* (1997), *Tachigali* was the most basal known genus of a clade within the Caesalpinioideae that gave rise to the subfamily Mimosoideae, where almost all species participate in root-nodule symbioses. The divergent status of the *Tachigali*-symbiont and these three indigenous isolates remains intriguing, especially since they were obtained from host species within the Papilionoideae subfamily. Initially Doyle *et al.* (1997) hypothesised that the evolution of nodulation within the Papilionoideae occurred in parallel with the origin or origins of the trait in the Caesalpinioideae/ Mimosoideae. However, in the light of evidence suggesting that nodulating flowering plants are more closely related than previously thought, the view has shifted towards a single origin for a predisposition for nodulation (Doyle, 1998). The exact origin of nodulation can however, only be refined in the presence of more phylogenetic sampling of both plant and symbiotic partners.

The failure of isolate PL10a, to group with any of the rhizobial genera was also evident in the whole cell protein profile from an earlier study by Kruger (1998) and a more detailed analysis of this isolate is necessary.

The isolation of different symbionts from the same nodule was previously reported (Dreyfus & Dommergues, 1981; Jenkins *et al.*, 1987; Dagutat, 1995) and is illustrated in this study since xhj12FR and xhj12S was isolated from the same root nodule of *Aspalathus linearis*.

Despite the lack of suitable discriminatory ability of 16S rDNA sequence analyses in some instances, it has proven sufficient to determine the diversity of the rhizobial symbionts associated with leguminous species in South Africa. Previous studies of only morphological traits could by no means have provided insight into the complex intra- and intergenic relationships among this diverse group of microorganisms. Although, most of the indigenous isolates showed good agreement with the known rhizobial genera, others demonstrated

extreme divergence from such genera. It is therefore reasonable to expect that more diverse rhizobia may be present in South African soils which may also be correlated to the diverse range of leguminous plants in the country.