CHAPTER 6

DETERMINATION OF THE DIVERSITY OF ROOT-NODULATING BACTERIA ASSOCIATED WITH *CYCLOPIA* SPP.

ABSTRACT

The diversity of root-nodulating isolates associated with 14 different *Cyclopia* spp. isolated from different localities in the geographic distribution of the *Cyclopia* genus were determined using 16S-23S IGS-RFLP analysis. With the exception of seven isolates, all the isolates grouped distantly from the α-*Proteobacteria* rhizobial reference strains. Partial 16S rDNA sequencing was performed to identify and classify the isolates. The sequencing data confirmed and corroborated the RFLP analysis. All the isolates except the seven α-*Proteobacteria* isolates belonged to the genus *Burkholderia*. A large number of the isolates belonged to the recently described root-nodulating species, *B. tuberum*. Several new *Burkholderia* genotypes were detected. The α-*Proteobacteria* isolates belonged to the genus *Bradyrhizobium*, *Rhizobium tropici* and one isolate displayed a novel genotype.

Keywords: honeybush tea, 16S-23S IGS-RFLP, partial 16S rDNA sequencing, *Burkholderia*

INTRODUCTION

The symbiotic association between legumes and the gram-negative bacteria collectively called rhizobia is an agricultural important association. The bacteria form nodules on the roots of the legumes and as a special adaptation for waterlogged regions on the stems of the legumes. In mature nodules nitrogen fixation and ammonia assimilation occur (Caetano-Anollés, 1997).

The taxonomy of the root-nodulating bacteria changed rapidly the last years as new techniques are employed and more legumes studied. Jordan (1984) included all rhizobia in the family *Rhizobiaceae* in the α-2-subgroup of the *Proteobacteria*. On the website (http://www.cme.msu.edu/bergeys/outline.prn.pdf) of Bergey's Manual rhizobia are included in several different families. *Rhizobium*, *Allorhizobium* and *Sinorhizobium* are placed in the family *Rhizobiaceae*, while *Mesorhizobium* is grouped in the family "Phyllobacteriaceae". *Bradyrhizobium* is placed in the family "Bradyrhizobiaceae", while *Azorhizobium* and the genus *Devosia* in which a newly nodulating species [*Devosia neptuniae* (Rivas *et al.*, 2003)] have been described, belong to the family *Hyphomicrobiaceae*. The other nodulating species of the α-*Proteobacteria* belong to the genus *Methylobacterium* in the family "Methylobacteriaceae" (http://www.cme.msu.edu/bergeys/outline.prn.pdf).

Recently, it became clear that the ability to nodulate and fix nitrogen is not restricted to the α-*Proteobacteria*, but that several species in the β-*Proteobacteria* acquired the ability as well (Chen *et al.*, 2001; Moulin *et al.*, 2001; Vandamme *et al.*, 2002). The two genera involved *Burkholderia* and *Ralstonia* belong to the families "Burkholderiaceae" and "Ralstoniaceae" respectively in the order "Burkholderiales" (Bergey's manual taxonomic list: http://www.cme.msu.edu/bergeys/outline/prn.pdf).

A traditional South African herbal infusion, commonly referred to as honeybush tea is manufactured from the leaves, stems and flowers of mainly *Cyclopia intermedia* (Kouga bush tea) and *C. subternata* (synonym of *C. falcata*) (bush tea) (De Nysschen *et al.*, 1996). The commercial cultivation of several species (*C. intermedia*, *C. subternata*, *C. maculata*, *C.sessiliflora* and *C. genistoides*) is investigated to guard against the overexploitation of the natural populations (Du Toit and Joubert, 1998).

Cyclopia is a genus consisting of 19 accepted species (List of accepted names: http:/www.ildis.org/LegumeWeb/6.00/names/npall/npall_201.shtml) endemic to the fynbos region of South Africa. The genus *Cyclopia* (tribe Podalyrieae, subtribe Podalyriinae) belong to the "genistoid alliance" in the Papilionoideae (Schutte and Van Wyk, 1998), of which the members produce characteristic quinolizidine alkaloids (Polhill 1994; Van Wyk and Schutte, 1995, Van Wyk, 2003). The plants grow in the coastal regions of the Western and Eastern Cape Provinces, from Darling to Port Elizabeth, bounded on the north by the Cederberg, Koue Bokkeveld, Klein Swartberg, Groot Swartberg and Kouga mountain ranges. Most of the species have limited distribution ranges and special habitat preferences (Du Toit *et al.*, 1998).

There is a need for good inoculant strains for the *Cyclopia* commercial plantings. The study of the root-nodulating bacteria associated with different natural populations of *Cyclopia* species covering the geographical distribution of the genus provides a collection of possible inoculant strains and knowledge of the diversity of these symbionts. The strains might also be used for other crop plants grown in similar environmental conditions, since the strains are adapted to such conditions. Some of the strains have been isolated from soil with a pH as low as 3.1 (J. Bloem, personal communication).

The aim of this study was to identify and determine the diversity of the *Cyclopia* nodule isolates. In this study, 16S-23S IGS-RFLP was used to identify, type and differentiate between closely related strains. This method has been used by several researchers (Laguerre *et al.*, 1996; LeBlond-Bourget *et al.*, 1996; Khbaya *et al.*, 1998; Guo *et al.*, 1999; Doignon-Bourcier *et al.*, 2000; Grundmann *et al.*, 2000; Bala *et al.*, 2002). Partial 16S rDNA sequences was used to identify and determine the taxonomic position of the isolates. This method has been used by several researchers to determine and confirm the identity of new isolates (Lafay and Burdon, 1998; Vinuesa *et al.*, 1998; McInroy *et al.*, 1999; Van Berkum and Fuhrmann, 2000; Mehta and Rosato, 2001; Odee *et al.*, 2002; Qian *et al.*, 2003).

MATERIALS AND METHODS

Bacterial strains used

The strains used in this study were received from the Agricultural Research Council-Plant Protection Research Institute (Private Bag X134, Pretoria, 0001, South Africa) [Table 6.1] and the Botany Department (University of Cape Town, Rondebosch, 7701, Cape Town, South Africa) [Table 6.2]. All the strains received from the University of Cape Town were authenticated nodulating strains. A selection of the strains received from the ARC was used in plant nodulation studies and confirmed as root-nodulating. Reference strains of the different rhizobial genera used in this study were obtained from the culture collections of the Laboratorium voor Microbiologie (LMG), University of Gent, Gent, Belgium, the United States Department of Agriculture (USDA), *Rhizobium* Culture Collection, Maryland, USA and the Laboratoire des Symbioses Tropicales et Méditerranéennes (STM), Montpellier, France.

Maintenance of bacterial 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 5-7 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.

Extraction of genomic DNA

A modified method for proteinase-K (Roche Molecular Biochemicals) treated cells as described by Laguerre *et al.* (1997) was used. A fresh culture of each strain, which had been checked for purity, was streaked on a tryptone yeast (TY) agar slant [0.5% (m/v) tryptone (Difco), 0.3% (m/v) yeast extract (Biolab), 0.13% (m/v) $CaCl_2.6H_2O$ (UniLab), 1.5% (m/v) bacteriological agar] in a screw-cap tube. TY reduces slime formation by the rhizobia. The

strains were incubated at 28 °C and checked for sufficient growth. Sterile 4.5 ml dH_2O was added to the slant growth to harvest the cells. An inoculation loop was used to aid the release of cells clinging to the agar. The volume of the water added was adjusted according to the amount of growth. Less water was used if the growth was poor and *vice versa*. The cell-suspension was collected in a clear plastic tube and vortexed to ensure a uniform suspension. The absorbancy of the suspension was measured with dH_2O as the spectophotometric blank at 620 nm. A formula was used to determine the volume of the cellsuspension to be treated further. The volume to be used in ml is equal to 0.2 divided by the absorbancy at 620 nm. Two tubes of the same strain were filled with the appropriate volume of cells and centrifuged at 13 000 *g* for 5 minutes at 4 °C. The supernatant was discarded and the excess media blotted dry. One of the tubes was stored at -20 °C for future use. In the second tube, 100 µl ddH_2O, 100 µl Tris-HCl (10 mM, pH 8.2) and 10 µl proteinase-K (15 mg/ml) (Roche Molecular Biochemicals) were added to the cell pellet. The mixture was incubated at 55 °C overnight. In order to inactivate the proteinase-K the mixture was boiled for 10 minutes. The cell lysates were stored at -20 °C until needed.

Table 6.1: Authenticated root-nodulating isolates of indigenous *Cyclopia* species included in this study received from the ARC-PPRI.

Isolate	Host species	Locality	Isolate	Host species	Locality
CS 1	*C. subternata*	Dennehoek, Joubertina	**Cses 1**	*C. sessiliflora*	Plattekloof, Heidelberg
CS 2	*C. subternata*	Dennehoek, Joubertina	**Cses 2**	*C. sessiliflora*	Plattekloof, Heidelberg
CS 3	*C. subternata*	Dennehoek, Joubertina	**Cses 3**	*C. sessiliflora*	Plattekloof, Heidelberg
CS 5	*C. subternata*	Dennehoek, Joubertina	**Cses 4**	*C. sessiliflora*	Plattekloof, Heidelberg
CS 6	*C. subternata*	Dennehoek, Joubertina	**Cses 5**	*C. sessiliflora*	Plattekloof, Heidelberg
CS 7	*C. subternata*	Dennehoek, Joubertina	**Cses 6**	*C. sessiliflora*	Plattekloof, Heidelberg
CI 1	*C. intermedia*	Dennehoek, Joubertina	**Cses 7**	*C. sessiliflora*	Plattekloof, Heidelberg
CI 2	*C. intermedia*	Dennehoek, Joubertina	**CF 1**	*C. falcata*	Large Winterhoek mountain Porterville
CI 2b	*C. intermedia*	Dennehoek, Joubertina	**CG 1**	*C. genistoides*	Silwerstroomstrand, Darlin
CI 3	*C. intermedia*	Dennehoek, Joubertina	**CG 4**	*C. genistoides*	Rondeberg, Darling
CI 4b	*C. intermedia*	Dennehoek, Joubertina	**Clong 1**	*C. longifolia*	Thornhill, Humansdorp
CI 6	*C. intermedia*	Onverwacht, Garcia Pass	**Clong 2**	*C. longifolia*	Thornhill, Humansdorp
CI 9	*C. intermedia*	Onverwacht, Garcia Pass	**Clong 3**	*C. longifolia*	Thornhill, Humansdorp
Cint S2*	*C. subternata* #	Dennehoek, Joubertina	**Clong 4**	*C. longifolia*	Thornhill, Humansdorp
Cint I1*	*C. intermedia* #	Dennehoek, Joubertina	**Clong 5**	*C. longifolia*	Thornhill, Humansdorp
Cint I2*	*C. intermedia* #	Dennehoek, Joubertina	**CM 1**	*C. maculata*	Paarlberg, Paarl
Cint I4*	*C. intermedia* #	Dennehoek, Joubertina	**CM 2**	*C. maculata*	Garcia Pass, Riversdal
Csub I1**	*C. intermedia* #	Dennehoek, Joubertina	**CM 3**	*C. maculata*	Garcia Pass, Riversdal
Csub I5**	*C. intermedia* #	unknown	**CB 2**	*C. buxifolia*	Helderberg, Somerset-Wes
Csub S1**	*C. subternata* #	Dennehoek, Joubertina	**CD 1**	*C. dregeana*	Du Toitskloof, Paarl
Csub S3**	*C. subternata* #	Dennehoek, Joubertina	**CD 4**	*C. dregeana*	Du Toitskloof, Paarl
Cses I1***	*C. intermedia* #	Dennehoek, Joubertina	**CD 9**	*C. dregeana*	Du Toitskloof, Paarl
Cses I2***	*C. intermedia* #	Dennehoek, Joubertina	**CD 10**	*C. dregeana*	Du Toitskloof, Paarl

Table 6.1: continued

Isolate	Host species	Locality	Isolate	Host species	Locality
Cses S1gr.***	*C. subternata* #	Dennehoek, Joubertina	**CD 11**	*C. dregeana*	Du Toitskloof, Paarl
Cses S1kl.***	*C. subternata* #	Dennehoek, Joubertina	**CD 12a**	*C. dregeana*	Du Toitskloof, Paarl
Cses S2gr.***	*C. subternata* #	Dennehoek, Joubertina	**CD 13**	*C. dregeana*	Du Toitskloof, Paarl
Cses S2kl.***	*C. subternata* #	Dennehoek, Joubertina	**Cpub 4**	*C. pubescens*	Next to N1, Port Elizabeth
Cses S3***	*C. subternata* #	Dennehoek, Joubertina	**Cpub 5**	*C. pubescens*	Next to N1, Port Elizabeth
Cses S7***	*C. subternata* #	Dennehoek, Joubertina	**Cpub 6**	*C. pubescens*	Next to N1, Port Elizabeth
Cmey 1	*C. meyeriana*	Kunje, Citrusdal	**Cplic 1**	*C. plicata*	Mannetjiesberg, Uniondale

* Re-isolated from plant inoculation test performed on *C. intermedia*

** Re-isolated from plant inoculation test performed on *C. subternata*

*** Re-isolated from plant inoculation test performed on *C. sessiliflora*

Original host plant of strain

Table 6.2: Authenticated root-nodulating strains isolated from indigenous *Cyclopia* species received from the Botany Department (UCT) and included in this study.

Isolate	Host species	Locality	Isolate	Host species	Locality
UCT 2	*C. genistoides*	Rein's Farms	**UCT 35**	*C. glabra*	Matroosberg
UCT 3	*C. genistoides*	Rein's Farms	**UCT 36**	*C. galioides*	Cape Point
UCT 4	*C. genistoides*	Rein's Farms	**UCT 37**	*C. galioides*	Cape Point
UCT 5	*C. genistoides*	Pearly Beach	**UCT 38**	*C. galioides*	Cape Point
UCT 6	*C. genistoides*	Pearly Beach	**UCT 39**	*C. galioides*	Cape Point
UCT 7	*C. genistoides*	Pearly Beach	**UCT 40**	*C. galioides*	Cape Point
UCT 8	*C. genistoides*	Betty's Bay	**UCT 41**	*C. plicata*	Kougaberg
UCT 9	*C. genistoides*	Betty's Bay	**UCT 42**	*C. plicata*	Kougaberg
UCT 10	*C. genistoides*	Rondeberg	**UCT 43**	*C. meyeriana*	Hottentots Holland mountains
UCT 11	*C. genistoides*	Rondeberg	**UCT 44**	*C. meyeriana*	Hottentots Holland mountains
UCT 13	*C. genistoides*	Rondeberg	**UCT 45**	*C. meyeriana*	Bains Kloof
UCT 14	*C. genistoides*	Rondeberg	**UCT 46**	*C. meyeriana*	Hottentots Holland mountains
UCT 15	*C. genistoides*	Constantiaberg	**UCT 47**	*C. glabra*	unknown
UCT 16	*C. genistoides*	Constantiaberg	**UCT 48**	*C. maculata*	Jonkershoek
UCT 17	*C. genistoides*	Constantiaberg	**UCT 49**	*C. genistoides*	Constantiaberg
UCT 18	*C. genistoides*	Constantiaberg	**UCT 50**	*C. sessiliflora*	Callie's farm, Heidelberg
UCT 19	*C. genistoides*	Constantiaberg	**UCT 52**	*C. plicata*	unknown
UCT 20	*C. genistoides*	Paardeberg	**UCT 53**	*C. plicata*	unknown
UCT 21	*C. genistoides*	Paardeberg	**UCT 55**	*C. plicata*	unknown
UCT 22	*C. maculata*	Jonkershoek	**UCT 56**	*C. meyeriana*	Hottentots Holland mountains
UCT 24	*C. maculata*	Jonkershoek	**UCT 57**	*C. subternata*	Port Alfred Pass
UCT 25	*C. intermedia*	Swartberg Pass	**UCT 58a**	*C. subternata*	Port Alfred Pass
UCT 26	*C. intermedia*	Swartberg Pass	**UCT 60**	*C. meyeriana*	Bains Kloof
UCT 27bii	*C. subternata*	Waboomskraal farm (wild tea)	**UCT 61**	*C. subternata*	Garcia Pass, Riversdal
UCT 28	*C. subternata*	Waboomskraal farm (wild tea)	**UCT 62**	*C. genistoides*	Pearly Beach

Table 6.2: continued

Isolate	Host species	Locality	Isolate	Host species	Locality
UCT 29	*C. sessiliflora*	Callie's farm, Heidelberg	**UCT 63**	*C. genistoides*	Betty's Bay
UCT 30	*C. sessiliflora*	Callie's farm, Heidelberg	**UCT 67**	*C. glabra*	unknown
UCT 31	*C. sessiliflora*	Grootvadersbosch	**UCT 69**	*C. glabra*	unknown
UCT 32	*C. buxifolia*	McGregor	**UCT 70**	*C. maculata*	Jonkershoek
UCT 33	*C. buxifolia*	McGregor	**UCT 71**	*C. glabra*	unknown
UCT 34	*C. glabra*	Matroosberg	**UCT 73**	*C. genistoides*	Betty's Bay

Table 6.3: Reference strains included in the IGS RFLP study

Reference strain	Strain number	Reference strain	Strain number
A. caulinodans	LMG 6465^T	***M. huakuii***	USDA 4779^T
A. caulinodans	USDA 4892^T	***M. huakuii***	LMG 14107^T
B. elkanii	USDA 76^T	***R. etli***	LMG 17827^T
B. japonicum	LMG 6138^T	***R. galegae***	LMG 6214^T
B. japonicum	USDA 6^T	***R. tropici***	LMG 9503^T
***Bradyrhizobium* sp.**	LMG 8319	***S. medicae***	LMG 18864^T
Burkholderia phymatum	STM 815^T	***S. meliloti***	LMG 6133^T
Burkholderia tuberum	STM 678^T	***S. saheli***	USDA 4893^T
M. ciceri	LMG 14989^T		

^{T} Type strain

Amplification of the 16S-23S IGS region and 16S rDNA gene

The 16S-23S IGS regions of the different strains including the reference strains (Table 6.3) were amplified with the primers FGPS1490 and FGPS132 (Table 6.4) as described by Laguerre *et al.* (1996). In a 50 µl PCR reaction mix the following were added: 5 µl of the cell lysate, 50 pmol of each primer, 250 µM of each dNTP, 1.5 mM $MgCl_2$ and 0.5 U Supertherm Taq DNA polymerase (Southern Cross Biotechnology). The amplification reaction was done in a Perkin Elmer GeneAmp PCR System 2400 thermocycler with the following profile: an initial 3 minutes of denaturation at 95 °C, 35 cycles of denaturation at 94 °C for 30 seconds, annealing at 55 °C for 30 seconds and extension at 72 °C for 1 minute followed by a final extension step at 72 °C for 5 minutes. Aliquots (5 µl) of the PCR reactions were examined to determine the size, purity and concentration of the products with horizontal agarose electrophoresis (Sambrook *et al.*, 1989) using 0.9% (m/v) agarose gels (Promega) in a 1 x TAE buffer (40 mM Tris-HCl, 20 mM NaOAc and 1 mM EDTA pH 8.5) stained with ethidium bromide (10 mg/ml). IGS PCR products were visualised by UV fluorescence (results

not shown). The standard marker 1 Kb PLUS DNA Ladder (GibcoBRL®) was included on each gel.

Amplification of the 16S rDNA gene of selected strains (Table 5) were performed with the primers fD1 and rP2 (Table 6.4) as described by Weisburg *et al.* (1991). The linker sequences of the primers were not included in the primer synthesis, since cloning of the products was not anticipated. These shorter primers were thus designated fD1SHRT and rP2SHRT. The PCR mixture of each strain contained: 5 µl of the cell lysate, 50 pmol of each primer, 250 µM of each dNTP, 1.5 mM $MgCl_2$ and 0.5 U Gold Taq DNA polymerase (Southern Cross Biotechnology) in a 50 µl reaction volume. The PCR reactions were done on a Perkin Elmer GeneAmp PCR System 2400 thermocycler using the same thermal profile as used in the amplification of the IGS region. The concentration, purity and size of the products were evaluated by running an aliquot (5 µl) of each reaction on 0.9% (m/v) horizontal agarose gels (Promega) (results not shown). The standard marker, molecular marker VI (Roche Molecular Biochemicals), was included on each gel.

Table 6.4: Primers used in the amplification and/or sequencing of the different genes analysed in this study.

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

* All the primers were synthesised by Roche Molecular Biochemicals, Mannheim, Germany

The 16S PCR product of each strain was purified, since any traces of unincorporated dNTPs, primers, etc. can negatively influence the 16S sequencing reaction. The products were purified using a Qiagen QIAquick PCR Purification kit (Southern Cross Biotechnology). Purification reactions were done as prescribed by the manufacturer. The concentration and purity of each purification reaction was verified visually. An aliquot (1 µl) of each purified 16S PCR product was run on 0.9% (m/v) horizontal agarose gels (Promega) (results not shown). On each gel, a standard marker, molecular marker VI (Roche Molecular Biochemicals) was included.

16S-23S IGS-RFLP

The IGS PCR products without prior purification were restricted with four tetrameric restriction enzymes. The enzymes *Alu*I, *Cfo*I, *Hae*III and *Msp*I (Roche Molecular Biochemicals) were each used to digest an aliquot of the products. In each 10 µl reaction volume the following was added: 5 µl of the PCR product, 5 U of the enzyme as well as the optimal restriction buffer for each enzyme as prescribed by the manufacturer. The reactions were incubated overnight at 37 °C. The restricted products were analysed on a 3.5% (m/v) horizontal agarose gel in a Hybaid Maxi Gel System for 180 minutes at 80V. Molecular weight marker VIII (Roche Molecular Biochemicals) was loaded as a standard in specified lanes on each gel. All the profiles were analysed visually, grouping isolates with similar bands in the same enzyme profile type. Additionally the gel-files were analysed with GelcomparII (Applied Maths, Kortrijk, Belgium) using the molecular weight marker VIII as the standard lane. The Dice coefficient (Nei and Lei, 1979) was used to calculate a distance matrix for each enzyme and the unweighted pair group method with arithmetic mean (UPGMA) was used to construct a dendrogram. The data of all four enzymes were combined with GelcomparII (Applied Maths, Kortrijk, Belgium) as described by the manufacturer and presented as a UPGMA constructed dendrogram.

16S rDNA sequencing

The partial sequence of each purified amplified 16S rDNA product of the chosen isolates were determined with the internal forward primer 16SRNAII-S (Kuhnert *et al.*, 1996) and the reverse primer rP2SHRT (Weisburg *et al.*, 1991) using the ABI Prism BigDye™ Terminator Cycle Sequencing Ready Reaction kit (AmpliTaq® DNA Polymerase, FS) (PE Applied Biosystems). Each 5 µl sequencing reaction contained the following: 2 µl of the ready reaction mix supplied with the kit which contains the dye terminators, dNTP's, AmpliTaq® DNA polymerase, $MgCl_2$ and Tris-HCl buffer pH 9.0; 12.5 pmol primer and approximately 100 ng template DNA. The sequencing reactions were carried out in a Perkin Elmer GeneAmp PCR System 2400 thermocycler with the following thermal profile: an initial denaturation at 96 °C for 5 seconds followed by 25 cycles of denaturation (96 °C for 10 seconds), annealing (50 °C for 5 seconds) and extension (60 °C for 4 minutes). The products were precipitated using the protocol as suggested by the manufacturer. The reaction tubes

were placed on ice, while 4 µl sterile ddH_2O and 16 µl ice cold absolute ethanol were added to the sequencing reaction mix. The tubes were vortexed briefly and placed in the dark for 30 minutes to aid the precipitation of the sequencing products. This was followed by centrifugation at maximum speed for 30 minutes. The supernatant was discarded and any excess moisture was removed. Washing of the pellet was done with 50 µl ice cold 70% (v/v) ethanol. The tubes were centrifuged for 5 minutes at maximum speed. After discarding the supernatant and removing any excess moisture the pellets was vacuum dried for 15 minutes. The tubes were then stored at -20 °C until used. 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 deionised 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.

Phylogenetic analysis of the 16S rDNA sequences

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). Both strands were sequenced with the primers used and the strands could be aligned to correct ambiguous positions. The ClustalX programme (Thompson *et al.*, 1997) was used to analyse the edited sequences as well as the reference sequences obtained from GenBank (Table 6.5). A distance matrix was constructed by pair-wise alignment of the sequences. 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 phylogenetic tree was visualised with NJplot (Perrière and Gouy, 1996).

Table 6.5: Reference sequences obtained from Genbank[1] included in the partial 16S rDNA sequence analysis.

Reference strain	Strain number	Host plant or relevant characteristics	Genbank[1] Accession number
Agrobacterium rhizogenes	LMG 152	NS	X67224
Allorhizobium undicola	LMG 11875^T	*Neptunia natans*	Y17047
Azorhizobium caulinodans	LMG 6465^T	*Sesbania rostrata*	X67221
Bradyrhizobium elkanii	USDA 76^T	*Glycine max*	U35000
Bradyrhizobium japonicum	LMG 6138^T	*Glycine max*	X66024
Bradyrhizobium liaoningense	LMG 18230^T	*Glycine max*	AJ250813
Bradyrhizobium yuanmingense	CCBAU 10071^T	*Lespedeza cuneata*	AF193818
Burkholderia ambifaria	MVPC 1/4	*B. cepacia* complex	AY028444
Burkholderia andropogonis	ATCC 23061^T	Sorghum (*Sorghum bicolor)*	X67037
Burkholderia anthina	R-4183^T	Rhizosphere soil, *B. cepacia* complex	AJ420880
Burkholderia brasilensis	M130	Plant-associated N_2-fixer	AJ238360
Burkholderia caledonica	LMG 19076^T	Rhizosphere soil	AF215704
Burkholderia caribiensis	LMG 18531^T	Vertisol microaggregates	Y17009
Burkholderia caryophylli	ATCC 25418^T	Carnation (*Dianthus caryophyllus)*	AB021423
Burkholderia cepacia	LMG 12615	Cystic fibrosis sputum, plant associated genomovar III	AF265235
Burkholderia fungorum	LMG 16225^T	*Phanerochaete chrysosporium*	AF215705
Burkholderia gladioli	LMG 11626	Fermented coconut	U96934
Burkholderia glathei	LMG 14190^T	Fossil lateritic soil	U96935
Burkholderia glumae	LMG 2196^T	Rice (*Oryza sativa)*	U96931
Burkholderia graminis	AUS 35	Rhizosphere	U96941
Burkholderia kururiensis	LMG 19447^T	Trichloroethylene degrader	AB024310
Burkholderia mallei	NCTC 10260	NS	AF110187
Burkholderia multivorans	LMG 13010^T	*B. cepacia* complex	Y18703
Burkholderia phenazinium	LMG 2247^T	Soil enriched with threonine	U96936
Burkholderia phymatum	STM 815^T	*Machaerium lunatum*	AJ302312
Burkholderia plantarii	LMG 9035^T	*Oryza sativa* pathogen	U96933
Burkholderia pseudomallei	V686	Soil	AF093052
Burkholderia pyrrocinia	LMG 14191^T	NS	U96930
Burkholderia sacchari	LMG 19450^T	Soil from sugarcane plantation	AF263278
Burkholderia sordicola	SNU 020123	Associated with white rot fungus *Phanerochaete sordicola*	AF512827
Burkholderia stabilis	LMG 14294^T	Formerly *B. cepacia* complex IV	AF148554
Burkholderia thailandensis	E264^T	*Pseudomallei* group	U91838
Burkholderia tropicalis	Ppe8^T	Plant-associated N_2-fixer	AJ420332
Burkholderia tuberum	STM 678^T	*Aspalathus carnosa*	AJ302311
Burkholderia ubonensis	GTC-P3-415	NS	AB030584
Burkholderia vietnamensis	LMG 10929^T	N_2-fixer from rice rhizophere	AF097534
Devosia neptuniae	J1^T	*Neptunia natans*	AF469072
Ensifer adhaerens	LMG 20582	NS	AY040360
Kaistia adipata	Chj 404^T	Rhizobiaceae group	AY039817
Mesorhizobium amorphae	ACCC 19665	*Amorpha fruticosa*	AF041442
Mesorhizobium chacoense	PR-5^T	*Prosopis alba*	AJ278249
Mesorhizobium ciceri	UPM-Ca7^T	*Cicer arietinum*	U07934
Mesorhizobium huakuii	IAM 14158^T	*Astragalus sinicus*	D12797
Mesorhizobium loti	LMG 6125^T	*Lotus corniculatus*	X67229
Mesorhizobium mediterraneum	UPM-Ca36^T	*Cicer arietinum*	L38825
Mesorhizobium plurifarium	LMG 11892^T	*Acacia senegal*	Y14158
Mesorhizobium tianshanense	A-1BS^T	*Glycyrrhiza pallidiflora*	Y71079
Methylobacterium nodulans	ORS 2060^T	*Crotalaria podocarpa*	AF220763

Table 6.5: continued

Reference strain	Strain number	Host plant or relevant characteristics	Genbank[1] Accession number
Pandoraea norimbergensis	NS	Alkaliphilic sulphur oxidiser	Y09879
Ralstonia picketti	MSP 3	Rhizosphere, soil	AB004790
Ralstonia solanacearum	ATCC 11696	*Lycopersicon lycopersicum*	X67036
Ralstonia taiwanensis	LMG 19424^T	*Mimosa pudica*	AF300324
Rhizobium etli	CFN 42^T	*Phaseolus vulgaris*	U28916
Rhizobium galegae	USDA 3394	*Galega officinalis*	AF025853
Rhizobium gallicum	R602sp^T	*Phaseolus vulgaris*	U86343
Rhizobium giardinii	H152^T	*Phaseolus vulgaris*	U86344
Rhizobium hainanensis	I66^T	*Desmodium sinuatum*	U71078
Rhizobium huautlense	USDA 4900^T	*Sesbania herbacae*	AF025852
Rhizobium indigoferae	CCBAU 71042^T	*Indigofera amblyantha*	AY034027
Rhizobium leguminosarum	LMG 8820	*Phaseolus vulgaris*	X67227
Rhizobium mongolense	USDA 1844^T	*Medicago ruthenica*	U89817
Rhizobium sullae	IS123^T	*Hedysarum coronarium*	Y10170
Rhizobium tropici	CIAT 899^T	*Phaseolus vulgaris*	U89832
Rhizobium yanglingense	CCBAU 71462	*Coronilla varia*	AF195031
Sinorhizobium arboris	HAMBI 1552^T	*Prosopis chilensis*	Z78204
Sinorhizobium fredii	LMG 6217^T	*Glycine max*	X67231
Sinorhizobium kostiense	HAMBI 1489^T	*Acacia senegal*	Z78203
Sinorhizobium kummerowiae	CCBAU 71714^T	*Kummerowia stipulacea*	AY034028
Sinorhizobium medicae	A321^T	*Medicago truncatula*	L39882
Sinorhizobium meliloti	LMG 6133^T	*Medicago sativa*	X67222
Sinorhizobium morelense	Lc04^T	*Leucaena leucocephala*	AY024335
Sinorhizobium saheli	LMG 7837^T	*Sesbania pachycarpa*	X68390
Sinorhizobium terangae	LMG 6463	*Sesbania rostrata*	X68387
Sinorhizobium xinjiangensis	IAM 14142	*Glycine max*	D12796

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

T Type strain

ACCC Agricultural Center of Culture Collection, Chinese Academy of Agriculture, Beijing, China

ATCC American Type Culture Collection, Rockville, Maryland, USA

CCBAU Culture Collection of Beijing Agricultural University, Beijing, People's Republic of China

CFN Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Mexico

CIAT Rhizobium Collection, Centro International de Agricultura Tropical, Cali, Colombia

DSM Deutsche Sammlung von Mikroorganismen, Braunschweig, Germany

HAMBI Culture Collection of the Department of Applied Chemistry and Microbiology, University of Helsinki, Helsinki, Finland

IAM Institute of Applied Microbiology, University of Tokyo, Tokyo, Japan

IFO Institute for Fermentation, Osaka, Japan

LMG BCCM™/LMG Bacteria Collection, Laboratorium voor Microbiologie, University of Gent, Gent, Belgium

NCIMB National Collections of Industrial and Marine Bacteria, Aberdeen, Scotland, UK

NCPPB National Collection of Plant Pathogenic Bacteria, Harpenden Laboratory, Hertfordshire, UK

NCTC	National Collection of Type Cultures, Central Public Health Laboratory, London, UK
ORS	ORSTOM Collection, Institut Français de Recherche Scientifique pour le Développement en Coopération, Dakar, Senegal
STM	Laboratoire des Symbioses Tropicales et Méditerranéennes, Montpellier, France
UPM	Universidad Politécnica Madrid, Spain
USDA	United States Department of Agriculture, *Rhizobium* Culture Collection, Beltsville Agricultural Research Center, Beltsville, MD, USA

RESULTS

IGS PCR

The IGS products of the isolates were successfully amplified with the primers FGPS1490 and FGPS132 as used by Laguerre *et al.* (1997) [results not shown]. Amplification products of the *Cyclopia* isolates varied in the range of approximately 700 bp to 1250 bp. The isolates in cluster 2A and 2B (see Fig. 6.1) had amplification products of approximately 730-780 bp. The amplification products of the isolates of cluster 2C are larger and range from 820 to 860 bp. All the IGS products of the isolates in cluster 4A and 4B have bands in the range of 710-780 bp. For the isolates in cluster 4C two additional less prominent bands of approximately 700 bp and 600 bp were amplified in addition to the approximately 800 bp prominent band. However, isolate UCT 30 which groups into cluster 4C has only the approximately 800 bp band. Isolates of cluster 4D have a band in the range of 800 bp. The IGS products of the isolates: UCT 42 (~ 970 bp), UCT 50 (~ 1250 bp), UCT 53 (~ 900 bp), UCT 55 (~ 970 bp), Cmey 1 (~950 bp), Cplic 1(~ 1250bp) and Cpub 4 (~ 950 bp) which group into cluster 1 were significantly larger in comparison with the *Cyclopia* isolates of cluster 2 and cluster 4.

The reference strains included in this study gave different size products. The root-nodulating *Burkholderia* species, *B. tuberum* STM 678^T and *B. phymatum* STM 815^T gave IGS products of approximately 750 bp. The IGS products of the *Azorhizobium caulinodans*, *Bradyrhizobium* and the *Mesorhizobium huakuii* reference strains were in the range of approximately 900 bp. In the *M. ciceri* strain a larger product than that of *M. huakuii* in the magnitude of 1000 bp was amplified. The *Rhizobium* and *Sinorhizobium* species strains included in the study gave larger products than the other rhizobial strains. *R. etli*, *R. galegae*, *S. medicae*, *S. meliloti* and *S. saheli* strains gave an IGS product in the range of approximately

1250 bp. The IGS product of *R. tropici* LMG 9503^T was approximately 100 bp smaller than that of the other *Rhizobium-Sinorhizobium* strains.

IGS-RFLP

The enzymes *Alu*I, *Cfo*I, *Hae*III and *Msp*I generated 17, 20, 9 and 22 different restriction profiles respectively, excluding the profiles generated for the rhizobial reference strains. The discrimination level of *Hae*III was the lowest. Four clusters could be distinguished on the dendrogram constructed from the combined profiles of the four restriction enzymes (Fig. 6.1). All the root-nodulating α-*Proteobacteria* reference strains included in this study grouped in cluster 1, while cluster 2 contained most of the *Cyclopia* isolates and *Burkholderia tuberum* STM 678^T. *Burkholderia phymatum* STM 815^T was the only isolate of cluster 3. No reference strains clustered in cluster 4. Cluster 1, which contained the different rhizobial genera, is the most heterogeneous collection of strains. Cluster 2 and cluster 4 are relatively homogeneous clusters, sharing similar restriction profiles for specific enzymes.

Table 6.6: Restriction enzyme (*Alu*I, *Cfo*I, *Hae*III and *Msp*I) profiles of the amplified intergenic spacer regions of the *Cyclopia* isolates and the reference strains included in the IGS RFLP study.

Isolate number	Host plant / species designation	*Alu*I	*Cfo*I	*Hae*III	*Msp*I
CS 1	*C. subternata*	b	g	NS	k
CS 2	*C. subternata*	b	g	NS	k
CS 3	*C. subternata*	f	j	a	h
CS 5	*C. subternata*	b	g	NS	k
CS 6	*C. subternata*	b	h	NS	NS
CS 7	*C. subternata*	b	g	NS	k
CI 1	*C. intermedia*	c	a	NS	NS
CI 2	*C. intermedia*	f	j	a	h
CI 2b	*C. intermedia*	b	g	NS	k
CI 3	*C. intermedia*	f	i	a	a
CI 4b	*C. intermedia*	e	k	b	l
CI 6	*C. intermedia*	c	a	NS	NS
CI 9	*C. intermedia*	f	i	a	a
Cint S2	*C. subternata*	b	g	NS	k
Cint I1	*C. intermedia*	f	i	a	a
Cint I2	*C. intermedia*	f	j	a	a
Cint I4	*C. intermedia*	e	k	b	l
Csub I1	*C. intermedia*	f	i	a	a
Csub S1	*C. subternata*	b	g	NS	k
Csub S3	*C. subternata*	b	g	NS	k
Cses 1	*C. sessiliflora*	f	i	a	a
Cses 2	*C. sessiliflora*	b	b	NS	j

Table 6.6: continued

Isolate number	Host plant / species designation	*Alu*I	*Cfo*I	*Hae*III	*Msp*I
Cses 3	*C. sessiliflora*	f	i	a	a
Cses 4	*C. sessiliflora*	b	b	NS	j
Cses 5	*C. sessiliflora*	b	b	NS	j
Cses 6	*C. sessiliflora*	b	b	NS	j
Cses 7	*C. sessiliflora*	b	b	NS	j
Cses I1	*C. intermedia*	f	i	a	a
Cses I2	*C. intermedia*	f	j	a	a
Cses S1gr.	*C. subternata*	f	j	a	a
Cses S1kl.	*C. subternata*	b	g	NS	k
Cses S2gr.	*C. subternata*	b	g	NS	k
Cses S2kl.	*C. subternata*	b	g	NS	k
Cses S3	*C. subternata*	c	a	NS	NS
Cses S7	*C. subternata*	b	k	b	l
CF 1	*C. falcata*	b	g	NS	k
CG 1	*C. genistoides*	f	j	a	m
CG 4	*C. genistoides*	f	e	a	m
Clong 1	*C. longifolia*	c	a	NS	NS
Clong 2	*C. longifolia*	c	a	NS	NS
Clong 3	*C. longifolia*	f	j	a	a
Clong 4	*C. longifolia*	c	a	NS	NS
Clong 5	*C. longifolia*	c	a	NS	NS
CM 1	*C. maculata*	a	a	NS	NS
CM 2	*C. maculata*	c	a	NS	NS
CM 3	*C. maculata*	c	a	NS	NS
CB 2	*C. buxifolia*	q	a	NS	NS
CD 1	*C. dregeana*	e	k	d	l
CD 4	*C. dregeana*	e	k	d	l
CD 9	*C. dregeana*	b	s	NS	NS
CD 10	*C. dregeana*	e	l	c	n
CD 11	*C. dregeana*	e	l	a	n
CD 12a	*C. dregeana*	e	l	a	n
CD 13	*C. dregeana*	c	a	NS	b
Cpub 4	*C. pubescens*	i	m	h	r
Cpub 5	*C. pubescens*	f	n	a	m
Cpub 6	*C. pubescens*	f	i	a	m
Cmey 1	*C. meyeriana*	n	c	g	q
Cplic 1	*C. plicata*	j	o	f	s
Csub I5	*C. intermedia*	f	j	a	a
UCT 2	*C. genistoides*	f	e	a	m
UCT 3	*C. genistoides*	f	j	a	m
UCT 4	*C. genistoides*	f	j	a	d
UCT 5	*C. genistoides*	f	e	a	e
UCT 6	*C. genistoides*	f	j	a	e'
UCT 7	*C. genistoides*	f	j	a	e
UCT 8	*C. genistoides*	f	e	a	c
UCT 9	*C. genistoides*	f	j	a	c
UCT 10	*C. genistoides*	f	j	a	c
UCT 11	*C. genistoides*	f	a	a	m
UCT 13	*C. genistoides*	f	a	a	m
UCT 14	*C. genistoides*	f	a	a	m
UCT 15	*C. genistoides*	f	j	a	h
UCT 16	*C. genistoides*	f	j	a	h
UCT 17	*C. genistoides*	f	j	a	h
UCT 18	*C. genistoides*	f	j	a	h

Table 6.6: continued

Isolate number	Host plant / species designation	*Alu*I	*Cfo*I	*Hae*III	*Msp*I
UCT 19	*C. genistoides*	f	j	a	d
UCT 20	*C. genistoides*	d	a	NS	NS
UCT 21	*C. genistoides*	d	a	NS	NS
UCT 22	*C. maculata*	p	a	NS	NS
UCT 24	*C. maculata*	f	j	a	i
UCT 25	*C. intermedia*	f	i	a	a
UCT 26	*C. intermedia*	e	l	c	n
UCT 27bii	*C. subternata*	f	i	a	a
UCT 28	*C. subternata*	g	t	a	o
UCT 29	*C. sessiliflora*	b	k	a	p
UCT 30	*C. sessiliflora*	k	p	NS	t
UCT 31	*C. sessiliflora*	f	j	a	h
UCT 32	*C. buxifolia*	f	i	a	a
UCT 33	*C. buxifolia*	f	i	a	a
UCT 34	*C. glabra*	a	a	NS	b
UCT 35	*C. glabra*	a	q	NS	NS
UCT 36	*C. galioides*	f	e	a	d
UCT 37	*C. galioides*	f	j	a	c
UCT 38	*C. galioides*	f	j	a	e
UCT 39	*C. galioides*	f	j	a	e
UCT 40	*C. galioides*	f	j	a	e′
UCT 41	*C. plicata*	b	l	NS	NS
UCT 42	*C. plicata*	m	r	g	f
UCT 43	*C. meyeriana*	a	a	NS	b
UCT 44	*C. meyeriana*	a	a	NS	b
UCT 45	*C. meyeriana*	a	a	NS	b
UCT 46	*C. meyeriana*	a	a	NS	b
UCT 47	*C. glabra*	b	a	NS	b
UCT 48	*C. maculata*	a	a	NS	NS
UCT 49	*C. genistoides*	f	j	a	h
UCT 50	*C.sessiliflora*	h	d	e	g
UCT 52	*C. plicata*	g	l	a	v
UCT 53	*C. plicata*	l	m	g	u
UCT 55	*C. plicata*	m	r	g	r
UCT 56	*C. meyeriana*	a	a	NS	b
UCT 57	*C. subternata*	c	a	NS	b
UCT 58a	*C. subternata*	c	a	NS	b
UCT 60	*C. meyeriana*	a	a	NS	b
UCT 61	*C. subternata*	p	a	NS	NS
UCT 62	*C. genistoides*	f	e	a	e
UCT 63	*C. genistoides*	f	e	a	h
UCT 67	*C. glabra*	a	a	NS	b
UCT 69	*C. glabra*	a	a	NS	b
UCT 70	*C. maculata*	f	e	a	i
UCT 71	*C. glabra*	b	a	NS	b
UCT 73	*C. genistoides*	f	j	a	e
LMG 6465^T	*Azorhizobium caulinodans*	A	U	J	D
USDA 4892^T	*Azorhizobium caulinodans*	A	U	J	D
USDA 76^T	*Bradyrhizobium elkanii*	I	L	R	A
LMG 6138^T	*Bradyrhizobium japonicum*	K	C	g	O
USDA 6^T	*Bradyrhizobium japonicum*	K	S	g	F
LMG 8319	*Bradyrhizobium* sp.	J	M	g	N
STM 815^T	*Burkholderia phymatum*	o	f	i	b
STM 678^T	*Burkholderia tuberum*	f	e	a	d
LMG 14989^T	*Mesorhizobium ciceri*	V	Q	C	X

Table 6.6: continued

Isolate number	Host plant / species designation	*Alu*I	*Cfo*I	*Hae*III	*Msp*I
USDA 4779^T	*Mesorhizobium huakuii*	P	T	A	H
LMG 14107^T	*Mesorhizobium huakuii*	P	H	A	T
LMG 17827^T	*Rhizobium etli*	R	J	E	U
LMG 6214^T	*Rhizobium galegae*	Q	I	F	S
LMG 9503^T	*Rhizobium tropici*	T	O	D	V
LMG 18864^T	*Sinorhizobium medicae*	D	X	M	J
LMG 6133^T	*Sinorhizobium meliloti*	F	Z	O	L
USDA 4893^T	*Sinorhizobium saheli*	C	W	L	E

NS no site for the restriction enzyme used, product remained uncut

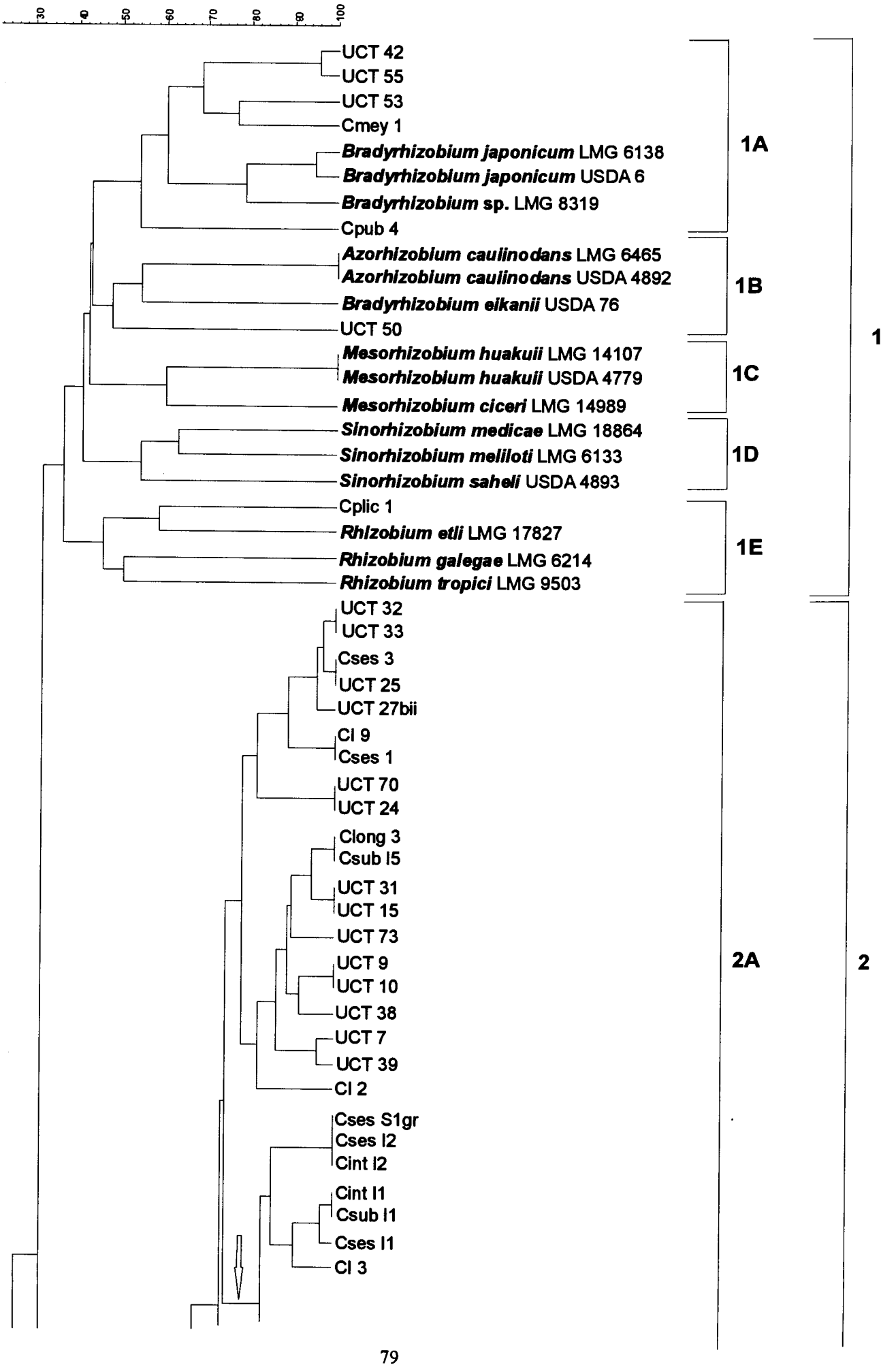
Cluster 1

Cluster 1 was divided into five sub-clusters to aid the discussion of the results. In this cluster all the rhizobial reference strains of the α-*Proteobacteria* grouped. The cluster contained seven *Cyclopia* isolates, which grouped in cluster 1A, 1B and 1E. In cluster 1C, the *Mesorhizobium* reference strains clustered, while the *Sinorhizobium* strains clustered in 1D.

Cluster 1A

The isolates UCT 42, UCT 53, UCT 55, Cmey1 and Cpub 4 grouped in cluster 1A together with the *Bradyrhizobium japonicum* and *Bradyrhizobium* sp. reference strains. The isolates were isolated from three different *Cyclopia* species: *C. plicata* (UCT 42, UCT 53 and UCT 55), *C. meyeriana* (Cmey 1) and *C. pubescens* (Cpub 4). UCT 42 and UCT 55 have the same restriction profiles for the enzymes *Alu*I, *Cfo*I, *Hae*III and *Msp*I and thus most likely have similar genotypes. The restriction profiles of UCT 53 show that the *Hae*III profile of the isolate is similar to that of UCT 42 and UCT 55. However, the resolution power of *Hae*III and *Msp*I has been found to be lower than that of the other two enzymes. Inspection of the profiles showed that Cmey 1 shares some similar bands with UCT 53, but not similar profiles. Isolate Cpub 4 clustered with a low similarity value (~55%) in cluster 1A. The isolate has different enzyme profiles for all four enzymes, which differ from that of the other isolates in cluster 1A.

Figure 6.1 (next page): **UPGMA dendrogram constructed from the combined restriction profiles of the amplified 16S-23S IGS PCR products generated with the enzymes *Alu*I, *Cfo*I, *Hae*III and *Msp*I. The x-axis shows the correlation values between the isolates and displays similarity values for convenience.**


30
40
50
60
70
80
90
100
UCT 42
UCT 55
UCT 53
Cmey 1
Bradyrhizobium japonicum LMG 6138
Bradyrhizobium japonicum USDA 6
Bradyrhizobium sp. LMG 8319
Cpub 4
Azorhizobium caulinodans LMG 6465
Azorhizobium caulinodans USDA 4892
Bradyrhizobium elkanii USDA 76
UCT 50
Mesorhizobium huakuii LMG 14107
Mesorhizobium huakuii USDA 4779
Mesorhizobium ciceri LMG 14989
Sinorhizobium medicae LMG 18864
Sinorhizobium meliloti LMG 6133
Sinorhizobium saheli USDA 4893
Cplic 1
Rhizobium etli LMG 17827
Rhizobium galegae LMG 6214
Rhizobium tropici LMG 9503
UCT 32
UCT 33
Cses 3
UCT 25
UCT 27bii
CI 9
Cses 1
UCT 70
UCT 24
Clong 3
Csub I5
UCT 31
UCT 15
UCT 73
UCT 9
UCT 10
UCT 38
UCT 7
UCT 39
CI 2
Cses S1gr
Cses I2
Cint I2
Cint I1
Csub I1
Cses I1
CI 3
1A
1B
1C
1D
1E
1
2A
2

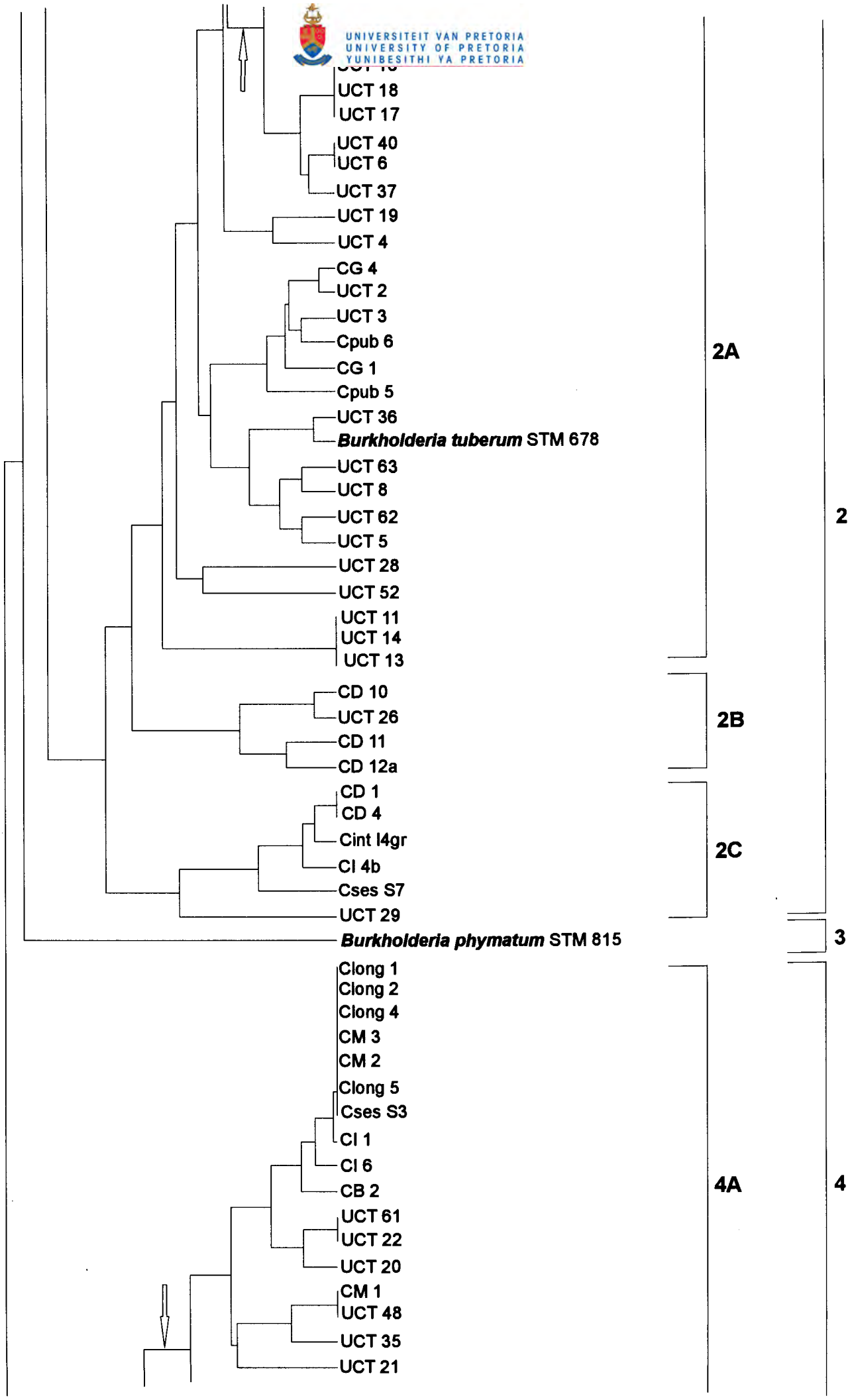


Figure 6.1: continued

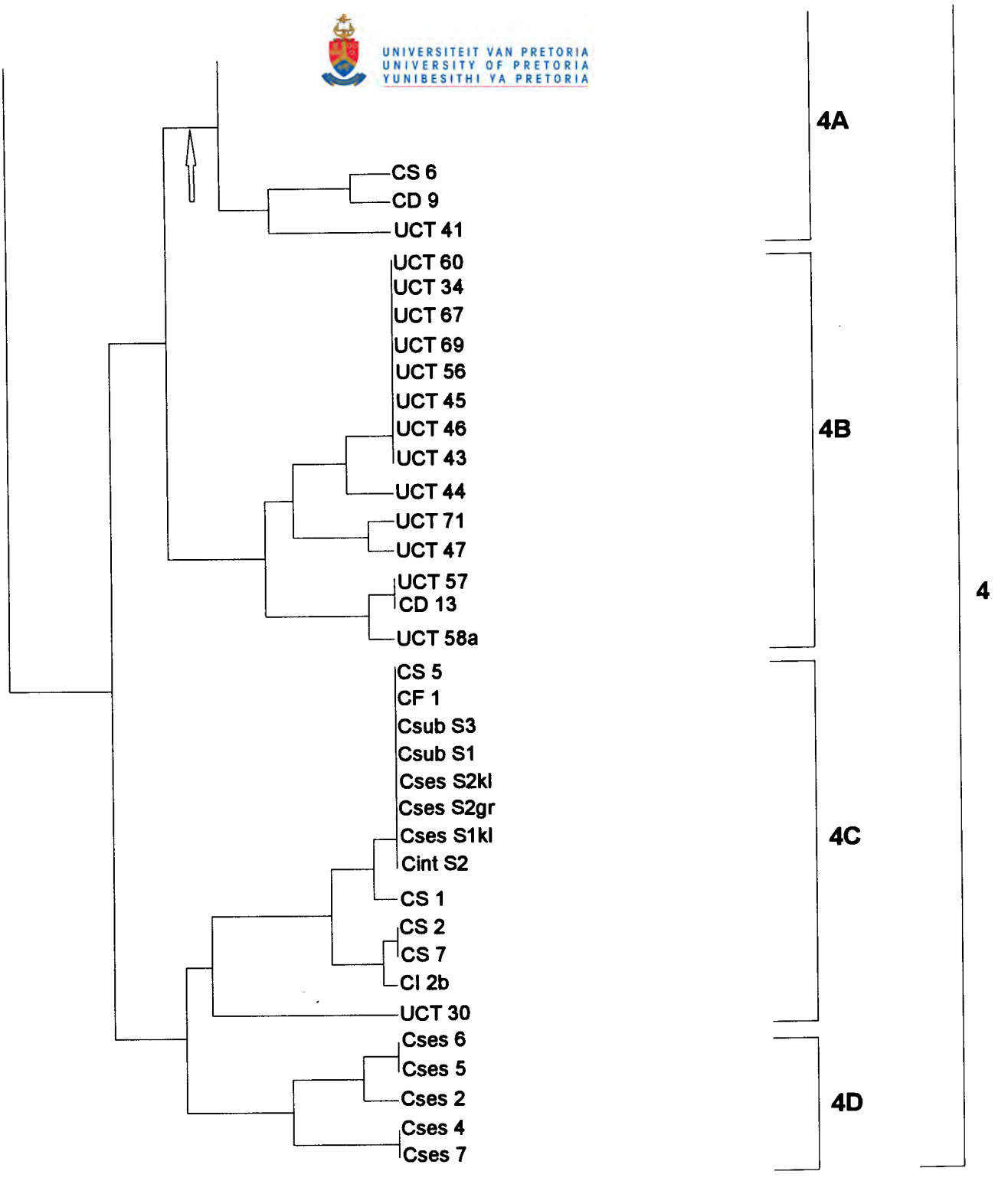


Figure 6.1: continued

Cluster 1B

In cluster 1B the *Azorhizobium caulinodans* and *Bradyrhizobium elkanii* reference strains clustered, though the association between the *A. caulinodans* strains and the *B. elkanii* strain is low (~55%). As an outgroup the *C. sessiliflora* isolate UCT 50 grouped distantly (~ 45%) in cluster 1B.

Cluster 1E

The reference strains of *R. etli*, *R. galegae* and *R. tropici* grouped in cluster 1E forming a *Rhizobium* group. Isolate Cplic 1 from *C. plicata* grouped in the *Rhizobium*-group having a different IGS profile type, but sharing some bands with the other isolates. The closest relative of Cplic 1 in the cluster was *R. etli* LMG 17827^T sharing 60% similarity.

Cluster 2

Cluster 2 consisted of the largest collection of *Cyclopia* strains and the sub-clusters 2A, 2B and 2C can be distinguished based on the IGS-RFLP type. The reference strain *Burkholderia tuberum* STM 678^T grouped in cluster 2. Root-nodulating isolates from ten different *Cyclopia* species grouped in this cluster. Most of the isolates have identical IGS *Alu* I and *Hae*III restriction profiles. However, the relationship of the isolates was resolved with the restriction enzymes *Cfo*I and *Msp*I and 20 different IGS-RFLP types can be distinguished.

Cluster 2A

Cluster 2A represents the largest collection of homogeneous *Cyclopia* isolates from nine different *Cyclopia* species (*C. buxifolia*, *C. galioides*, *C. genistoides*, *C. intermedia*, *C. longifolia*, *C. maculata*, *C. plicata*, *C. pubescens*, *C. sessiliflora* and *C. subternata*). All the isolates of cluster 2A displayed the same restriction profile for *Hae*III. With the exception of UCT 28 (*C. subternata*) and UCT 52 (*C. plicata*) all the isolates had the same *Alu*I restriction enzyme profile. These isolates had unique profiles for *Cfo*I and *Msp*I. The isolates UCT 11, UCT 13 and UCT 14, all from *C. genistoides* shared a *Cfo*I profile with isolates of cluster 4.

Cluster 2B

The isolates of this cluster were isolated from *C. dregeana* and *C. intermedia*. Only UCT 26 was isolated from *C. intermedia*. Cluster 2B is a highly related cluster having a similarity value of 80%. The profiles of most the isolates in this cluster differ from those of the other isolates in cluster 2. However, all the isolates of cluster 2B, except isolates CD 10 and UCT 26 have the common *Hae*III restriction profile of cluster 2.

Cluster 2C

Four *Cyclopia* species (*C. dregeana*, *C. intermedia*, *C. sessiliflora* and *C. subternata*) are represented in this cluster. Except for isolates Cint I4gr, CI 4b and Cses S7, which have the common *Hae*III restriction profile, the other isolates of this cluster have different IGS-RFLP restriction patterns. However, some of the bands are shared with the profiles of the other isolates of cluster 2.

Cluster 3

The reference strain *Burkholderia phymatum* STM 815^T is the only isolate in this cluster. The strain has the common *Alu*I restriction profile of cluster 2 and displayed the same *Msp*I restriction profile as the isolates in cluster 4C, but differed from all the isolates analysed in the restriction profiles obtained with *Cfo*I and *Hae*III.

Cluster 4

No reference strain grouped in cluster 4. Four sub-groups can be distinguished in cluster 4. The isolates of cluster 4 had been isolated from twelve different *Cyclopia* species. All the isolates in cluster 4 lacked an enzyme site for *Hae*III.

Cluster 4A

Cluster 4A represents isolates from nine different *Cyclopia* spp. (*C. buxifolia*, *C. dregeana*, *C. genistoides*, *C. glabra*, *C. intermedia*, *C. longifolia*, *C. maculata*, *C. plicata* and *C. subternata*). All the isolates of cluster 4A lacked an enzyme site for *Msp*I in addition to

having no *Hae*III restriction site. The isolates formed a relatively homogeneous group where the genotypic differences of the isolates could be resolved with the aid of *Alu*I and/or *Cfo*I in some instances.

Cluster 4B

All the isolates in cluster 4B have the same *Msp*I and *Cfo*I profiles in addition to the common characteristic of cluster 4, the absence of any *Hae*III restriction sites. Three different genotypes could be distinguished in this cluster based on the profiles created with *Alu*I. The isolates were isolated from four different *Cyclopia* species (*C. dregeana*, *C. glabra*, *C. meyeriana* and *C. subternata*).

Cluster 4C

Most of the isolates in cluster 4C displayed the same *Alu*I restriction profile and all shared the characteristic lack of a cleavage site for *Hae*III. Two genotypes could be resolved in this group based on the profiles created with the enzymes *Cfo*I and *Msp*I. Isolate UCT 30 (*C. sessiliflora*) which clustered on a separate branch in the cluster displayed different restriction profiles for *Alu*I, *Cfo*I and *Msp*I from that of the other isolates in cluster 4C, though some bands are shared with the other profiles. The isolates in this cluster were isolated from four *Cyclopia* species (*C. falcata*, *C. intermedia*, *C. sessiliflora* and *C. subternata*).

Cluster 4D

Cluster 4D is a homogeneous collection of isolates from *Cyclopia sessiliflora*. The isolates displayed the same profiles for all four enzymes and the differences shown on the dendrogram could be attributed to different size IGS amplification products. The isolates had the same *Alu*I profile as the isolates in cluster 4C, but differed in the profiles generated with *Cfo*I and *Msp*I.

16S PCR

The primers fD1 and rP2 (Weisburg *et al.*, 1991) were able to amplify the 16S rDNA gene of the selected isolates (Table 6.7). The size of the amplification products of the 16S rDNA gene were in the range of approximately 1500 bp (results not shown).

Table 6.7 ***Cyclopia*** **isolates included in the determination of the partial 16S sequence data.**

Isolate	GenBank Accesion number	IGS cluster	Isolate	GenBank Accesion number	IGS cluster
CB 2	AY178059	4A	Cses 5	AY178106	4D
CD 1	AY178094	2C	UCT 11	AY178107	2A
CD 10	AY178083	2B	UCT 15	AY178068	2A
CD 12a	AY178096	2B	UCT 2	AY178073	2A
CD 13	AY178095	4B	UCT 21	AY178057	4A
CD 9	AY178076	4A	UCT 27bii	AY178084	2A
CG 4	AY178097	2A	UCT 28	AY178085	2A
CI 1	AY178060	4A	UCT 29	AY178062	2C
CI 2	AY178069	2A	UCT 30	AY178067	4C
CI 3	AY178072	2A	UCT 31	AY178074	2A
CI 4b	AY178098	2C	UCT 34	AY178056	4B
CI 6	AY178099	4A	UCT 42	AY178077	1A
CI 9	AY178100	2A	UCT 43	AY178055	4B
Clong 1	AY178061	4A	UCT 50	AY178082	1B
Clong 3	AY178070	2A	UCT 52	AY178086	2A
CM 1	AY178058	4A	UCT 53	AY178078	1A
Cmey 1	AY178079	1A	UCT 56	AY178054	4B
Cplic 1	AY178081	1E	UCT 57	AY178087	4B
Cpub 4	AY178080	1A	UCT 58a	AY178088	4B
Cpub 5	AY178101	2A	UCT 61	AY178089	4A
Cpub 6	AY178071	2A	UCT 62	AY178090	2A
CS 2	AY178065	4C	UCT 63	AY178092	2A
CS 3	AY178102	2A	UCT 67	AY178091	4B
CS 6	AY178066	4A	UCT 70	AY178075	2A
Cses 1	AY178103	2A	UCT 71	AY178064	4B
Cses 2	AY178104	4D	UCT 73	AY178093	2A
Cses 3	AY178105	2A	UCT 8	AY178108	2A
Cses 4	AY178063	4D			

16S rDNA sequence analysis

The sequencing reactions conducted with the internal forward primer 16SRNAII-S (Kuhnert *et al.*, 1996) and the reverse primer rP2SHRT (Weisburg *et al.*, 1991) were able to give an unambiguous DNA sequence for each isolate of approximately 700 bp. The last part of each strand had ambiguous positions, since the sequencer had problems to distinguish the correct signal. However, the ambiguous positions could be resolved using the other strand. The

sequences were deposited in the GenBank database. The relevant accession numbers and IGS cluster type can be seen in Table 6.7 for the *Cyclopia* isolates included in this study.

Phylogenetic relationship of the *Cyclopia* isolates within the α- and β-*Proteobacteria*

In order to simplify the dendrogram only the relevant rhizobial reference strains and several *Burkholderia* species were included in the analysis to reveal the possible affinities of the *Cyclopia* isolates. All the sequences of the reference strains used were edited to include the same part of the 16S rDNA gene in the sequence analysis.

The tree reconstructed with the partial 16S rDNA sequence data revealed two prominent lineages corresponding to the α- and β- subclass of the *Proteobacteria*. The isolates Cplic 1 (*C. plicata*), UCT 50 (*C. sessiliflora*), UCT 53 (*C. plicata*), Cmey 1 (*C. meyeriana*), UCT 42 (*C. galioides*) and Cpub 4 (*C. pubescens*) grouped in the α-*Proteobacteria* cluster. All the other *Cyclopia* isolates belonged to the β-*Proteobacteria*.

Isolates Cplic 1 and UCT 50 clustered in the *R. tropici-Agrobacterium rhizogenes* branch within the *Rhizobium* lineage. Within the *Bradyrhizobium* lineage the three isolates UCT 53 (*C. plicata*), Cmey1 (*C. meyeriana*) and UCT 42 (*C. galioides*) grouped. The closest neighbour of isolate UCT 53 is *B. japonicum* LMG 6138^T, while Cmey 1 formed a separate branch and isolate UCT 42 is closer related to *B. liaoningense*. Comparison of the sequence data of Cpub 4 (*C. pubescens*) with that of the data available in GenBank using the BLAST algorithm (Altschul *et al.*, 1990) revealed that the sequence is most related to "*Kaistia adipata*", which led to the inclusion of this isolate in the sequence analysis. The phylogenetic tree revealed that *Kaistia adipata* is the closest neighbour of isolate Cpub 4 sharing 96.9% sequence similarity.

The *Cyclopia* isolates belonged to six different clusters in the β-*Proteobacteria* (see Fig. 6.2). All the isolates in cluster I corresponded to IGS RFLP cluster 2. Isolates of cluster II belonged to IGS cluster 2C and cluster 4D. The isolates from the different IGS clusters grouped on two separate branches in cluster II. In cluster III, all the isolates displayed the genotype of IGS cluster 4. Isolate CS 2 (*C. subternata*) belonged to cluster IV. Two isolates

Figure 6.2 (next page): **Unrooted neighbour-joining tree reconstructed from partial 16S rDNA sequence data to show the phylogenetic relationships between the *Cyclopia* isolates and some reference strains of the α- and β-*Proteobacteria*. Horizontal branch lengths are proportional to the phylogenetic distances, while the vertical branches are non-informative. The scale bar indicates 2% nucleotide difference and bootstrap values higher than 600 are indicated. Abbreviations: *B.* = *Burkholderia*, *M.* = *Mesorhizobium*, *R.* = *Rhizobium* and *S.* = *Sinorhizobium***

UCT 71 (*C. glabra*) and CS 6 (*C. subternata*) both from IGS cluster 4, formed cluster V, a separate cluster lacking any reference isolates. Another isolate from IGS cluster 4 CD 9 (*C. dregeana*) grouped in cluster VI. The tree was simplified by excluding the sequences of all the α-*Proteobacteria* in a separate analysis and thus enabling a better understanding of the phylogenetic position of the *Cyclopia* isolates within the β-*Proteobacteria*, specifically the genus *Burkholderia*.

Phylogenetic relationship of the *Cyclopia* isolates within the β-*Proteobacteria*

The *Cyclopia* isolates belonged to the same six different clusters (see Fig. 6.3) in the simplified tree as seen in the more complex tree (see Fig. 6.2). Clusters I and II shared a common ancestor and more distantly, a common ancestor was shared by clusters I, II and III. Most of the *Cyclopia* isolates belonged to these clusters. In clusters IV, V and VI only four *Cyclopia* isolates grouped.

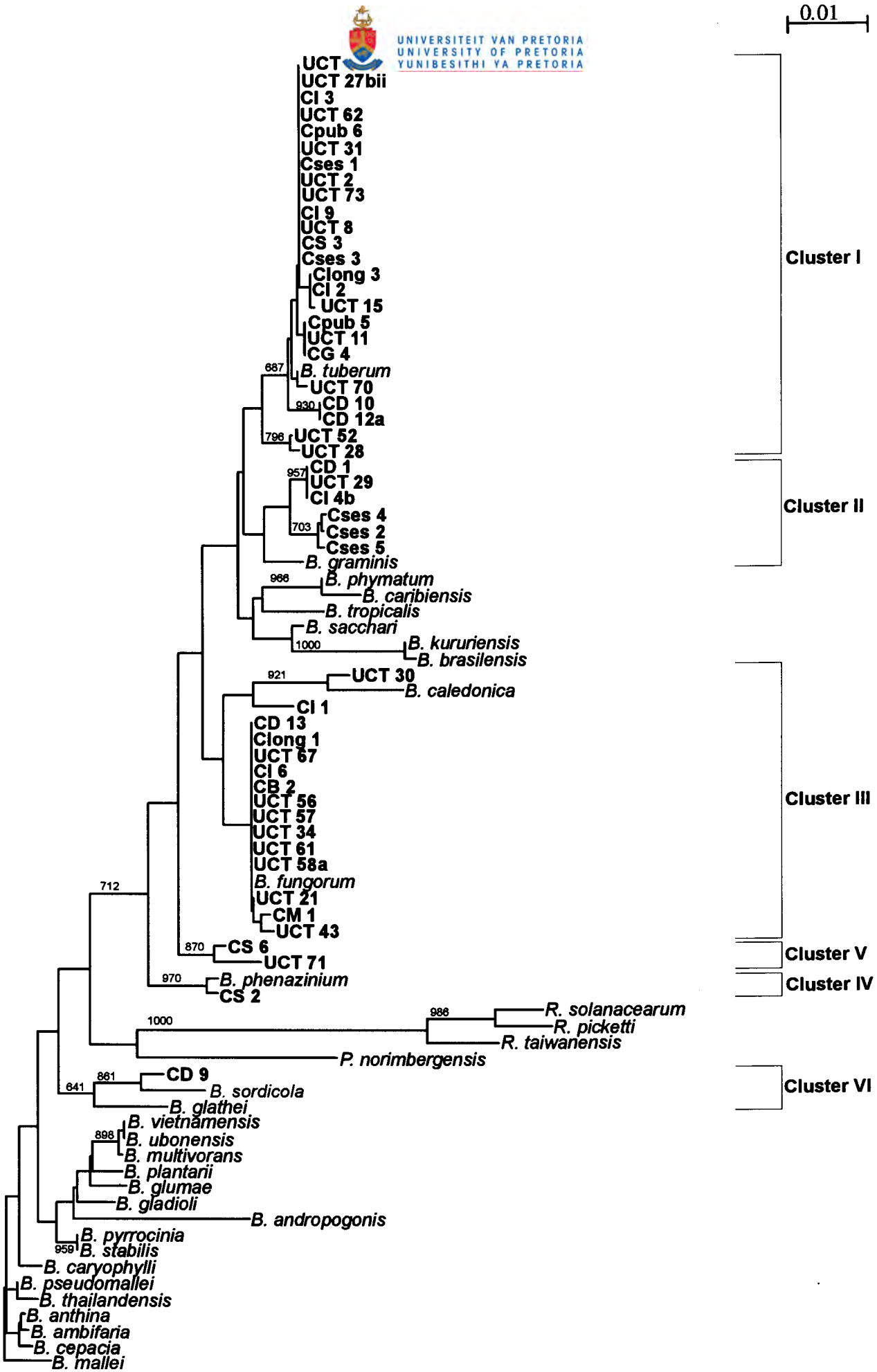
Cluster I

Isolates from nine *Cyclopia* species (*C. dregeana*, *C. genistoides*, *C. intermedia*, *C. longifolia*, *C. maculata*, *C. plicata*, *C. pubescens*, *C. sessiliflora* and *C. subternata*) belonged to this cluster. The cluster is a highly related group of isolates sharing high 16S rDNA sequence similarity values. A single reference strain, *Burkholderia tuberum* STM 678^T was included in this cluster. Isolates CD 10 and CD 11a (both from *C. dregeana*) formed a separate branch in cluster I sharing 99.5% sequence similarity with *B. tuberum*. Another separate branch was formed by the isolates UCT 52 (*C. plicata*) and UCT 28 (*C. subternata*) sharing 99.2% and 99.1% sequence similarity respectively with *B. tuberum*.

Cluster II

Isolates from two different IGS clusters belonged to cluster II. The isolates from IGS cluster 2C and 4D grouped on separate branches in cluster II. Sequence similarities between the isolates of the two branches ranged from 99.3%-99.4%. *Burkholderia graminis* AUS 35, a rhizosphere organism was the reference strain, which belonged to this cluster. However, *B. graminis* grouped as an outgroup of the cluster. *B. graminis* was closer related to the isolates

Figure 6.3 (next page): **Unrooted neighbour-joining tree reconstructed from the partial 16S rDNA sequence data of the *Cyclopia* isolates and some reference strains of the genera *Burkholderia*, *Pandoraea* and *Ralstonia* to show the phylogenetic relationship of the *Cyclopia* isolates within the β-*Proteobacteria*. Branch lengths reflect phylogenetic distances between the isolates, while the vertical branches are non-informative. The scale bar shows 1% sequence difference. Bootstrap probabilities higher than 600 are indicated at the respective nodes. Numbering of clusters as found in Fig. 6.2 were retained in this tree.**



CD1 (*C. dregeana*), UCT 29 (*C. sessiliflora*) and CI 4b (*C. intermedia*) than to the other isolates sharing 99% sequence similarity with the previously mentioned isolates.

Cluster III

Three lineages corresponding to two *Burkholderia* species, *B. fungorum* LMG 16225^T and *B. caledonica* LMG 19076^T can be seen. Isolate CI 1 (*C. intermedia*) formed the third lineage, without a reference strain. The isolate shared 97.6% and 98.7% sequence similarity with *B. caledonica* and *B. fungorum* respectively. The closest neighbour of isolate UCT 30 (*C. sessiliflora*) was *B. caledonica* sharing 98.8% sequence similarity. Isolates collected from nine *Cyclopia* species (*C. buxifolia*, *C. dregeana*, *C. genistoides*, *C. glabra*, *C. intermedia*, *C. longifolia*, *C. maculata*, *C. meyeriana* and *C. subternata*) belonged to the *B. fungorum* lineage. These isolates shared high sequence similarity values ranging from 99.7%-100%.

Cluster IV

Burkholderia phenazinium LMG 2247^T, an isolate obtained from soil enriched with threonine and isolate CS 2 (*C. subternata*) grouped in cluster IV sharing 99.7% sequence similarity. The closest neighbours of these strains were the two *Cyclopia* isolates of cluster V.

Cluster V

In cluster V the two *Cyclopia* isolates CS 6 (*C. subternata*) and UCT 71 (*C. glabra*) grouped. These isolates shared 99.3% sequence similarity.

Cluster VI

Isolate CD 9 (*C. dregeana*) belonged to cluster VI. The reference strains *B. sordicola* SNU 020123 and *B. glathei* LMG 14190^T clustered in this group. *B. sordicola* and *B. glathei* shared 98.9% and 98.2% sequence similarity with isolate CD 9 respectively.

DISCUSSION

The root-nodulating rhizobial reference species belonged to several distinct lineages in the α-*Proteobacteria*. The genera *Azorhizobium* and *Bradyrhizobium* were more related to each other than to the genera *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* as was also found by other researchers (De Lajudie *et al.*, 1998a; Velázquez *et al.*, 1998; Wang *et al.*, 1999b). *Methylobacterium nodulans* and *Devosia neptuniae*, both new α-*Proteobacteria* root-nodulating species formed separate lineages unrelated to the other root-nodulating genera. The inclusion of reference strains from other genera of the α-*Proteobacteria*, as well as the use of full-length or near full-length sequences, in the comparative sequence analysis would enhance the resolution of the relationships of the α-*Proteobacteria* rhizobial species and the other α-*Proteobacteria* species.

Only seven *Cyclopia* isolates belonged to the α-*Proteobacteria*, one of which (Cpub 4) could not be identified yet, but according to BLAST results (Altschul *et al.*, 1990) and comparative sequence analysis, the closest relative was "*Kaistia adipata*". Two isolates UCT 50 and Cplic 1 were found related to the *Rhizobium tropici-Agrobacterium rhizogenes* branch based on sequence data. Several researchers found this close association of *R. tropici* and *A. rhizogenes* (Chen *et al.*, 1997; Terefework *et al.*, 1998). In the IGS-RFLP study, UCT 50 grouped in the same cluster as *Azorhizobium caulinodans* and *B. elkanii*. This might reflect the inability of IGS-RFLP to differentiate between genera. Willems *et al.* (2001b) reported the inability of the IGS sequence analysis to distinguish between genera, since the strains of the genera *Nitrobacter* and *Blastobacter* grouped in the genus *Bradyrhizobium*. Based on the comparative sequence analysis UCT 50 and isolate Cplic 1 (*C. plicata*) is most probably *R. tropici* strains. Isolate UCT 53 is most related to *Bradyrhizobium japonicum*, while the identity of UCT 42 and Cmey 1 is not that clear from the 16S rDNA tree. However, these strains clearly belonged to the genus *Bradyrhizobium*.

The soil from which the *Cyclopia* isolates have been collected was very acidic (J. Bloem, personal communication). The finding of mainly *Bradyrhizobium* strains would be expected, since the slow-growing strains are better adapted to these environmental conditions (Graham *et al.*, 1994). *Rhizobium tropici* strains are also more acid-tolerant than other fast-growing rhizobial species (Martínez-Romero *et al.*, 1991; Graham *et al.*, 1994). The higher acid

tolerance of *Bradyrhizobium* strains and *Rhizobium tropici* strains would thus explain the finding of these bacteria in the root-nodules of some *Cyclopia* plants. Lafay and Burdon (1998) also reported the isolation of these acid-tolerant rhizobia from plants growing in areas with low pH soils in Australia. It can thus be assumed that the *Burkholderia* isolates also have high acid-tolerance.

Based on the results of both techniques the identity of the *Cyclopia* isolates of 16S rDNA cluster I could be proposed as strains of the species *Burkholderia tuberum* (Moulin *et al*, 2001; Vandamme *et al*, 2002). This would also include all the isolates, which grouped in IGS cluster 2A. It is clear from the IGS-RFLP dendrogram that this collection of strains forms a highly homogeneous group. However, isolates CD 10, CD 12a, UCT 28 and UCT 52 might be members of two additional *Burkholderia* species closely related to *B. tuberum* based on the separate branches that they formed in cluster I.

The identity of the isolates in cluster II is not clear from the tree reconstructed from the 16S sequence data. The significance of the branching in the same lineage as *Burkholderia graminis*, a species described by Viallard *et al.* (1998), was not supported with a bootstrap value higher than 50%. The addition of more similar sequences or the use of longer sequences might change the association.

All the isolates in cluster III, except isolates UCT 30 and CI 1 could be strains belonging to the species *Burkholderia fungorum* described by Coenye *et al.* (2001). However, the *B. fungorum* branch was supported by only 55% of the 1000 replicates generated with the bootstrap analysis. The use of partial sequence data might hinder the differentiation between closely related isolates, since it does not reflect true relationships. Thus to draw conclusions on the species affiliation of these isolates full-length 16S sequence analysis should be done.

The significance of the branching of clusters IV and V, as determined with bootstrap is 97% and 87% respectively. The high significance level and sequence similarity would support the identity of isolate CS 2 as a possible strain of *Burkholderia phenazinium* (Viallard *et al.*, 1998). No definite conclusions can however be made on the identity of isolates CS 6 and UCT 71.

The branching pattern of cluster VI was found in 64% of the 1000 bootstrap generated replicates, but the branch leading to CD 9 and *B. sordicola* was found highly significant (86%). The use of full-length sequence analysis would be able to unequivocally determine the correct species affiliation of isolate CD 9.

None of the isolates was related to the other root-nodulating *Burkholderia* species, *Burkholderia phymatum* (Moulin *et al.*, 2001 and Vandamme *et al.*, 2002). Based on the 16S rDNA sequence data, the closest neighbour of *B. phymatum* was *B. caribiensis*, as previously found by Vandamme *et al.* (2002). Vandamme *et al.* (2002) also identified root-nodulating isolates from tropical legumes as members of the species *B. caribiensis* and *B. cepacia* genomovar VI.

Other researchers have shown that IGS-RFLP analysis is a useful method for determining the diversity of bacterial populations (Laguerre *et al.*, 1996; Vinuesa *et al.*, 1998; Diouf *et al.*, 2000, Doignon-Bourcier *et al.*, 2000). In this study, it has been found that IGS-RFLP analysis was an easy and reproducible method for the diversity determination of the *Cyclopia* isolates, even showing intraspecific differences between the strains. The partial 16S rDNA sequencing analysis corroborated the results of the IGS-RFLP analysis. In a study conducted by Willems *et al.* (2001a), the researchers found an agreement between the clustering of the same *Bradyrhizobium* strains obtained with IGS sequence analysis and the clustering obtained from IGS-RFLP analysis as done by Doignon-Bourcier *et al.* (2000). Willems *et al.* (2001a) proved in the study of *Bradyrhizobium* strains that the groupings obtained with IGS sequence analysis and AFLP (amplified fragment length polymorphism) analysis correlated with data generated with DNA homology analysis. It might thus be possible to draw conclusions on different genomic species from the IGS-RFLP dendrogram if a threshold value for species delineation could be determined.

The different size ranges of the amplified IGS products were expected, since the IGS products of several bacteria vary in length due to the insertion of tRNA genes (Gürtler and Stanisich, 1996). The rRNA operon is also present in multiple copies and the insertion of tRNA genes could explain the length differences between IGS products of the same strain (Gürtler and Stanisich, 1996; Laguerre *et al.*, 1996; LeBlond-Bourget *et al.*, 1996).

CHAPTER 7

PHYLOGENETIC, SYMBIOTIC AND PHENOTYPIC CHARACTERISATION OF SOME *BURKHOLDERIA* SPP. ISOLATES

ABSTRACT

The phylogenetic position of some *Burkholderia* strains isolated from different *Cyclopia* species was determined using near full-length 16S rDNA sequencing. The data showed the identity of several isolates as *B. tuberum*. Several possible novel *Burkholderia* species were found. However, DNA homology studies remain to be done to confirm and delineate the novel species. All the *nodA* sequences of the isolates displayed high sequence similarity. The *nodA* sequence of *B. tuberum* isolated from *Aspalathus carnosa* shared highest sequence similarity with the *Cyclopia* isolates. The phenotypic study confirmed the isolates as members of the genus *Burkholderia* and thus clearly different from the rhizobial genera included in the analysis.

Keywords: *Burkholderia*, 16S rDNA sequencing, *nodA* sequencing, substrate utilisation patterns

INTRODUCTION

The study of more legume hosts from diverse environmental conditions opened the door for new discoveries. The understanding of the rhizobium-legume symbiosis changed significantly in recent years. The description of β-*Proteobacteria* (Chen *et al.*, 2001; Moulin *et al.*, 2001 and Vandamme *et al.*, 2002), methylotrophic bacteria (Sy *et al.*, 2001; Jaftha *et al.*, 2002), *Devosia neptuniae* (Rivas *et al.*, 2002; Rivas *et al.*, 2003) and the budding bacteria *Blastobacter denitrificans* capable of root-nodulation changed the rhizobium taxonomy.

The genus *Burkholderia* contains plant and animal pathogens (Brett *et al.*, 1998), obligate endosymbionts of Rubiaceae and Myrsinaceae hosts (Van Oevelen *et al.*, 2002), an endosymbiont of *Gigaspora margarita* (arbuscular mycorrhizal fungus) [Minerdi *et al.*, 2001] strains capable of bioremediation (Fain and Haddock, 2001), biocontrol (Trân Van *et al.*, 2000; Peix *et al.*, 2001) and plant growth promotion (Trân Van *et al.*, 2000; Peix *et al.*, 2001; Ciccillo *et al.*, 2002). Strains of the genus can fix nitrogen (Gillis *et al.*, 1995; Minerdi *et al.*, 2001). *B. vietnamensis* was described for nitrogen-fixing isolates from rice in Vietnam (Gillis *et al.*, 1995). *B. brasilensis*, *B. kururiensis*, *B. tropicalis* (Marin *et al.*, 2003) in addition to *B. phymatum* and *B. tuberum* (Moulin *et al.*, 2001; Vandamme *et al.*, 2002) are nitrogen-fixers.

The 16S rDNA molecule can be used to identify and determine the phylogenetic position of isolates. The use of full-length sequences is essential for phylogenetic conclusions (Ludwig *et al.*, 1998). The limitation of 16S sequence data is the relatively conserved nature of the molecule, since closely related species cannot be differentiated with 16S data (Vandamme *et al.*, 1996; LeBlond-Bourget *et al.*, 1996; Rosselló-Mora and Amann, 2001; Stackebrandt *et al.*, 2002). However, this technique has been widely used to identify and determine the phylogenetic position of isolates. Lafay and Burdon (1998) used the technique to rapidly identify novel rhizobial isolates from scrubby legumes in Southeastern Australia, while Terefework *et al.* (1998) used the technique to determine the phylogenetic position of *Rhizobium galegae* in the *Rhizobiaceae*.

Rhizobia recognise specific signals from legumes, which activate the regulatory NodD proteins and induce the *nod* gene expression. The common genes (*nodABC*) are involved in the formation of the backbone of the Nod-factor (lipo-chitooligosaccharide signal), which induces specific infection and nodulation in legumes. NodA is an acyltransferase, which

transfers a fatty acyl chain to the acetyl-free C-2 carbon of the non-reducing end of the oligosaccharide molecule (Perret *et al.*, 2000; Hirsch *et al.*, 2001). NodA is also host-specific, since it transfers specific acyl chains (Ritsema *et al.*, 1996). The *nodA* gene is found as a single copy in rhizobia. The phylogenetic trees reconstructed from the nod genes, *nodA*, *nodB*, *nodC* and *nodD*, agree with each other, but differ from that of the 16S rRNA (Haukka *et al.*, 1998). The study of the *nodA* gene is a rapid technique to determine the host range of isolates (Haukka *et al.*, 1998; Ba *et al.*, 2002).

Numerical taxonomy provides descriptive phenotypic information about strains (Vandamme *et al.*, 1996). Several authors used this technique to characterise strains and differentiate between different phenotypes of a strain. McInroy *et al.* (1999) used Biolog™ and partial 16S rRNA sequencing to characterise rhizobia isolated from African acacias and other tropical woody legumes. In a study on the diversification of *Pseudomonas corrugata* 2140 Barnett *et al.* (1999) used Biolog™ GN microplates (Biolog Inc., Hayward, California, USA) to identify new phenotypes.

MATERIALS AND METHODS

Bacterial strains used

The strains used in this study were received from the Agricultural Research Council's-Plant Protection Research Institute (Private Bag X134, Pretoria, 0001, South Africa) and the Botany Department of the University of Cape Town (Rondebosch, 7701, Cape Town, South Africa) [Table 7.1]. All the *Cyclopia* isolates included in this study were analysed with IGS-RFLP analysis and partial 16S rDNA sequence analysis. The reference strains of *Burkholderia tuberum* and *B. phymatum* used in this study were obtained from the culture collection of the Laboratoire des Symbioses Tropicales et Méditerranéennes (STM), Montpellier, France. All the sequences of the reference strains used in the 16S rDNA and *nodA* sequence analysis were obtained from the Genbank database of the National Centre for Biotechnology (NCBI) [website address: www.ncbi.nlm.nih.gov/Genbank/].

Table 7.1: List of isolates from *Cyclopia* spp. included in the 16S rDNA and partial *nodA* sequence analysis

Isolate	Host species	Locality	16S rDNA accession number[1]	*nodA* accession number[1]
CB 2	*C. buxifolia*	Helderberg, Somerset-West	AY178059	AY189248
CD 9	*C. dregeana*	Du Toitskloof, Paarl	AY178076	AY189250
CI 1	*C. intermedia*	Dennehoek, Joubertina	AY178060	AY189253
CI 2	*C. intermedia*	Dennehoek, Joubertina	AY178069	AY189229
CI 3	*C. intermedia*	Dennehoek, Joubertina	AY178072	AY189254
Clong 1	*C. longifolia*	Thornhill, Humansdorp	AY178061	AY189228
Clong 3	*C. longifolia*	Thornhill, Humansdorp	AY178070	AY189273
CM 1	*C. maculata*	Paarlberg, Paarl	AY178058	AY189256
Cpub 6	*C. pubescens*	Next to N1, Port Elizabeth	AY178071	AY189274
CS 2	*C. subternata*	Dennehoek, Joubertina	AY178065	AY189259
CS 6	*C. subternata*	Dennehoek, Joubertina	AY178066	AY189261
Cses 4	*C. sessiliflora*	Plattekloof, Heidelberg	AY178063	AY189230
UCT 15	*C. genistoides*	Constantiaberg	AY178068	AY189275
UCT 2	*C. genistoides*	Rein's Farms	AY178073	AY189267
UCT 21	*C. genistoides*	Paardeberg	AY178057	AY189276
UCT 29	*C. sessiliflora*	Callie's farm, Heidelberg	AY178062	AY189266
UCT 30	*C. sessiliflora*	Callie's farm, Heidelberg	AY178067	AY189268
UCT 31	*C. sessiliflora*	Grootvadersbosch	AY178074	AY189240
UCT 34	*C. glabra*	Matroosberg	AY178056	AY189241
UCT 43	*C. meyeriana*	Hottentots Holland mountains	AY178055	AY189271
UCT 56	*C. meyeriana*	Hottentots Holland mountains	AY178054	AY189245
UCT 70	*C. maculata*	Jonkershoek	AY178075	AY189277
UCT 71	*C. glabra*	unknown	AY178064	AY189278

[1] Genbank (www.ncbi.nlm.nih.gov/Genbank/)

Maintenance of bacterial 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 5-7 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.

Extraction of genomic DNA

A modified method for proteinase-K (Roche Molecular Biochemicals) treated cells as described by Laguerre *et al.* (1997) was used. A pure fresh culture of each strain was streaked

on a tryptone yeast (TY) agar slant [0.5% (m/v) tryptone (Difco), 0.3% (m/v) yeast extract (Biolab), 0.13% (m/v) $CaCl_2.6H_2O$ (UniLab), 1.5% (m/v) bacteriological agar] in a screw-cap tube. The strains were incubated at 28 °C and checked for sufficient growth. Sterile 4.5 ml dH_2O was added to the slant growth to harvest the cells. An inoculation loop was used to aid the release of cells clinging to the agar. The volume of the water added was adjusted according to the amount of growth. Less water was used if the growth was poor and *vice versa*. The cell-suspension was collected in a clear plastic tube and vortexed to ensure a uniform suspension. The absorbancy of the suspension was measured with dH_2O as the spectrophotometric blank at 620 nm. A formula was used to determine the volume of the cells to be treated further. The volume to be used in ml was equal to 0.2 divided by the abosorbancy at 620 nm. Two tubes of the same strain were filled with the appropriate volume of cells and centrifuged at 13 000 *g* for 5 min at 4 °C. The supernatant was discarded and the excess media blotted dry. One of the tubes was stored at -20 °C for future use. In the second tube, 100 µl ddH_2O, 100 µl Tris-HCl (10 mM, pH 8.2) and 10 µl proteinase-K (15 mg/ml) (Roche Molecular Biochemicals) were added to the cell pellet. The mixture was incubated at 55 °C overnight. In order to inactivate the proteinase-K the mixture was boiled for 10 minutes. The cell lysates were stored at -20 °C until needed.

Amplification of the 16S rDNA and the partial *nodA* gene

Amplification of the 16S rDNA gene of selected strains (Table 7.1) was performed with the primers fD1 and rP2 (Table 7.2) as described by Weisburg *et al.* (1991). The linker sequences of the primers were not included in the primer synthesis, since no cloning reactions were anticipated. These shorter primers were thus designated fD1SHRT and rP2SHRT. The PCR mixture of each strain contained: 5 µl of the cell lysate, 50 pmol of each primer, 250 µM of each dNTP, 1.5 mM $MgCl_2$ and 0.5 U Gold Taq DNA polymerase (Southern Cross Biotechnology) in a 50 µl reaction volume. The amplification reactions were performed on a Perkin Elmer GeneAmp PCR System 2400 thermocycler using the following thermal profile: initial denaturation at 95 °C for 3 minutes, followed by 35 cycles of denaturation at 94 °C for 30 seconds, annealing at 55 °C for 30 seconds and extension at 72 °C for 1 minute followed by a final extension step at 72 °C for 5 minutes. Aliquots (5 µl) of the amplified products were evaluated with horizontal agarose gel electrophoresis (Sambrook *et al.*, 1989) using 0.9% (m/v) agarose gels (Promega) in a 1X TAE buffer (40 mM Tris-HCl, 20 mM NaOAc

and 1 mM EDTA pH 8.5) stained with ethidium bromide (10 mg/ml) [results not shown]. Molecular marker VI (Roche Molecular Biochemicals) was included on each gel as a standard lane.

The 16S PCR product of each strain was purified to remove any traces of unincorporated dNTPs, primers, etc. which could negatively influence the 16S sequencing reaction. The products were purified using a Qiagen QIAquick PCR Purification kit (Southern Cross Biotechnology). Purification reactions were done as prescribed by the manufacturer. The concentration and purity of each purification reaction was verified visually. An aliquot (1 µl) of each purified 16S PCR product was run on 0.9% (m/v) horizontal agarose gels (Promega) [results not shown]. On each gel, a standard marker, molecular marker VI (Roche Molecular Biochemicals) was included.

Table 7.2: Primers used in the amplification and/or sequencing of the 16S rDNA and the nodA.

Primer name*	Primer sequence (5'-3')#	Target gene	Reference
fD1SHRT	5'-AGAGTTTGATCCTGGCTCAG-3'	16S rDNA	Weisburg *et al.*, 1991
rP2SHRT	5'-ACGGCTACCTTGTTACGACTT-3'	16S rDNA	Weisburg *et al.*, 1991
16SRNAII-S	5'-GTGTAGCGGTGAAATGCGTAG-3'	16S rDNA	Kuhnert *et al.*, 1996
16SRNAVI-S	5'-CTACGCATTTCACCGCTACAC-3'	16S rDNA	Kuhnert *et al.*, 1996
NodAunivF145u	5'-TGGGCSGGNGCNAGRCCBGA-3'	nodA	Moulin *et al.*, 2001
NodAR.brad	5'-TCACARCTCKGGCCCGTTCCG-3'	nodA	Moulin *et al.*, 2001

* All the primers were synthesised by Roche Molecular Biochemicals, Mannheim, Germany

\# Abbreviations: B= G/C/T, K= G/T, N= A/G/C/T, R= A/G, S= G/C

The amplification of the *nodA* gene was performed with primer set NodAunivF145u and NodAR.brad (Table 7.2) as used by Moulin *et al.* (2001) in the first report of *Burkholderia* strains capable of root-nodulation. In each 50 µl amplification reaction the following was added: 5 µl of the cell lysate, 50 pmol of each primer, 250 µM of each dNTP, 1.5 mM $MgCl_2$ and 0.5 U Taq DNA polymerase (Southern Cross Biotechnology). The following thermal profile was used: a hot start at 95 °C for 3 minutes, then 35 cycles of denaturation (94 °C for 30 seconds), annealing (55 °C for 45 seconds) and extension (72 °C for 1 minute), followed by a final extension step (72 °C for 5 minutes). The reactions were performed on a Perkin Elmer GeneAmp PCR System 2400 thermocycler. The success of the amplification reactions was checked with horizontal agarose gel electrophoresis (Sambrook *et al.*, 1989) using the method described for 16S rDNA amplification.

The *nodA* amplification products were purified to remove all traces of inhibitors of the sequencing reaction, as well as to ensure the presence of a single product for sequencing. The products were purified with a combined method using the binding buffer of the High Pure Purification PCR kit (Roche Molecular Biochemicals) and the columns and chemicals from the Qiagen QIAquick PCR Purification kit (Southern Cross Biotechnologies). The total volume of the amplification product of each isolate was run on a 0.9% (m/v) agarose gel (Promega) in a 1X TAE buffer (40 mM Tris-HCl, 20 mM NaOAc and 1 mM EDTA pH 8.5) stained with ethidium bromide (10 mg/ml) (Sambrook *et al.*, 1989). The molecular marker VI was included on each gel. The fragment of the correct size was excised from the gel and the weight of the agarose was determined for each isolate. To each fragment, 300 µl binding buffer (Roche Molecular Biochemicals) for each 100 mg of agarose gel was added. The mixture was vortexed to aid in dissolving the agarose. The tubes were incubated at 50-60 °C for 10 minutes, while vortexing the tubes every 2-3 min. After the incubation period, 150 µl of isopropanol (Merck) per 100 mg of agarose was added to each tube. This mixture was added to the Qiagen QIAquick PCR Purification kit columns. The rest of the purification procedure was done as prescribed by the manufacturer. The success of the purification reactions was verified using the method as described for 16S rDNA (results not shown).

Sequence analysis of the 16S rDNA and the partial *nodA* gene

The near full-length sequence of each purified amplified 16S rDNA product of the chosen isolates were determined with the internal forward primer 16SRNAII-S (Kuhnert *et al.*, 1996), the reverse primer rP2SHRT (Weisburg *et al.*, 1991), the internal reverse primer 16SRNAVI-S (Kuhnert *et al.*, 1996) and the forward primer fD1SHRT (Weisburg *et al.*, 1991) using the ABI Prism BigDye™ Terminator Cycle Sequencing Ready Reaction kit (AmpliTaq® DNA Polymerase, FS) (PE Applied Biosystems). In the sequencing reactions of the *nodA* gene the forward primer NodAunivF145u was used. Each 5 µl sequencing reaction contained the following: 2 µl of the ready reaction mix supplied with the kit which contains the dye terminators, dNTP's, AmpliTaq® DNA polymerase, $MgCl_2$ and Tris-HCl buffer pH 9.0; 12.5 pmol primer and approximately 100 ng template DNA. The sequencing reactions were carried out in a Perkin Elmer GeneAmp PCR System 2400 thermocycler with the following thermal profile: an initial denaturation at 96 °C for 5 seconds followed by 25 cycles of denaturation (96 °C for 10 seconds), annealing (50 °C for 5 seconds) and extension (60 °C for

4 minutes). The products were precipitated using the protocol as suggested by the manufacturer. 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 de-ionised 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.

Phylogenetic analysis of the 16S rDNA and *nodA* sequences

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). Both strands were sequenced with the primers used and the strands could be aligned to correct ambiguous positions. The resulting two unambiguous strands were overlapped to give a continuous near full-length sequence for each isolate. The *nodA* sequences were checked visually to see that the peaks and the corresponding nucleotides were correct. The ClustalX programme (Thompson *et al.*, 1997) was used to analyse the edited sequences as well as the reference sequences obtained from GenBank (Table 7.3 and Table 7.4). A distance matrix was constructed by pair-wise alignment of the sequences. 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. A bootstrap confidence analysis was performed on 1000 replicates to determine the reliability of the tree topologies (Felsenstein, 1985). The phylogenetic trees were visualised with NJplot (Perrière and Gouy, 1996).

Table 7.3 Reference strains obtained from GenBank included in the comparative 16S sequence analysis

Reference strain	Strain number	Host plant or relevant characteristics	GenBank[1] Accession number
Burkholderia ambifaria	MVPC 1/4	*B. cepacia* complex	AY028444
Burkholderia andropogonis	ATCC 23061^T	Sorghum (*Sorghum bicolor)*	X67037
Burkholderia anthina	R-4183^T	Rhizosphere soil, *B. cepacia* complex	AJ420880
Burkholderia brasilensis	M130	Plant-associated N_2-fixer	AJ238360
Burkholderia caledonica	LMG 19076^T	Rhizosphere soil	AF215704
Burkholderia caribiensis	LMG 18531^T	Vertisol microaggregates	Y17009
Burkholderia caryophylli	ATCC 25418^T	Carnation (*Dianthus caryophyllus)*	AB021423
Burkholderia cenocepacia	LMG 16656^T	Cystic fibrosis patients, plant associated *B. cepacia* genomovar III	AF148556
Burkholderia cepacia	ATCC 25416^T	Cystic fibrosis patients, B. cepacia genomovar I	AF097530
Burkholderia fungorum	LMG 16225^T	*Phanerochaete chrysosporium*	AF215705
Burkholderia gladioli	ATCC 10248^T	*Gladiolus* sp.	X67038
Burkholderia glathei	LMG 14190^T	Fossil lateritic soil	U96935
Burkholderia glumae	LMG 2196^T	Rice (*Oryza sativa)*	U96931
Burkholderia graminis	AUS 35	Rhizosphere	U96941
Burkholderia hospita	LMG 20598^T	Agricultural soil	AY040365
Burkholderia kururiensis	LMG 19447^T	Trichloroethylene degrader	AB024310
Burkholderia mallei	NCTC 10260	NS	AF110187
Burkholderia multivorans	LMG 13010^T	*B. cepacia* complex	Y18703
Burkholderia phenazinium	LMG 2247^T	Soil enriched with threonine	U96936
Burkholderia phymatum	STM 815^T	*Machaerium lunatum*	AJ302312
Burkholderia plantarii	LMG 9035^T	*Oryza sativa* pathogen	U96933
Burkholderia pseudomallei	V686	Soil	AF093052
Burkholderia pyrrocinia	LMG 14191^T	soil	U96930
Burkholderia sacchari	LMG 19450^T	Soil from sugarcane plantation	AF263278
Burkholderia sordicola	SNU 020123	Associated with white rot fungus *Phanerochaete sordicola*	AF512827
Burkholderia stabilis	LMG 14294^T	Formerly *B. cepacia* complex IV	AF148554
Burkholderia terricola	LMG 20594^T	Agricultural soil	AY040362
Burkholderia thailandensis	E264^T	*Pseudomallei* group	U91838
Burkholderia tropicalis	Ppe8	Plant-associated N_2-fixer	AJ420332
Burkholderia tuberum	STM 678^T	*Aspalathus carnosa*	AJ302311
Burkholderia ubonensis	GTC-P3-415	NS	AB030584
'Burkholderia unamae'	MT1-641^T	maize	AY221956
Burkholderia vietnamensis	LMG 10929^T	N_2-fixer from rice rhizophere	AF097534
Candidatus Burkholderia kirkii	Strain19536779	*Psychotria kirkii* var. *tarambassica*	AF475063
Pandoraea norimbergensis	NS*	Alkaliphilic sulphur oxidiser	Y09879
Ralstonia picketti	MSP 3	Rhizosphere, soil	AB004790
Ralstonia solanacearum	ATCC 11696	*Lycopersicon lycopersicum*	X67036
Ralstonia taiwanensis	LMG 19424^T	*Mimosa pudica*	AF300324

* See footnotes of Table 7.4

Table 7.4 ***NodA*** **sequences obtained from GenBank included in the comparative** ***nodA*** **sequence analysis**

Reference strain	Strain number	Host plant	Accession number GenBank[1]
Azorhizobium caulinodans	ORS 571[T]	*Sesbania rostrata*	L18897
Bradyrhizobium elkanii	USDA 94	NS	U04609
Bradyrhizobium japonicum	110spc4	NS	AF322013
Bradyrhizobium **sp.**	NC92	*Arachis hypogaea*	U33192
Bradyrhizobium **sp.**	WM9	*Lupinus* sp.	AF222753
Bradyrhizobium **sp.**	ANU289	*Parasponia* sp.	X03720
Bradyrhizobium **sp.**	ORS 285	Photosynthetic	AF284858
Bradyrhizobium **sp.**	ORS 287	*Aeschynomene afraspera*	AJ437607
Bradyrhizobium **sp.**	ORS 301	*Aeschynomene americana*	AJ437608
Bradyrhizobium **sp.**	ORS 302	*Aeschynomene pfundii*	AJ437609
Bradyrhizobium **sp.**	ORS 304	*Aeschynomene elaphroxylon*	AJ437610
Bradyrhizobium **sp.**	ORS 309	*Aeschynomene uniflora*	AJ437611
Bradyrhizobium **sp.**	ORS 336	*Aeschynomene afraspera*	AJ437612
Bradyrhizobium **sp.**	ORS 364	*Aeschynomene nilotica*	AJ437613
Burkholderia tuberum	STM 678[T]	*Aspalathus carnosa*	AJ302321
Mesorhizobium ciceri	USDA 3383	*Hedysarum boreale*	AJ250140
Mesorhizobium loti	NS	NS	L06241
Mesorhizobium mediterraneum	USDA 3392	NS	AJ250141
Mesorhizobium plurifarium	ORS 1096	*Acacia tortilis* subsp. *raddiana*	AJ302678
Mesorhizobium **sp.**	BR3804	*Chamaecrista ensiformis*	Z95249
Mesorhizobium **sp.**	DW0366	*Acacia polycantha*	Z95248
Mesorhizobium **sp.**	7653R	*Astragalus sinicus*	AJ249353
Mesorhizobium **sp.**	N33	*Oxytropis arctobia*	U53327
Mesorhizobium tianshanense	USDA 3592	NS	AJ250142
Methylobacterium nodulans	ORS 2060[T]	*Crotalaria podocarpa*	AF266748
Rhizobium etli	CFN 42[T]	*Phaseolus vulgaris*	NC_004041
Rhizobium galegae	HAMBI 1174	*Galega orientalis*	X87578
Rhizobium leguminosarum **bv.** ***phaseoli***	NS	NS	M58625
Rhizobium leguminosarum **bv.** ***trifolii***	ANU843	NS	X03721
Rhizobium leguminosarum **bv.** ***viciae***	NS	NS	Y00548
Rhizobium tropici	CFN 299	*Phaseolus* sp.	X98514
Sinorhizobium arboris	HAMBI 1700	*Acacia senegal*	Z95235
Sinorhizobium fredii	USDA 257	NS	M73699
Sinorhizobium kostiense	HAMBI 1489[T]	*Acacia senegal*	Z95236
Sinorhizobium meliloti	NS	NS	X01649
Sinorhizobium saheli	ORS 609	*Sesbania cannabina*	Z95241
Sinorhizobium **sp.**	NGR234	Broad host range	AE000076
Sinorhizobium **sp.**	BR827	*Leucaena leucocephala*	Z95232
Sinorhizobium **sp.**	BR4007	*Prosopis juliflora*	Z95240
Sinorhizobium **sp.**	M6	*Prosopis* sp.	Z95233
Sinorhizobium **sp.**	ORS 1085	*Acacia tortilis* subsp. *raddiana*	AJ302677
Sinorhizobium terangae	ORS 1009	*Acacia laeta*	Z95237

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

[T] Type strain

NS Not stated

ATCC American Type Culture Collection, Rockville, Maryland, USA

CFN	Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Mexico
HAMBI	Culture Collection of the Department of Applied Chemistry and Microbiology, University of Helsinki, Helsinki, Finland
LMG	BCCM™/LMG Bacteria Collection, Laboratorium voor Microbiologie, University of Gent, Gent, Belgium
NCTC	National Collection of Type Cultures, Central Public Health Laboratory, London, UK
ORS	ORSTOM Collection, Institut Français de Recherche Scientifique pour le Développement en Coopération, Dakar, Senegal
STM	Laboratoire des Symbioses Tropicales et Méditerranéennes, Montpellier, France
USDA	United States Department of Agriculture, *Rhizobium* Culture Collection, Beltsville Agricultural Research Center, Beltsville, MD, USA

Numerical taxonomy

The substrate utilisation patterns of selected *Cyclopia* isolates (Table 7.5) and the two root-nodulating *Burkholderia* species were determined with Biolog GN Microplates™ (Biolog Inc., Hayward, USA). Each plate has 96 wells, comprising a negative control and 95 preselected carbon sources (Appendix A). The method as prescribed by the manufacturer was used to grow the cultures and inoculate the microplates. A positive reaction depends on the reduction of tetrazolium violet to form a purple dye in wells where oxidation of the carbon source takes place. A positive reaction was scored as one (1), while a negative reaction was scored as zero (0). These data were analysed with the simple matching coefficient in the Bionum programme (Applied Maths, Kortrijk, Belgium). A dendrogram was constructed from the distance values with the unweighted pair group method with arithmetic mean (UPGMA) in Gelcompar 4.0 (Applied Maths, Kortrijk, Belgium). Substrate utilisation profiles of the rhizobial reference strains included in the analysis were obtained from a previous study (Kruger, 1998).

Table 7.5 Isolates included in the substrate utilisation determination

Isolate	Host plant	Isolate	Host plant
Burkholderia tuberum STM 678^T	*Aspalathus carnosa*	CS 6	*Cyclopia subternata*
Burkholderia phymatum STM 815^T	*Machaerium lunatum*	Cses 4	*Cyclopia sessiliflora*
CB 2	*Cyclopia buxifolia*	UCT 30	*Cyclopia sessiliflora*
CD 9	*Cyclopia dregeana*	UCT 34	*Cyclopia glabra*
CI 3	*Cyclopia intermedia*	UCT 70	*Cyclopia maculata*
CS 2	*Cyclopia subternata*	UCT 71	*Cyclopia glabra*

RESULTS

Amplification of the 16S rDNA

The primers fD1 and rP2 (Weisburg *et al.*, 1991) were able to amplify the 16S rDNA gene of the selected isolates. The size of the amplification products of the 16S rDNA gene was in the range of approximately 1500 bp (results not shown).

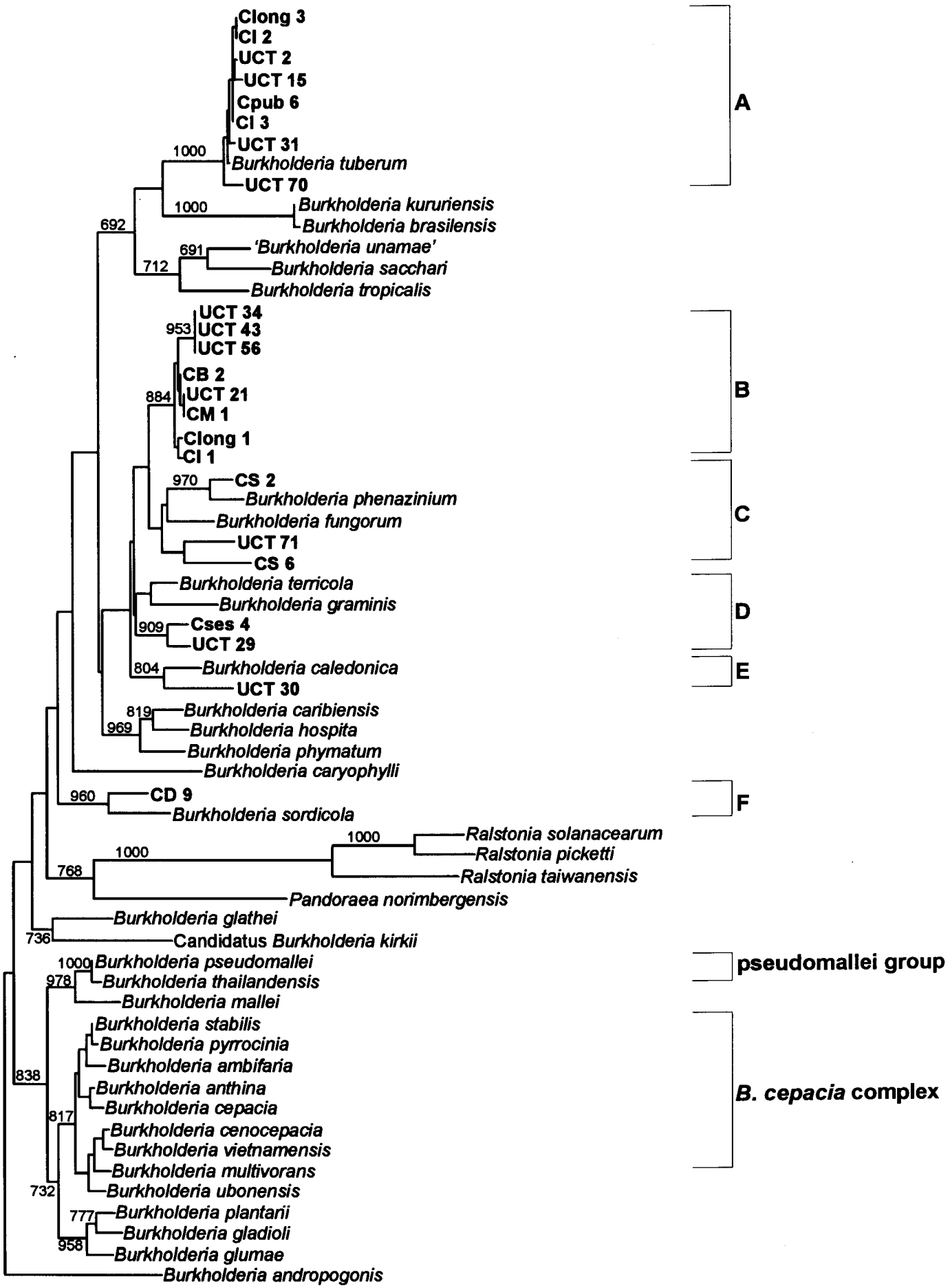
16S rDNA sequence analysis

The sequencing reactions conducted with the two primer sets, namely the internal forward primer 16SRNAII-S (Kuhnert *et al.*, 1996) and the reverse primer rP2 (Weisburg *et al.*, 1991), the forward primer fD1 (Weisburg *et al.*, 1991) and the internal reverse primer 16SRNAVI-S (Kuhnert *et al.*, 1996) were able to give two unambiguous DNA sequences for each isolate of approximately 700 bp. The last part of each strand had ambiguous positions, since the sequencer had problems to distinguish the correct signal. However, the ambiguous positions could be resolved using the other strand. It was possible to overlap the two strands, since the end and beginning of the respective strands were the same sequence. The sequences were deposited in the GenBank database (see Table 7.1 for accession numbers).

Phylogenetic relationship of *Cyclopia* isolates in the β-*Proteobacteria* determined with near full-length 16S rDNA sequence analysis

The *Cyclopia* isolates clustered in six different groups based on the comparative 16S sequence analysis (see Fig. 7.1). Three lineages could be recognised to which the *Cyclopia* isolates belong. Clusters B, C, D and E belonged to one lineage, while cluster A belonged to a lineage, which shared a common ancestor with the *Burkholderia* spp.: *B. kururiensis*, *B. brasilensis*, *B. sacchari* and *B. tropicalis*. The undescribed possibly new species from maize roots, *B. unamae* was the closest neighbour of *B. sacchari* showing sequence similarity values of 98.5%. Cluster F formed a separate lineage.

Figure 7.1 (next page): **Unrooted phylogenetic tree reconstructed with the neighbour-joining method from the comparative 16S rDNA sequence analysis of the *Cyclopia* isolates and reference strains of the genus *Burkholderia*. Reference strains of the genera *Pandoraea* and *Ralstonia* were included for clarity. Branch lenghts are proportional to the phylogenetic distances, while the vertical branches are non-informative. The scale bar shows 1% nucleotide difference. Bootstrap values higher than 600 are indicated.**


Clong 3
CI 2
UCT 2
UCT 15
Cpub 6
CI 3
UCT 31
Burkholderia tuberum
UCT 70
Burkholderia kururiensis
Burkholderia brasilensis
'Burkholderia unamae'
Burkholderia sacchari
Burkholderia tropicalis
UCT 34
UCT 43
UCT 56
CB 2
UCT 21
CM 1
Clong 1
CI 1
CS 2
Burkholderia phenazinium
Burkholderia fungorum
UCT 71
CS 6
Burkholderia terricola
Burkholderia graminis
Cses 4
UCT 29
Burkholderia caledonica
UCT 30
Burkholderia caribiensis
Burkholderia hospita
Burkholderia phymatum
Burkholderia caryophylli
CD 9
Burkholderia sordicola
Ralstonia solanacearum
Ralstonia picketti
Ralstonia taiwanensis
Pandoraea norimbergensis
Burkholderia glathei
Candidatus Burkholderia kirkii
Burkholderia pseudomallei
Burkholderia thailandensis
Burkholderia mallei
Burkholderia stabilis
Burkholderia pyrrocinia
Burkholderia ambifaria
Burkholderia anthina
Burkholderia cepacia
Burkholderia cenocepacia
Burkholderia vietnamensis
Burkholderia multivorans
Burkholderia ubonensis
Burkholderia plantarii
Burkholderia gladioli
Burkholderia glumae
Burkholderia andropogonis
1000
1000
692
691
712
953
884
970
909
804
819
969
960
1000
1000
768
736
1000
978
838
817
732
777
958
A
B
C
D
E
F
pseudomallei group
B. cepacia complex

Cluster A

In cluster A, *Cyclopia* isolates from six different *Cyclopia* species (*C. genistoides*, *C. intermedia*, *C. longifolia*, *C. maculata*, *C. pubescens* and *C. sessiliflora*) grouped with the root-nodulating isolate from *Aspalathus carnosa*, *Burkholderia tuberum*. The *Cyclopia* isolates shared high sequence similarities with the reference strain *B. tuberum*, ranging from 99.6-99.9%. Cluster A was found well resolved with a bootstrap value of 100%. *B. kururiensis* was the closest neighbour of the clade sharing 97% sequence similarity with *B. tuberum*.

Cluster B

Cluster B is a highly related cluster in which no reference strain clustered. This cluster was supported by a significant bootstrap value of 95.3%. Isolates from seven different *Cyclopia* species (*C. buxifolia*, *C. genistoides*, *C. glabra*, *C. intermedia*, *C. longifolia*, *C. maculata*, *C. meyeriana*) belonged to this cluster. Isolates UCT 43 and UCT 56 (both from *C. meyeriana*) and UCT 34 (*C. glabra*) shared 100% sequence similarity with each other and differed most from the rest of the cluster.

Cluster C

In cluster C, three groups could be distinguished. On one branch isolate CS 2 (*C. subternata*) and the soil organism *B. phenazinium* grouped, while *B. fungorum* formed another branch and the two *Cyclopia* isolates UCT 71 (*C. glabra*) and CS 6 (*C. subternata*) formed the other branch. Unlike cluster A and cluster B, cluster C was not so closely related. Isolate CS 2 shared the highest similarity with *B. phenazinium* (99.1%). The two isolates UCT 71 and CS 6 shared 98.2% sequence similarity, while sharing 98.1% and 97.9% sequence similarity with *B. fungorum* respectively. Cluster C was not a significant and repeatable cluster, though the branch leading to the CS 2 and *B. phenazinium* was significant (97%). The branch leading to isolates UCT 71 and CS 6 was supported with a low bootstrap value (52.4%).

Cluster D

The reference strains *B. terricola* and *B. graminis* grouped together with the isolates Cses 4 and UCT 29 (both from *C. sessiliflora*) in cluster D. The two reference strains formed a separate branch, while the two *Cyclopia* isolates formed another branch. The clustering of the two *Cyclopia* isolates was supported with a highly significant bootstrap value of 90.9%. These two isolates shared 99.3% sequence similarity. The reference strain *B. terricola* shared 98.6% and 98.5% with Cses 4 and UCT 29 respectively.

Cluster E

The rhizosphere isolate, *B. caledonica* and isolate UCT 30 (*C. sessiliflora*) belonged to cluster E, sharing 98.4% sequence similarity. Cluster E was a well-resolved grouping, since it was supported with a bootstrap value of 80.4%.

Cluster F

In cluster F, the reference strain *B. sordicola* and isolate CD 9 (*C. dregeana*) clustered. Isolate CD 9 shared 98.5% sequence similarity with *B. sordicola*. Cluster F formed a well-resolved clade with a high bootstrap value (96%).

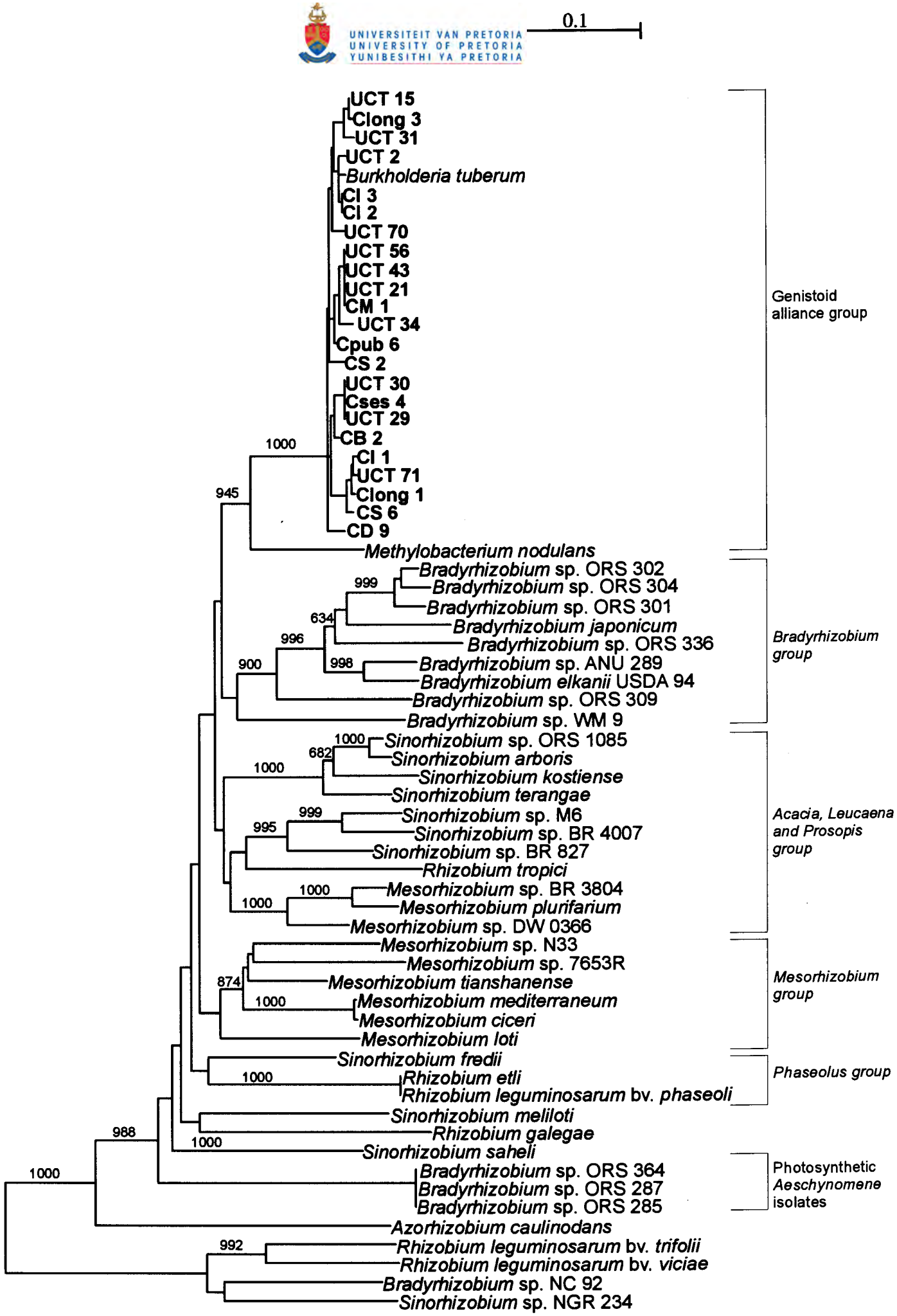
***NodA* PCR**

The primer set (NodAunivF145u and NodAR.brad) used was able to amplify the *nodA* gene resulting in a fragment size of 455 bp as was expected from the results of Moulin *et al.* (2001). In some of the isolates, non-specific fragments were also amplified, since the primers are degenerate.

***NodA* sequence analysis**

The forward primer used was able to determine the sequence of the *nodA* gene. The sequencing results were checked visually by comparing the peaks and the called nucleotides, since the sequencer sometimes calls two nucleotides for a single peak. The sequences were also edited using the *nodA* sequence of *Burkholderia tuberum* as a reference sequence. Any

Figure 7.2 (next page): **Unrooted neighbour-joining tree reconstructed from comparative partial *nodA* sequence analysis. Horizontal branch lengths reflect phylogenetic distances, while the vertical branch lengths are non-informative and set for clarity. The scale bar indicates 10% nucleotide difference. Bootstrap values found in more than 600 of the 1000 replications are shown.**



ambiguous positions were corrected. A strand of approximately 390 bp of unambiguous positions was obtained. The sequences were deposited in the GenBank database (see Table 7.1 for accession numbers).

All the *nodA* genes of the *Cyclopia* isolates formed a single well-resolved clade based on the comparative sequence analysis (see Fig. 7.2). This clade was supported with a 100% bootstrap value. In the clade, the *nodA* gene of *Burkholderia tuberum* also clustered. The *nodA* gene of the α-*Proteobacteria* species, *Methylobacterium nodulans* shared 83.5% sequence similarity with the *Burkholderia* spp. clade.

Numerical taxonomy

The substrate utilisation of 95 carbon sources of the selected isolates could be determined with the Biolog microplates. The 95 carbon sources were divided into the 11 groups as done by Garland and Mills (1991). In Table 7.6, the different oxidation patterns of these isolates are shown. None of the isolates could utilise 2,3-butanediol, phenylethylamine, putrescine, glycyl-L-aspartic acid, gentobiose, maltose, D-melibiose, turanose, glucose-1-phosphate or α-cyclodextrin. All the isolates could utilise glycerol (an alcohol), succinamic acid (an amide), eleven of the 20 amino acids, urocanic acid (an aromatic chemical), bromo-succinic acid (a brominated chemical), thirteen of the 28 carbohydrates, thirteen of the 24 carboxylic acids, both esters tested for and finally two of the five polymers.

The substrate utilisation data were used to construct a dendrogram (see Fig. 7.3) to show the phenotypic similarities of the isolates in a schematic format. The data of the rhizobial isolates were included in the analysis to show the phenotypic differences/similarities between the known rhizobial isolates and the *Burkholderia* isolates. Three main clusters could be distinguished in the dendrogram. The first cluster contained all the *Cyclopia* isolates as well as the *Burkholderia* reference strains. In the second cluster, species from the genera *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* grouped, while the third cluster contained strains of the *Bradyrhizobium* genus. Isolates in cluster 1 shared 80% similarity, while the isolates of clusters 2 and 3 shared 70.5% and 85% similarity respectively.

In cluster 1a, UCT 34, CB 2 and UCT 71 grouped with *B. phymatum* (STM 815). Isolates UCT 34 and CB 2 displayed high phenotypic similarity (93.5%). Inspection of the substrate

utilisation pattern revealed that CB 2 and UCT 34 could not use D-raffinose, citric acid and sucrose, while *B. phymatum* and UCT 71 did. Isolates CB 2 and UCT 34 could utilise itaconic acid, while *B. phymatum* and UCT 71 could not. UCT 71 is the only *Burkholderia* isolate tested capable of growth on inosine and β-methyl D-glucoside. *B. phymatum* did not utilise glycyl-L-glutamic acid, uridine or α-keto-glutaric acid, while the other isolates in the cluster did. The phosphorylated chemical, D, L-α-glycerol phosphate was uniquely used by *B. phymatum*.

Isolates Cses 4 and UCT 30 formed cluster 1b and displayed 91.5% similarity. Isolate UCT 30 could uniquely utilise cellobiose. Isolate Cses 4 was unable to utilise L-ornithine, *N*-acetyl-D-galactosamine, D-trehalose, γ-hydroxybutyric acid, α-keto-valeric acid or D-saccharic acid, while UCT 30 could utilise all the previously named substrates.

Cluster 1c contained CD 9 and CS 2, which shared 88.5% phenotypic similarity. The isolates could be distinguished based on their substrate utilisation. CD 9 could utilise glucuronamide, D-serine, xylitol, D-glucoronic acid, glucose-6-phosphate, while CS 2 was not able to utilise these substrates, but could utlise N-acetyl-galactosamine, γ-hydroxybutyric acid and dextrin. The isolates CD 9 and CS 2 were able to uniquely utilise *i*-erythritol and α-D-lactose respectively.

In cluster 1d isolates UCT 70, CI 3, CS 6 and *Burkholderia tuberum* (STM 678) grouped at an overall similarity value of 89.5%. *B. tuberum*, UCT 70 and CI 3 shared 93.5% phenotypic similarity. These three strains could be distinguished from CS 6 as well as the other *Burkholderia* isolates included in the analysis based on their inability to utilise D, L-carnitine. CS 6 was unable to grow on L-alanyl-glycine and α-hydroxybutyric acid, while these substrates were utilised by all *Burkholderia* isolates studied.

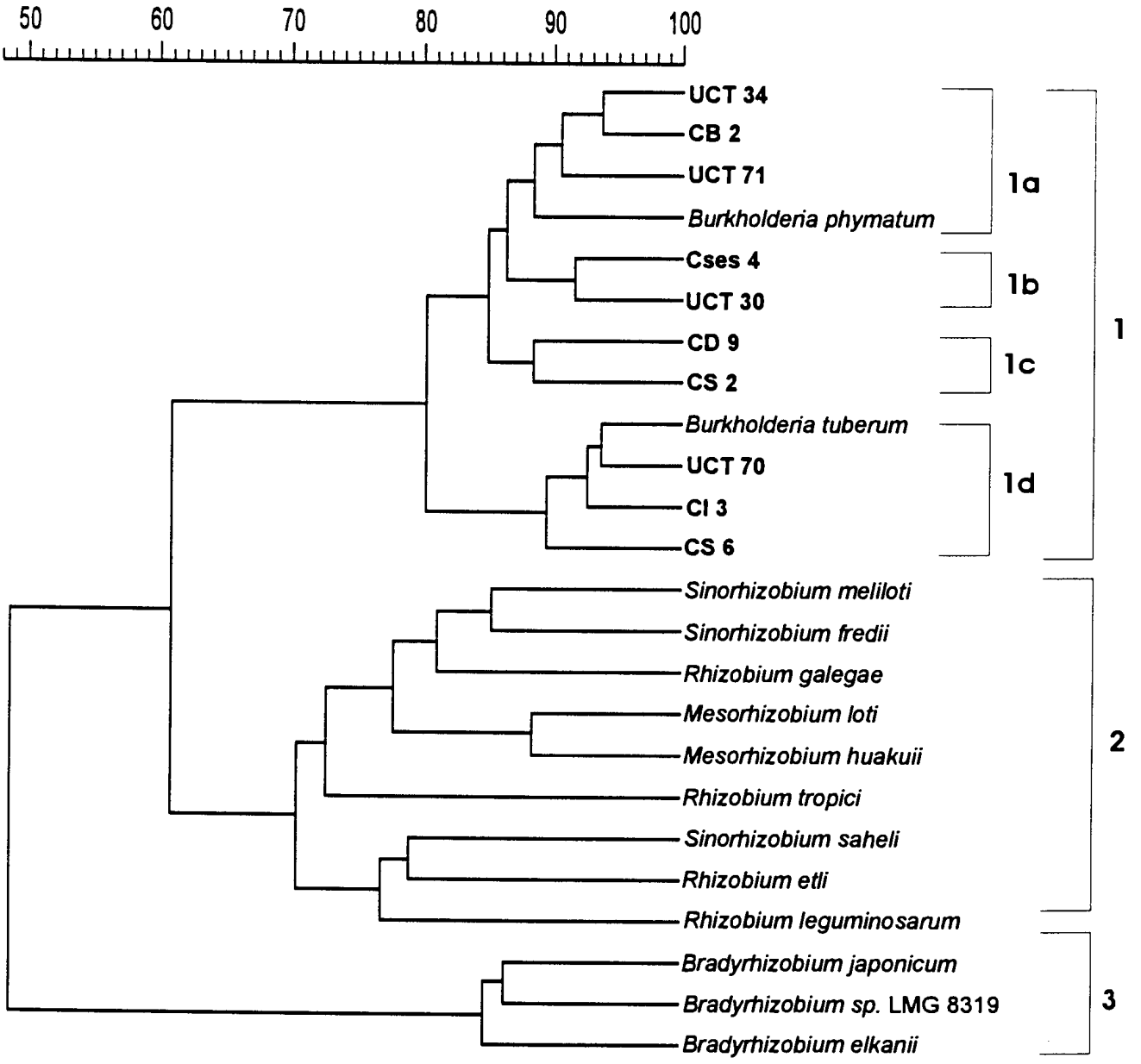


Figure 7.3: **Schematic representation of the substrate utilisation patterns of the *Cyclopia* isolates and the *Burkholderia* spp. included in the analysis. The rhizobial strains were included as references. The x-axis shows the correlation between the isolates and displays similarity values for convenience.**

Table 7.6: Oxidation patterns of the different carbon sources utilised by selected *Cyclopia* isolates and the *Burkholderia* sp. included in the study

Carbon sources	STM 678^T	STM 815^T	CB 2	CD 9	CI 3	CS 2	CS 6	Cses 4	UCT 30	UCT 34	UCT 70	UC 71
Alcohols												
2,3-butanediol	-	-	-	-	-	-	-	-	-	-	-	-
glycerol	+	+	+	+	+	+	+	+	+	+	+	+
Amides												
alaninamide	-	+	+	+	-	+	-	+	+	+	-	+
glucuronamide	-	-	+	+	+	-	-	-	-	-	+	-
succinamic acid	+	+	+	+	+	+	+	+	+	+	+	+
Amines												
2-amino-ethanol	-	+	+	+	-	+	-	+	+	+	-	+
phenylethylamine	-	-	-	-	-	-	-	-	-	-	-	-
putrescine	-	-	-	-	-	-	-	-	-	-	-	-
Amino acids												
D-alanine	+	+	+	+	+	+	+	+	+	+	+	+
L-alanine	+	+	+	+	+	+	+	+	+	+	+	+
L-alanyl-glycine	+	+	+	+	+	+	-	+	+	+	+	+
L-asparagine	+	+	+	+	+	+	+	+	+	+	+	+
L-aspartic acid	+	+	+	+	+	+	+	+	+	+	+	+
L-glutamic acid	+	+	+	+	+	+	+	+	+	+	+	+
glycyl-L-aspartic acid	-	-	-	-	-	-	-	-	-	-	-	-
glycyl-L-glutamic acid	+	-	+	+	-	+	-	-	-	+	+	+
L-histidine	+	+	+	+	+	+	+	+	+	+	+	+
hydroxy-L-proline	+	+	+	+	-	+	-	+	+	+	-	+
L-leucine	+	+	+	+	+	+	+	+	+	+	+	+
L-ornithine	-	+	+	-	-	-	-	-	+	+	-	+
L-phenylalanine	+	+	+	+	+	+	+	+	+	+	+	+
L-proline	+	+	+	+	+	+	+	+	+	+	+	+
L-pyroglutamic acid	+	+	+	+	+	+	+	+	+	+	+	+
D-serine	-	+	+	+	+	-	-	-	-	+	-	+
L-serine	+	+	+	+	+	+	+	+	+	+	+	+
L-threonine	+	+	+	+	+	+	-	+	+	+	+	+
D,L-carnitine	-	+	+	+	-	+	+	+	+	+	-	+
γ-amino butyric acid	-	+	+	+	-	+	-	+	+	+	-	+
Aromatic chemicals												
inosine	-	-	-	-	-	-	-	-	-	-	-	+
thymidine	-	-	-	-	-	+	-	-	-	-	-	-
uridine	-	-	+	+	-	+	-	-	-	+	-	+
urocanic acid	+	+	+	+	+	+	+	+	+	+	+	+
Brominated chemicals												
bromo-succinic acid	+	+	+	+	+	+	+	+	+	+	+	+
Carbohydrates												
N-acetyl-D-galactosamine	-	+	+	-	-	+	-	-	+	+	-	+
N-acetyl-D-glucosamine	+	+	+	+	+	+	+	+	+	+	+	+
adonitol	+	+	+	+	+	+	+	+	+	+	+	+
L-arabinose	+	+	+	+	+	+	+	+	+	+	+	+
L-arabitol	+	+	+	+	+	+	+	+	+	+	+	+
cellobiose	-	-	-	-	-	-	-	-	+	-	-	-
i-erythritol	-	-	-	+	-	-	-	-	-	-	-	-
D-fructose	+	+	+	+	+	+	+	+	+	+	+	+
L-fucose	+	+	+	+	+	+	+	+	+	+	+	+
D-galactose	+	+	+	+	+	+	+	+	+	+	+	+
gentobiose	-	-	-	-	-	-	-	-	-	-	-	-
α-D-glucose	+	+	+	+	+	+	+	+	+	+	+	+
m-inositol	+	+	+	+	+	+	+	+	+	+	+	+

Table 7.6 continued

Carbon sources	STM 678^T	STM 815^T	CB 2	CD 9	CI 3	CS 2	CS 6	Cses 4	UCT 30	UCT 34	UCT 70	UC 71
α-D-lactose	-	-	-	-	-	+	-	-	-	-	-	-
lactulose	+	-	-	+	+	+	+	-	-	-	+	-
maltose	-	-	-	-	-	-	-	-	-	-	-	-
D-mannitol	+	+	+	+	+	+	+	+	+	+	+	+
D-mannose	+	+	+	+	+	+	+	+	+	+	+	+
D-melibiose	-	-	-	-	-	-	-	-	-	-	-	-
β-methyl D-glucoside	-	-	-	-	-	-	-	-	-	-	-	+
psicose	+	+	+	+	-	+	-	+	+	+	+	+
D-raffinose	-	+	-	-	-	-	-	-	+	-	-	+
L-rhamnose	+	+	+	+	+	+	+	+	+	+	+	+
D-sorbitol	+	+	+	+	+	+	+	+	+	+	+	+
sucrose	-	+	-	-	-	-	-	+	+	-	-	+
D-trehalose	-	+	+	-	-	-	-	-	+	+	-	+
turanose	-	-	-	-	-	-	-	-	-	-	-	-
xylitol	-	+	+	+	-	-	-	+	+	+	+	+
Carboxylic acids												
acetic acid	+	+	+	+	+	+	+	+	+	+	+	+
cis-aconitic acid	+	+	+	+	+	+	+	+	+	+	+	+
citric acid	+	+	-	+	+	+	+	-	-	-	+	+
formic acid	+	+	+	+	+	+	+	+	+	+	+	+
D-galactonic acid lactone	+	+	+	+	+	+	+	+	+	+	+	+
D-galacturonic acid	+	+	+	+	+	+	+	-	-	-	+	+
D-gluconic acid	+	+	+	+	+	+	+	+	+	+	+	+
D-glucosaminic acid	+	+	+	+	+	+	+	+	+	+	+	+
D-glucoronic acid	+	+	+	+	+	-	+	-	-	-	+	+
α-hydroxybutyric acid	+	+	+	+	+	+	-	+	+	+	+	+
β-hydroxybutyric acid	+	+	+	+	+	+	+	+	+	+	+	+
γ-hydroxybutyric acid	-	-	-	-	-	+	-	-	+	+	-	-
p-hydroxyphenylacetic acid	+	+	+	+	+	+	+	+	+	+	+	+
itaconic acid	+	-	+	-	+	-	-	-	+	+	+	-
α-keto-butyric acid	+	+	+	+	-	+	-	+	+	+	+	+
α-keto-glutaric acid	+	-	+	-	-	-	-	+	+	+	-	+
α-keto-valeric acid	-	-	-	-	-	-	-	-	+	+	-	+
D,L-lactic acid	+	+	+	+	+	+	+	+	+	+	+	+
malonic acid	-	+	+	+	-	+	-	+	+	+	+	+
propionic acid	+	+	+	+	+	+	+	+	+	+	+	+
quinic acid	+	+	+	+	+	+	+	+	+	+	+	+
D-saccharic acid	+	+	+	+	+	+	+	-	+	+	+	+
sebacic acid	+	+	+	+	+	+	+	+	+	+	+	+
succinic acid	+	+	+	+	+	+	+	+	+	+	+	+
Esters												
mono-methylsuccinate	+	+	+	+	+	+	+	+	+	+	+	+
methylpyruvate	+	+	+	+	+	+	+	+	+	+	+	+
Phosphorylated chemicals												
D,L-α-glycerol phosphate	-	+	-	-	-	-	-	-	-	-	-	-
glucose-1-phosphate	-	-	-	-	-	-	-	-	-	-	-	-
glucose-6-phosphate	-	+	-	+	-	-	-	+	+	-	-	-
Polymers												
glycogen	-	+	+	+	-	+	-	-	-	+	+	+
α-cyclodextrin	-	-	-	-	-	-	-	-	-	-	-	-
dextrin	-	-	+	-	-	+	-	-	-	-	-	+
Tween-40	+	+	+	+	+	+	+	+	+	+	+	+
Tween-80	+	+	+	+	+	+	+	+	+	+	+	+

DISCUSSION

The nearly full-length 16S rDNA sequence data confirmed the identity of some of the *Cyclopia* isolates as strains of the species *Burkholderia tuberum* (Moulin *et al.*, 2001; Vandamme *et al.*, 2002). None of the *Cyclopia* isolates belonged to the species *B. phymatum*, which shared highest sequence similarity with *B. caribiensis* (Vandamme *et al.*, 2002). From the comparative sequence analysis, it is clear that the rest of the isolates included do not unequivocally belong to one of the *Burkholderia* species. Similar new isolates have to be included in the description of new species to refrain from describing one strain species.

Isolates of cluster B are clearly strains of a new *Burkholderia* species, possibly even two new *Burkholderia* species. These isolates shared high 16S sequence similarities with each other and belonged to a well-resolved clade. *B. fungorum* is the closest phylogenetic neighbour (98.6%) of the new species based on 16S sequence data. The two representative isolates used in the phenotypic utilisation profile analysis shared high similarity values (93.5%). The strains can also be differentiated based on the distinct restriction patterns obtained with IGS-RFLP analysis. A new species, *Burkholderia capensis* sp. nov. is provisionally proposed for the cluster B strains. A formal species proposal would be done after DNA-DNA hybridisation analysis and G + C content determinations have been performed (Vandamme *et al.*, 1996; Stackebrandt *et al.*, 2002). A more extensive phenotypic characterisation of several of the cluster B strains would also first have to be done to gather phenotypic information to aid the differentiation of this species from the other *Burkholderia* species.

The species of the genus *Burkholderia* show high 16S sequence similarity, which highlights the problem of distinguishing between closely-related species (Leblond-Bourget *et al.*, 1996). DNA-DNA hybridisation studies would have to be done to determine the taxonomic position of the isolates. DNA-DNA homology studies provide a consolidated measure to delineate bacterial species and the technique can be used to identify unknown isolates (Rosselló-Mora and Amann, 2001).

The basic topology of the phylogenetic tree agreed with that previously found in other studies (Brämer *et al.*, 2001; Fain and Haddock, 2001; Goris *et al.*, 2002; Van Oevelen *et al.*, 2002). The high sequence similarity between *B. kururiensis* and *B. brasilensis* has been reported by other researchers (Fain and Haddock, 2001; Marin *et al.*, 2003). *Pandoraea norimbergensis*

was previously described as a species of *Burkholderia*, but was transferred by Coenye *et al.* (2000) to the genus *Pandoraea*. From the comparative sequence data, it is clear that this transfer was warranted, since *P. norimbergensis* formed a separate lineage on the phylogenetic tree. *B. andropogonis* formed a distinct separate branch, which has been found by other researchers as well (Viallard *et al.*, 1998; Coenye *et al.*, 2001; Fain and Haddock, 2001; Van Oevelen *et al.*, 2002).

The numerical taxonomy study confirmed the 16S rDNA sequence analysis and showed the phenotype of the *Cyclopia* isolates as different from that of the known rhizobial genera as the dendrogram clearly separated the *Burkholderia* isolates and the rhizobial isolates. The dendrogram reflected the considerable difference in phenotype between the genus *Bradyrhizobium* and the genera *Mesorhizobium*, *Rhizobium* and *Sinorhizobium*. All the *Cyclopia* isolates as well as the *Burkholderia* spp. included could utilise D-galactose, D-glucose, glycerol, inositol, mannitol, D-mannose, sorbitol, L-arabinose and D-fructose. These carbon sources can be utilised by all *Burkholderia* strains (Viallard *et al.*, 1998). The inability to utilise maltose or D-turanose also confirmed the identity of the strains as *Burkholderia* strains (Viallard *et al.*, 1998).

All the *Cyclopia* isolates have nearly the same *nodA* gene, which indicates that the gene is relatively conserved in these organisms. Chaintreuil *et al.* (2001) reported the clear distinction between the *nodA* genes of the photosynthetic *Bradyrhizobium* isolates (ORS 285, ORS 287 and ORS 364) and the non-photosynthetic *Bradyrhizobium* isolates (ORS 301, ORS 302 and ORS 304) from *Aeschynomene*. This distinction and the forming of a separate lineage by the photosynthetic isolates, while the non-photosynthetic isolates belonged to the *Bradyrhizobium* clade was clear from the *nodA* gene tree. The conserved nature (100% sequence similarity) of the *nodA* gene of the photosynthetic isolates was also clear from the *nodA* tree as previously found by Chaintreuil *et al.* (2001). Rhizobia isolated from *Acacia*, *Leucaena* and *Prosopis* clustered in a well-resolved clade. *Rhizobium tropici* and *Mesorhizobium* sp. BR3804 also belonged to this clade as was reported by Ba *et al.* (2002). Strains of *R. tropici* can also nodulate *Leaucaena* sp. (Martínez-Romero *et al.*, 1991). As was previously reported by Zhang *et al.* (2000), the high sequence similarity between the *nodA* sequences of *Mesorhizobium ciceri* and *M. mediterraneum*, both isolated from *Cicer arietinum*, was also evident in the *nodA* tree reconstructed in this study.

From the phylogenetic tree based on partial *nodA* sequence data it is clear that different chromosomal backgrounds harbour the same symbiotic profile. The different *Cyclopia* species are not nodulated by a specific symbiotic genotype. Silva *et al.* (2003) studied the genetic structure of *R. etli* and *R. gallicum* strains in Mexico and concluded that the plant host impose selective pressure on the rhizobia which favours the maintenance of specific chromosomal and symbiotic combinations.

In their study on *Astragalus sinicus* rhizobia, Zhang *et al.* (2000) found that some rhizobia with different chromosomal genotypes had identical *nodA* genes, which suggest horizontal gene transfer of the *nod* genes between diverse rhizobia. It is possible that the *Burkholderia* isolates acquired the symbiotic genes through horizontal gene transfer from either *Bradyrhizobium* or *R. tropici* strains, which shared the niche with the *Burkholderia* organisms. The *Cyclopia* plants grow in soils with very low pH values (personal communication, J. Bloem). Curtis *et al.* (2002) reported the isolation of aciduric *Burkholderia* isolates from acidic soil capable of growth at pH ranges of approximately 3.5-8. These isolates shared high sequence similarity with *B. stabilis* (97%) and *B. fungorum* (98%). The nitrogen-fixing species *B. vietnamensis* was also isolated from acidic soil (Gillis *et al.*, 1995).

Burkholderia tuberum and *Methylobacterium nodulans* were both isolated from members of the genistoid alliance in the Papillionoideae. Members of the alliance produce characteristic quinolizidine alkaloids (Van Wyk, 2003). The characterisation of other members of this alliance might lead to the discovery of more novel associations.

CHAPTER 8

PHYLOGENETIC AND SYMBIOTIC CHARACTERISATION OF THE α-*PROTEOBACTERIA CYCLOPIA* ISOLATES

ABSTRACT

The isolates had been previously characterised with 16S-23S IGS-RFLP and partial 16S-sequencing analyses. In order to further investigate the phylogenetic position of these isolates near full-length 16S sequencing analysis was used. The symbiotic genotype of the isolates was determined with *nodA* sequence analysis. The isolates belonged to two *Bradyrhizobium* genomic species, *Rhizobium tropici* and a possible new genus in the α-*Proteobacteria*. All the isolates had been collected from acidic soil and the finding of *Bradyrhizobium* and *Rhizobium tropici* was thus expected. The significance of the isolate, which might be a member of a new genus, needs to be further investigated. The symbiotic genotype of all the isolates was similar to that of *Burkholderia tuberum*.

Keywords: 16S rDNA sequencing, *nodA* sequencing, *Bradyrhizobium*, *Rhizobium tropici*, acid-tolerant strains

INTRODUCTION

Legumes form a symbiotic association with root-nodulating bacteria, collectively called rhizobia. The gram-negative rhizobia belong to several genera in the α-*Proteobacteria*, namely *Allorhizobium* (De Lajudie *et al.*, 1998a), *Azorhizobium* (Dreyfus *et al.*, 1988), *Bradyrhizobium* (Jordan, 1984; Kuykendall *et al.*, 1992; Xu *et al.*, 1995; Yao *et al.*, 2002), *Mesorhizobium* (Jarvis *et al.*, 1997; De Lajudie *et al.*, 1998b; Wang *et al.*, 1999b; Veláquez *et al.*, 2001), *Sinorhizobium* (Chen *et al.*, 1988; De Lajudie *et al.*, 1994; Rome *et al.*, 1996b; Nick *et al.*, 1999; Wang *et al.*, 2002; Wei *et al.*, 2002, Toledo *et al.*, 2003) and *Rhizobium* (Jordan, 1984; Lindström, 1989; Martínez-Romero *et al.*, 1991; Segovia *et al.*, 1993; Amarger *et al.*, 1997; Chen *et al.*, 1997; Van Berkum *et al.*, 1998; Wang *et al.*, 1998; Tan *et al.*, 2001b; Wei *et al.*, 2002; Squartini *et al.*, 2002; Wei *et al.*, 2003). New species of genera unknown to possess the ability to nodulate were described in recent years. *Devosiae neptuniae* (Rivas *et al.*, 2003) and *Methylobacterium nodulans* (Sy *et al.*, 2001) are both capable of nodulation. Recently, the description of species in the β-*Proteobacteria* capable of root-nodulation (Chen *et al.*, 2001; Vandamme *et al.*, 2002) clearly showed that the ability to nodulate rather than the phylogenetic position in the α-*Proteobacteria* warranted the name rhizobia (Geiger and López-Lara, 2002).

The interaction between the plant and the bacteria is specific, since the specific Nod-factor, lipo-chitooligosaccharide or LCO, as well as the concentration is important for nodulation (Perret *et al.*, 2000). The flavonoids of the host plant are recognised by the nodD protein, which then activates the transcription of the *nod* genes. The common genes (*nodABC*) are found in all rhizobia, while the host-specific genes (*nodFE*, *nodH*, *nodSU* and *nodZ*) are found in different combinations in rhizobial species. The common genes are involved in the formation of the LCO backbone, while the host-specific genes are involved in the addition of specific substitutions (Perret *et al.*, 2000; Zhang *et al.*, 2000). However, it has been shown that *nodA* and *nodC* are also host-specific genes, since nodC determines the length of the Nod-factor, while nodA recognises and transfers different acyl chains to the lipo-chitooligosaccharide backbone (Perret *et al.*, 2000).

The phylogeny of the different symbiotic genes, *nodA*, *nodB*, *nodC* and *nodD* resemble each other (Ueda *et al.*, 1995; Haukka et al., 1998). There is a correlation between the phylogeny

of the *nod* genes and host plant range (Haukka *et al.*, 1998; Zhang *et al.*, 2000; Laguerre *et al.*, 2001). The *nodA* has been found to be a good symbiotic marker, since the gene is present in all rhizobia as a single copy and *nodA* analysis reflects Nod-factor features (Haukka *et al.*, 1998; Chaintreuil *et al.*, 2001; Ba *et al.*, 2002). Evidence shows that the *nodA* phylogeny is similar to that of the host plants, which suggests *nod* gene evolution under host constraint and thus possible coevolution of the symbiotic partners as quoted by Radeva *et al.* (2001). The use of 16S rDNA sequence analysis has been found to be an excellent way to determine the phylogenetic position of isolates down to the genus level and has been used extensively (De Lajudie *et al.*, 1998; Khbaya *et al.*, 1998; Terefework *et al.*, 1998; Tan *et al.*, 2001b). All the *Cyclopia* isolates included in this study have been previously characterised with 16S-23S IGS-RFLP and partial 16S sequence analysis. The aim of this study was to further investigate the phylogenetic position of these isolates. *NodA* sequence analysis was conducted to determine the symbiotic genotype of these isolates.

MATERIALS AND METHODS

Bacterial strains used

The strains used in this study (Table 8.1) were received from the Agricultural Research Council-Plant Protection Research Institute (Private Bag X134, Pretoria, 0001, South Africa) and the Botany Department, University of Cape Town (Rondebosch, 7701, Cape Town, South Africa). All the sequences of the reference strains used in the 16S rDNA and *nodA* sequence analysis were obtained from the GenBank database of the National Centre for Biotechnology (NCBI) [website address: www.ncbi.nlm.nih.gov/Genbank/].

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 5-7 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.

Table 8.1 Isolates included in the phylogenetic analysis of the 16S rDNA and *nodA*

Isolate	Host species	16S rDNA GenBank accession number	*NodA* GenBank accession number
UCT 42	*C. plicata*	AY178077	AY189242
UCT 50	*C. sessiliflora*	AY178082	AY189243
Cmey 1	*C. meyeriana*	AY178079	AY189257
Cplic 1	*C. plicata*	AY178081	AY189258
Cpub 4	*C. pubescens*	AY178080	AY189232

Extraction of genomic DNA

A modified method for proteinase-K (Roche Molecular Biochemicals) treated cells as described by Laguerre *et al.* (1997) was used. A fresh culture of each strain, which had been checked for purity, was streaked on a tryptone yeast (TY) agar slant [0.5% (m/v) tryptone (Difco), 0.3% (m/v) yeast extract (Biolab), 0.13% (m/v) $CaCl_2.6H_2O$ (UniLab), 1.5% (m/v) bacteriological agar] in a screw-cap tube. The strains were incubated at 28 °C and checked for sufficient growth. Sterile 4.5 ml dH_2O was added to the slant growth to harvest the cells. An inoculation loop was used to release cells clinging to the agar. The volume of the water added was adjusted according to the amount of growth. Less water was used if the growth was poor and *vice versa*. The cell-suspension was collected in a clear plastic tube and vortexed to ensure a uniform suspension. The absorbancy of the suspension was measured with dH_2O as the spectrophotometric blank at 620 nm. A formula was used to determine the volume of the suspension to be treated further. The volume to be used in ml is equal to 0.2 divided by the absorbancy at 620 nm. Two tubes of the same strain were filled with the appropriate volume of the suspension and centrifuged at 13 000 *g* for 5 minutes at 4 °C. The supernatant was discarded and the excess media blotted dry. One of the tubes was stored at -20 °C for future use. In the second tube, 100 µl ddH_2O, 100 µl Tris-HCl (10 mM, pH 8.2) and 10 µl proteinase-K (15 mg/ml) (Roche Molecular Biochemicals) were added to the cell pellet. The mixture was incubated at 55 °C overnight. In order to inactivate the proteinase-K the mixture was boiled for 10 minutes. The cell lysates were stored at -20 °C until needed.

Amplification of the 16S rDNA and the partial *nodA* genes

Amplification of the 16S rDNA gene of strains (Table 8.1) were performed with the primers fD1 and rP2 (Table 8.2) as described by Weisburg *et al.* (1991). The linker sequences of the

primers were not included in the primer synthesis. These shorter primers were designated fD1SHRT and rP2SHRT. The PCR mixture of each strain contained: 5 µl of the cell lysate, 50 pmol of each primer, 250 µM of each dNTP, 1.5 mM $MgCl_2$ and 0.5 U Gold Taq DNA polymerase (Southern Cross Biotechnology) in a 50 µl reaction volume. The PCR reactions were done on a Perkin Elmer GeneAmp PCR System 2400 thermocycler using the following thermal profile: initial denaturation at 95 °C for 3 minutes, followed by 35 cycles of denaturation (94 °C for 30 seconds), annealing (55 °C for 30 seconds) and extension (72 °C for 1 minute). This was followed by a final extension step at 72 °C for 5 minutes. The concentration, purity and size of the products were evaluated by running an aliquot (5 µl) of each reaction on 0.9% (m/v) horizontal agarose gels (Promega) (results not shown). The standard marker molecular marker VI (Roche Molecular Biochemicals) was included on each gel.

Table 8.2: Primers used in the amplification and/or sequencing of the 16S rDNA and the *nodA* genes

Primer name*	Primer sequence (5'-3')#	Target gene	Reference
fD1SHRT	5'-AGAGTTTGATCCTGGCTCAG-3'	16S rDNA	Weisburg *et al.*, 1991
rP2SHRT	5'-ACGGCTACCTTGTTACGACTT-3'	16S rDNA	Weisburg *et al.*, 1991
16SRNAII-S	5'-GTGTAGCGGTGAAATGCGTAG-3'	16S rDNA	Kuhnert *et al.*, 1996
16SRNAVI-S	5'-CTACGCATTTCACCGCTACAC-3'	16S rDNA	Kuhnert *et al.*, 1996
NodAunivF145u	5'-TGGGCSGGNGCNAGRCCBGA-3'	*nodA*	Moulin *et al.*, 2001
NodAR.brad	5'-TCACARCTCKGGCCCGTTCCG-3'	*nodA*	Moulin *et al.*, 2001

* All the primers were synthesised by Roche Molecular Biochemicals, Mannheim, Germany

\# Abbreviations: B= G/C/T, K= G/T, N= A/G/C/T, R= A/G, S= G/C

The partial *nodA* gene was amplified with the primers NodAunivF145u and NodAR.brad (Table 8.2) as used by Moulin *et al.* (2001) using the same PCR reaction mixture as described for 16S rDNA. The same thermal profile as used for the amplification of the 16S rDNA product was used, except that the annealing time was extended to 45 seconds. Analysis of the amplified product was done as described for the verification of the amplified 16S product (results not shown).

The amplification products of the 16S rDNA and *nodA* were purified, since any traces of unincorporated dNTPs, primers, etc. can negatively influence the sequencing reactions. The 16S products were purified using a Qiagen QIAquick PCR Purification kit (Southern Cross Biotechnology). Purification reactions were done as prescribed by the manufacturer. The

nodA primers are degenerate, resulting in additional amplification products. The products of the desired size were excised from the gel and then purified using the initial steps for gel extraction as prescribed by the manufacturers of the High Pure PCR purification kit (Roche Molecular Biochemicals). The gel solution (containing the desired fragment) was transferred to the columns of the Qiagen QIAquick PCR Purification kit (Southern Cross Biotechnologies). The protocol for the purification was then followed as prescribed by the manufacturers. The concentration and purity of each purification reaction was verified visually. An aliquot (1 µl) of each purified product was run on 0.9% (m/v) horizontal agarose gels (Promega) (results not shown). On each gel, a standard marker, molecular marker VI (Roche Molecular Biochemicals) was included.

16S rDNA and *nodA* sequencing

The sequences of the purified 16S rDNA and *nodA* products were determined using the ABI Prism BigDye™ Terminator Cycle Sequencing Ready Reaction kit (AmpliTaq^R DNA Polymerase, FS) (Perkin Elmer Applied Biosystems). The near full-length sequence of each 16S rDNA product was determined with the internal forward primer 16SRNAII-S (Kuhnert *et al.*, 1996), the forward primer fD1SHRT (Weisburg *et al.*, 1991), the internal reverse primer 16SRNAVI-S (Kuhnert *et al.*, 1996) and the reverse primer rP2SHRT (Weisburg *et al.*, 1991). The purified *nodA* products were sequenced with the forward primer NodAunivF145u (Moulin *et al.*, 2001). Each 5 µl sequencing reaction contained the following: 2 µl of the ready reaction mix supplied with the kit which contains the dye terminators, dNTP's, AmpliTaq^R DNA polymerase, $MgCl_2$ and Tris-HCl buffer pH 9.0; 12.5 pmol primer and approximately 100 ng template DNA. The sequencing reactions were carried out in a Perkin Elmer GeneAmp PCR System 2400 thermocycler with the following thermal profile: an initial denaturation at 96 °C for 5 seconds followed by 25 cycles of denaturation (96 °C for 10 seconds), annealing (50 °C for 5 seconds) and extension (60 °C for 4 minutes). The products were precipitated using the protocol as suggested by the manufacturer. 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 deionised 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.

Phylogenetic analysis of the 16S rDNA and *nodA* sequences

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). The *nodA* sequences were edited visually comparing the nucleotides and their corresponding peaks. Both strands of the 16S rDNA products were sequenced with the primers used and the strands could be aligned to correct ambiguous positions. The final edited two strands were overlapped in the ABI Prism Sequencing Navigator 1.0.1 computer programme to form a continuos sequence reading. The ClustalX programme (Thompson *et al.*, 1997) was used to analyse the edited sequences as well as the reference sequences obtained from GenBank (Table 8.3 and Table 8.4), which were suitably edited. A distance matrix was constructed by pair-wise alignment of the sequences. 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 phylogenetic trees were visualised with NJplot (Perrière and Gouy, 1996).

Table 8.3: Reference sequences obtained from Genbank[1] included in the partial 16S rDNA sequence analysis.

Reference strain	Strain number	Host plant or relevant characteristics	Genbank[1] Accession number
Acetobacter diazotrophicus	LMG 7603^T	*Saccharum officinarum* root	X75618
Afipia clevelandensis	NS	NS	M69186
Afipia felis	NS	NS	M65248
Agrobacterium larrymoorei	NS	NS	Z30542
Agrobacterium radiobacter	ATCC 19358^T	NS	AJ389904
Agrobacterium rhizogenes	LMG 152	NS	X67224
Agrobacterium rubi	IFO 13261	NS	D14503
Agrobacterium tumefaciens	LMG 196	NS	X67223
Agrobacterium vitis	NCPPB 3554	*Vitis vinifera*	D14502
***α-Proteobacterium* strain**	LMG 20591	Agricultural soil	AY040361
Allorhizobium undicola	LMG 11875^T	*Neptunia natans*	Y17047
Aquaspirillum magnetotacticum	NS	NS	M58171
Azorhizobium caulinodans	LMG 6465^T	*Sesbania rostrata*	X67221
Azospirillum brasilense	DSM 2298	NS	X79734
Azospirillum lipoferum	NCIMB 11861	NS	Z29619
Bartonella bacilliformis	NS	NS	M65249
Beijerinckia indica	ATCC 9039^T	Acid soil	M59060
Blastobacter denitrificans	LMG 8443^T	Surface water	S46917
Bradyrhizobium elkanii	USDA 76^T	*Glycine max*	U35000
Bradyrhizobium japonicum	LMG 6138^T	*Glycine max*	X66024

Table 8.3: continued

Reference strain	Strain number	Host plant or relevant characteristics	Genbank[1] Accession number
Bradyrhizobium liaoningense	LMG 18230^T	*Glycine max*	AJ250813
***Bradyrhizobium* genosp. A**	BDV 5028	*Bossiaea ensata*	Z94811
***Bradyrhizobium* genosp. O**	BDV 5840	*Gompholobium huegelii*	Z94823
***Bradyrhizobium* sp.**	Ppau 3-41	*Phaseolus pauciflorus*	AF384137
Bradyrhizobium yuanmingense	CCBAU 10071^T	*Lespedeza cuneata*	AF193818
Brucella neotomae	ATCC 23459	NS	L26167
Devosia neptuniae	J1^T	*Neptunia natans*	AF469072
Ensifer adhaerens	LMG 20582	NS	AY040360
Kaistia adipata	Chj 404^T	Rhizobiaceae group	AY039817
Mesorhizobium amorphae	ACCC 19665	*Amorpha fruticosa*	AF041442
Mesorhizobium chacoense	PR-5^T	*Prosopis alba*	AJ278249
Mesorhizobium ciceri	UPM-Ca7^T	*Cicer arietinum*	U07934
Mesorhizobium huakuii	IAM 14158^T	*Astragalus sinicus*	D12797
Mesorhizobium loti	LMG 6125^T	*Lotus corniculatus*	X67229
Mesorhizobium mediterraneum	UPM-Ca36^T	*Cicer arietinum*	L38825
Mesorhizobium plurifarium	LMG 11892^T	*Acacia senegal*	Y14158
Mesorhizobium tianshanense	A-1BS^T	*Glycyrrhiza pallidiflora*	Y71079
Methylobacterium nodulans	ORS 2060^T	*Crotalaria podocarpa*	AF220763
Mycoplana dimorpha	IAM 13154^T	Soil	D12786
Nitrobacter winogradskyi	ATCC 14123	NS	L35507
Ochrobactrum anthropi	IAM 14119	NS	D12794
Paracoccus denitrificans	LMG 4218^T	Garden soil enriched with 5% K-Na-tartrate + 2% KNO_3	X69159
Phyllobacterium myrsinacearum	IAM 13584	NS	D12789
Phyllobacterium rubiacearum	IAM 13587	NS	D12790
Rhizobium etli	CFN 42^T	*Phaseolus vulgaris*	U28916
Rhizobium galegae	USDA 3394	*Galega officinalis*	AF025853
Rhizobium gallicum	R602sp^T	*Phaseolus vulgaris*	U86343
Rhizobium giardinii	H152^T	*Phaseolus vulgaris*	U86344
Rhizobium hainanensis	I66^T	*Desmodium sinuatum*	U71078
Rhizobium huautlense	USDA 4900^T	*Sesbania herbacae*	AF025852
Rhizobium indigoferae	CCBAU 71042^T	*Indigofera amblyantha*	AY034027
Rhizobium leguminosarum	LMG 8820	*Phaseolus vulgaris*	X67227
Rhizobium loessense	CCBAU 7190B^T	*Astragalus complanatus*	AF364069
Rhizobium mongolense	USDA 1844^T	*Medicago ruthenica*	U89817
Rhizobium sullae	IS123^T	*Hedysarum coronarium*	Y10170
Rhizobium tropici	CIAT 899^T	*Phaseolus vulgaris*	U89832
Rhizobium yanglingense	CCBAU 71462	*Coronilla varia*	AF195031
Rhodobacter sphaeroides	IF0 12203^T	NS	D16425
Rhodoplanes roseus	NS	NS	D25313
Rhodopseudomonas palustris	ATCC 17001	NS	D25312
Rickettsia rickettsii	ATCC VR 891	NS	M21293
Sinorhizobium arboris	HAMBI 1552^T	*Prosopis chilensis*	Z78204
Sinorhizobium fredii	LMG 6217^T	*Glycine max*	X67231
Sinorhizobium kostiense	HAMBI 1489^T	*Acacia senegal*	Z78203
Sinorhizobium kummerowiae	CCBAU 71714^T	*Kummerowia stipulacea*	AY034028
Sinorhizobium medicae	A321^T	*Medicago truncatula*	L39882
Sinorhizobium meliloti	LMG 6133^T	*Medicago sativa*	X67222
Sinorhizobium morelense	Lc04^T	*Leucaena leucocephala*	AY024335
Sinorhizobium saheli	LMG 7837^T	*Sesbania pachycarpa*	X68390
Sinorhizobium terangae	LMG 6463	*Sesbania rostrata*	X68387
Sinorhizobium xinjiangensis	IAM 14142	*Glycine max*	D12796
Xanthobacter agilis	SA 35	NS	X94198
Xanthobacter autotrophicus	NS	NS	X94201
Xanthobacter flavus	NS*	NS	X94199

Table 8.4 ***NodA*** **sequences obtained from GenBank included in the comparative** ***nodA*** **sequence analysis**

Reference strain	Strain number	Host plant	Accession number GenBank[1]
Azorhizobium caulinodans	ORS 571[T]	*Sesbania rostrata*	L18897
Bradyrhizobium elkanii	USDA 94	NS	U04609
Bradyrhizobium japonicum	110spc4	NS	AF322013
Bradyrhizobium sp.	NC92	*Arachis hypogaea*	U33192
Bradyrhizobium sp.	WM9	*Lupinus* sp.	AF222753
Bradyrhizobium sp.	ANU289	*Parasponia* sp.	X03720
Bradyrhizobium sp.	ORS 285	Photosynthetic isolate	AF284858
Bradyrhizobium sp.	ORS 287	*Aeschynomene afraspera*	AJ437607
Bradyrhizobium sp.	ORS 301	*Aeschynomene americana*	AJ437608
Bradyrhizobium sp.	ORS 302	*Aeschynomene pfundii*	AJ437609
Bradyrhizobium sp.	ORS 304	*Aeschynomene elaphroxylon*	AJ437610
Bradyrhizobium sp.	ORS 309	*Aeschynomene uniflora*	AJ437611
Bradyrhizobium sp.	ORS 336	*Aeschynomene afraspera*	AJ437612
Bradyrhizobium sp.	ORS 364	*Aeschynomene nilotica*	AJ437613
Burkholderia tuberum	STM 678[T]	*Aspalathus carnosa*	AJ302321
Mesorhizobium ciceri	USDA 3383	*Hedysarum boreale*	AJ250140
Mesorhizobium loti	NZP 2213[T]	*Lotus corniculatus*	L06241
Mesorhizobium mediterraneum	USDA 3392	NS	AJ250141
Mesorhizobium plurifarium	ORS 1096	*Acacia tortilis* subsp. *raddiana*	AJ302678
Mesorhizobium sp.	BR3804	*Chamaecrista ensiformis*	Z95249
Mesorhizobium sp.	DW0366	*Acacia polycantha*	Z95248
Mesorhizobium sp.	7653R	*Astragalus sinicus*	AJ249353
Mesorhizobium sp.	N33	*Oxytropis arctobia*	U53327
Mesorhizobium tianshanense	USDA 3592	NS	AJ250142
Methylobacterium nodulans	ORS 2060[T]	*Crotalaria podocarpa*	AF266748
Rhizobium etli	CFN 42[T]	*Phaseolus vulgaris*	NC_004041
Rhizobium galegae	HAMBI 1174	*Galega orientalis*	X87578
Rhizobium leguminosarum bv. ***phaseoli***	NS	NS	M58625
Rhizobium leguminosarum bv. ***trifolii***	ANU843	NS	X03721
Rhizobium leguminosarum bv. ***viciae***	NS	NS	Y00548
Rhizobium tropici	CFN 299	*Phaseolus* sp.	X98514
Sinorhizobium arboris	HAMBI 1700	*Acacia senegal*	Z95235
Sinorhizobium fredii	USDA 257	NS	M73699
Sinorhizobium kostiense	HAMBI 1489[T]	*Acacia senegal*	Z95236
Sinorhizobium meliloti	NS	NS	X01649
Sinorhizobium saheli	ORS 609	*Sesbania cannabina*	Z95241
Sinorhizobium sp.	NGR234	Broad host range	AE000076
Sinorhizobium sp.	BR827	*Leucaena leucocephala*	Z95232
Sinorhizobium sp.	BR4007	*Prosopis juliflora*	Z95240
Sinorhizobium sp.	M6	*Prosopis* sp.	Z95233
Sinorhizobium sp.	ORS 1085	*Acacia tortilis* subsp. *raddiana*	AJ302677
Sinorhizobium terangae	ORS 1009	*Acacia laeta*	Z95237

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

T Type strain

ACCC Agricultural Center of Culture Collection, Chinese Academy of Agriculture, Beijing, China

ATCC American Type Culture Collection, Rockville, Maryland, USA

CCBAU Culture Collection of Beijing Agricultural University, Beijing, People's Republic of China

CFN	Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Mexico
CIAT	Rhizobium Collection, Centro International de Agricultura Tropical, Cali, Columbia
DSM	Deutsche Sammlung von Mikroorganismen, Braunschweig, Germany
HAMBI	Culture Collection of the Department of Applied Chemistry and Microbiology, University of Helsinki, Helsinki, Finland
IAM	Institute of Applied Microbiology, University of Tokyo, Tokyo, Japan
IFO	Institute for Fermentation, Osaka, Japan
LMG	BCCM™/LMG Bacteria Collection, Laboratorium voor Microbiologie, University of Gent, Gent, Belgium
NCIMB	National Collections of Industrial and Marine Bacteria, Aberdeen, Scotland, UK
NCPPB	National Collection of Plant Pathogenic Bacteria, Harpenden Laboratory, Hertfordshire, UK
NCTC	National Collection of Type Cultures, Central Public Health Laboratory, London, UK
NZP	Applied Biochemistry Division, Department of Scientific and Industrial Research, Palmerston North, New Zealand
ORS	ORSTOM Collection, Institut Français de Recherche Scientifique pour le Développement en Coopération, Dakar, Senegal
STM	Laboratoire des Symbioses Tropicales et Méditerranéennes, Montpellier, France
UPM	Universidad Politécnica Madrid, Spain
USDA	United States Department of Agriculture, *Rhizobium* Culture Collection, Beltsville Agricultural Research Center, Beltsville, MD, USA

RESULTS

Amplification of the 16S rDNA and the *nodA* gene

The primers fD1SHRT and rP2SHRT (Weisburg *et al.*, 1991) were able to amplify the 16S rDNA gene of the isolates and size of the products corresponded to the expected size of approximately 1500 bp. The partial *nodA* gene was amplified with the primers NodAunivF145u and NodAR.brad used by Moulin *et al.*, (2001) and a product of the expected size of 455 bp was obtained. Due to the degenerate nature of the primers, faint additional bands were visible on the horizontal agarose gel electrophoresis. These bands did not hinder the sequencing reaction since the correct fragment was excised and purified.

Sequence analysis of the 16S rDNA and the *nodA* gene

The sequencing reactions of the 16S rDNA products conducted with the four primers were able to give an unambiguous DNA sequence for each isolate of approximately 1250 bp. The last part of each strand had ambiguous positions. However, the ambiguous positions could be resolved using the other strand. The two edited strands of each isolate could be overlapped to give an uninterrupted sequence. The partial *nodA* sequence could be determined with the primer used. An unambiguous strand of approximately 390 bp was obtained for each isolate. All the sequences were deposited in the GenBank database. The relevant accession numbers can be seen in Table 8.1.

Phylogenetic analysis of the 16S rDNA and the *nodA* gene

The comparative sequence analysis based on the 16S rDNA sequences reflected the polyphyletic nature of the rhizobia (see Fig. 8.1). The rhizobial genera formed five distinct lineages; the *Sinorhizobium* lineage, *Mesorhizobium* lineage, *Rhizobium* lineage, *Bradyrhizobium* lineage and the *Azorhizobium* lineage. The root-nodulating species, *Devosia neptuniae* and *Methylobacterium nodulans* belonged to two additional lineages. The relatively close phylogenetic relationship between the genera *Allorhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* is evident from the tree. The significant phylogenetic separation between *Azorhizobium* and *Bradyrhizobium* and the other rhizobial genera is evident from the tree.

Isolates Cplic 1 and UCT 50 clustered on the *Rhizobium tropici-Agrobacterium rhizogenes* branch, which formed a well-resolved clade supported by a bootstrap value of 100%. Isolate Cplic 1 displayed sequence similarity values of 99.2% and 99.5% with *R. tropici* and *A. rhizogenes* respectively. *A. rhizogenes* and *R. tropici* shared 99.6% and 99.0% with UCT 50.

Isolates UCT 42, UCT 53 and Cmey 1 did not belong to any of the described *Bradyrhizobium* reference strains. These isolates belonged to two different *Bradyrhizobium* genospecies based on 16S rDNA sequence analysis. Cmey 1 shared 100% 16S sequence similarity with the Australian *Bradyrhizobium* genosp. O. The branch leading to the two strains was supported with a bootstrap value of 100%. Isolates UCT 42 and UCT 53 shared 100% sequence

Figure 8.1 (next page): **Phylogenetic tree reconstructed with the neighbour-joining algorithm from a distance matrix of the comparative 16S rDNA sequences analysis of the *Cyclopia* isolates and representative reference strains from the α-*Proteobacteria*. The branch lengths are proportional to the phylogenetic divergence between isolates. The vertical branches are set for clarity and are non-informative. The scale bar indicates 2% sequence divergence. Bootstrap values higher than 600 are indicated on the tree.**

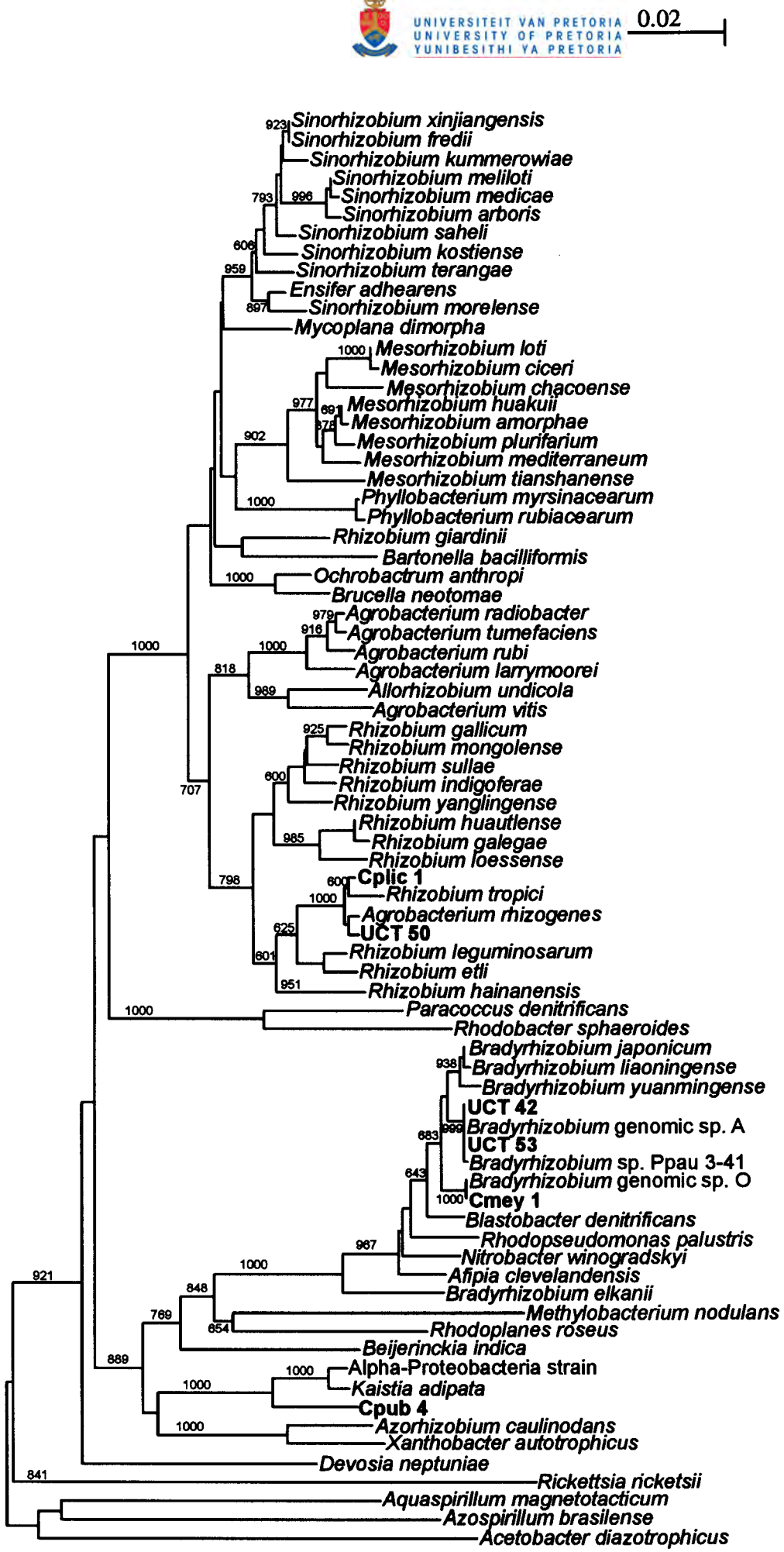
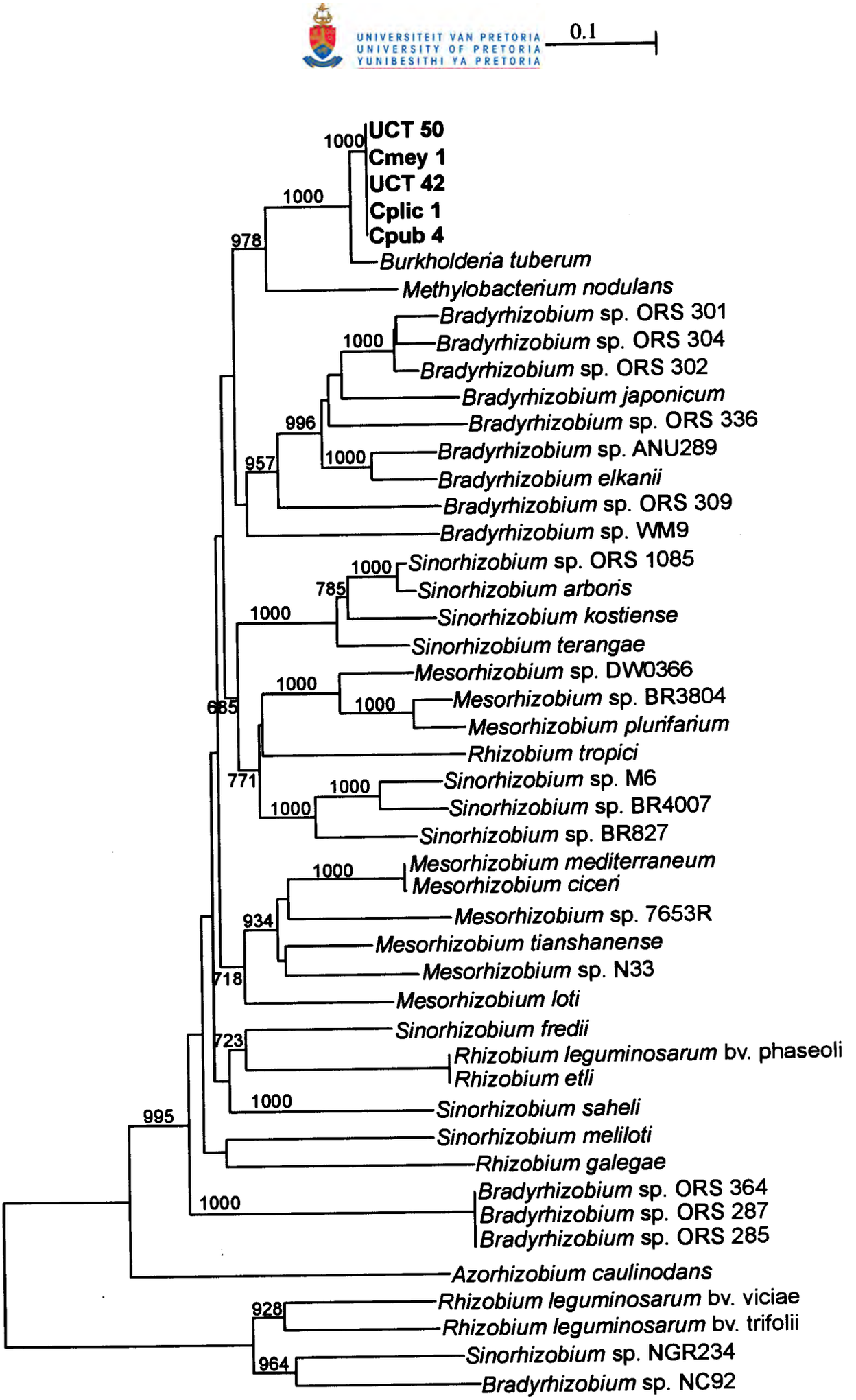

0.02
Sinorhizobium xinjiangensis
Sinorhizobium fredii
Sinorhizobium kummerowiae
Sinorhizobium meliloti
Sinorhizobium medicae
Sinorhizobium arboris
Sinorhizobium saheli
Sinorhizobium kostiense
Sinorhizobium terangae
Ensifer adhearens
Sinorhizobium morelense
Mycoplana dimorpha
Mesorhizobium loti
Mesorhizobium ciceri
Mesorhizobium chacoense
Mesorhizobium huakuii
Mesorhizobium amorphae
Mesorhizobium plurifarium
Mesorhizobium mediterraneum
Mesorhizobium tianshanense
Phyllobacterium myrsinacearum
Phyllobacterium rubiacearum
Rhizobium giardinii
Bartonella bacilliformis
Ochrobactrum anthropi
Brucella neotomae
Agrobacterium radiobacter
Agrobacterium tumefaciens
Agrobacterium rubi
Agrobacterium larrymoorei
Allorhizobium undicola
Agrobacterium vitis
Rhizobium gallicum
Rhizobium mongolense
Rhizobium sullae
Rhizobium indigoferae
Rhizobium yanglingense
Rhizobium huautlense
Rhizobium galegae
Rhizobium loessense
Cplic 1
Rhizobium tropici
Agrobacterium rhizogenes
UCT 50
Rhizobium leguminosarum
Rhizobium etli
Rhizobium hainanensis
Paracoccus denitrificans
Rhodobacter sphaeroides
Bradyrhizobium japonicum
Bradyrhizobium liaoningense
Bradyrhizobium yuanmingense
UCT 42
Bradyrhizobium genomic sp. A
UCT 53
Bradyrhizobium sp. Ppau 3-41
Bradyrhizobium genomic sp. O
Cmey 1
Blastobacter denitrificans
Rhodopseudomonas palustris
Nitrobacter winogradskyi
Afipia clevelandensis
Bradyrhizobium elkanii
Methylobacterium nodulans
Rhodoplanes roseus
Beijerinckia indica
Alpha-Proteobacteria strain
Kaistia adipata
Cpub 4
Azorhizobium caulinodans
Xanthobacter autotrophicus
Devosia neptuniae
Rickettsia ricketsii
Aquaspirillum magnetotacticum
Azospirillum brasilense
Acetobacter diazotrophicus

Figure 8.2 (next page): **Unrooted neighbour-joining tree reconstructed from the comparative sequence analysis of the partial *nodA* sequences of the *Cyclopia* isolates and reference *nodA* sequences obtained from GenBank. Horizontal branch lengths are proportional to the phylogenetic divergence between isolates, while the vertical branch lengths are non-informative and set for clarity. The scale bar indicates 10% sequence divergence. Bootstrap probabilities higher than 600 are indicated at the respective nodes.**



similarity with the Austalian *Bradyrhizobium* genosp. A and nearly 100% similarity with *Bradyrhizobium* sp. Ppau3-41 from Mexico. This clade was also supported with a high bootstrap value (99.9%). These reference strains were included in the comparative analysis since BLAST (Altschul *et al.*, 1990) results revealed high sequence similarity with the *Cyclopia* isolates.

Isolate Cpub 4 showed a sequence similarity value of approximately 97.1% with both the α-*Proteobacteria* strain LMG 20591 isolated from soil and '*Kaistia adipata*'. The GenBank database was searched for similar sequences with the BLAST algorithm (Altschul *et al.*, 1990) and these strains displayed the highest sequence similarity with Cpub 4. Strain '*Kaistia adipata*' and the α-*Proteobacteria* strain LMG 20591 showed a sequence similarity value of 99.3%. The phylogenetic closest rhizobial reference strain, *Azorhizobium caulinodans* showed a sequence similarity value of 92.7% with Cpub 4.

The *nodA* gene of all the *Cyclopia* isolates was highly conserved. The *nodA* gene of *Burkholderia tuberum* was the closest phylogenetic neighbour of the *Cyclopia nodA* gene, sharing 97% sequence similarity. The branch leading to the *nodA* genes of the *Cyclopia* isolates was well resolved (100%), while the clade of the *nodA* genes of *B. tuberum* and the *Cyclopia* isolates were also supported with a bootstrap value of 100%.

DISCUSSION

The polyphyletic nature of the rhizobia (the root-nodulating organisms) was reflected in the phylogenetic tree based on 16S rDNA sequence data. Genera (such as *Agrobacterium*, *Bartonella*, *Brucella*, *Mycoplana*, etc.) unable to nodulate cluster among the root-nodulating bacteria (Willems and Collins, 1993). Most of the grouping is similar to that obtained by other researchers (Tan *et al.*, 2001b; Wei *et al.*, 2002; Yao *et al.*, 2002; Willems *et al.*, 2003), however with some differences.

It is clear that an extensive revision of the genus *Bradyrhizobium* is inevitable. Phylogenetic trees show the polyphyletic nature of the genus *Bradyrhizobium* and the tree constructed in this study supported this.

The phylogenetic relationships within the rhizobial genera could be resolved, though the distinction between the species *Sinorhizobium fredii* and *S. xinjiangense* could not be made. This problem was also experienced by other researchers (Tan *et al.*, 2001; Wei *et al.*, 2002; Yao *et al.*, 2002), which raised questions about the taxonomic validity of *S. xinjiangense*. Peng *et al.* (2002) included novel *S. xinjiangense* isolates in a genotypic (16S sequencing, IGS sequencing and DNA-DNA hybridisation) and phenotypic (SDS-PAGE) analyses and managed to differentiate between these closely related species.

In the *Sinorhizobium* clade which was supported with a high bootstrap value, the two non-nodulating species, *Ensifer adhaerens* and *S. morelense* are more related to each other than to the other *Sinorhizobium* species as previously found by Wang *et al.* (2002) and Willems *et al.* (2003). Willems *et al.* (2003) have requested an opinion for the transfer of *Ensifer adhearens* to the genus *Sinorhizobium*.

The polyphyletic nature of the genus *Rhizobum* prompted Young *et al.* (2001) to give an emended description for the genus *Rhizobium*. The emended description of the genus *Rhizobium* contained all the current *Rhizobium* species, as well as all the species from *Agrobacterium* and the single species genus, *Allorhizobium*. The tree constructed in this study supports the separate genus status of *Allorhizobium*. *Agrobacterium vitis* was more related to *Allorhizobium undicola* than to the *Agrobacterium* type species as found by De Lajudie *et al.* (1998). A need for the revision of the taxonomic position of *A. vitis* was pointed out by De Lajudie *et al.* (1998).

Rhizobium giardinii formed a separate lineage as found by other researchers (Amarger *et al.*, 1997). This species might constitute a possible novel genus (Laguerre *et al.*, 2001). *R. galegae*, *R. huautlense* and *R. loessense* formed a well-supported clade within the *Rhizobium* group. Several authors argued for the formation of a possible new genus for *R. galegae* and its closest neighbours. However, Wei *et al.* (2003) suggested the inclusion of more *R. galegae*-related bacteria before a decision is made on the taxonomic position of these strains. In the phylogenetic tree reconstructed in this study, it would seem that *R. galegae*, *R. huautlense* and *R. loessense* clearly belong to the *Rhizobium* genus. However, this association might change with the addition of more sequences, since the *Rhizobium* clade (including *R. gallicum*, *R. mongolense*, *R. sullae*, *R. indigoferae*, *R. yanglingense*) in which *R. galegae*, *R. huautlense* and *R. loessense* cluster, is not supported by a significant bootstrap value and thus

not stable. The topology of the tree could also be influenced by the algorithm used to construct the tree (Young *et al.*, 2001).

The two *Cyclopia* isolates UCT 50 and Cplic 1 clearly belong to the species *R. tropici*, which is phylogenetically closer related to *Agrobacterium rhizogenes* than to the other *Rhizobium* species. This high sequence similarity between *R. tropici* and *A. rhizogenes* has been well documented (Laguerre *et al.*, 1994; Khbaya *et al.*, 1998; Terefework *et al.*, 1998; Wei *et al.*, 2003).

Except for isolate Cpub 4, the rest of the *Cyclopia* isolates (UCT 42, UCT 53 and Cmey 1) belong to the genus *Bradyrhizobium*. It is clear that the isolates do not belong to one of the described species, but rather to two genomic species, *Bradyrhizobium* genomic sp. A and *Bradyrhizobium* genomic sp. O. The genomic species are related to *B. japonicum* (Lafay and Burdon, 1998). The *Bradyrhizobium* genomic sp. A strains were the dominant genotype isolated in the study of Lafay and Burdon (1998) in their extensive study to determine the rhizobia nodulating indigenous scrubby legumes in Southeastern Australia. Only six of the 745 rhizobial strains isolated in the study of Lafay and Burdon (1998) belonged to *Bradyrhizobium* genomic sp. O. The finding of the same *Bradyrhizobium* genospecies in South Africa and Australia again shows that a specific chromosomal genotype is not restricted to one continent.

The 16S rDNA sequences of the *Bradyrhizobium* genus display high similarity values, which makes it difficult to resolve the close relationships within the genus (Van Berkum and Fuhrmann, 2000). The use of DNA-DNA hybridisation will aid to clarify the position of the *Bradyrhizobium* isolates from *Cyclopia* spp., since this method is still an important criterion for species delineation (Stackebrandt and Goebel, 1994). It is clear that an extensive revision of the genus *Bradyrhizobium* is inevitable. Phylogenetic trees show the polyphyletic nature of the genus *Bradyrhizobium*, since the species *B. japonicum*, *B. liaoningense* and *B. yuanmingense* are more related to genera such as *Afipia* and *Blastobacter* than *B. elkanii*. The existence of several genomic species in the genus shows the huge scope for taxonomic revision of this genus.

Isolate Cpub 4, '*Kaistia adipata*' and the alpha-Proteobacteria strain isolated from agricultural soil might possibly belong to a new species, since they share more than 97% sequence

homology. However, sequence homology of 97% is not a guarantee for species identity, since species might share high sequence similarity, but differ significantly in their genotype, resulting in low DNA homology values (Stackebrandt and Goebel, 1994). DNA-DNA hybridisations and the determination of the G + C content would have to be done to determine the possible species status of these strains, since these techniques are required for species description (Wayne *et al.*, 1987).

The determination of the *nodA* sequence was an easy method to study host plant range. The *nodA* gene is a single copy gene (Haukka *et al.*, 1998) and direct PCR sequencing could thus be done. The topology of the *nodA* tree is in broad agreement with that obtained by other researchers (Haukka *et al.*, 1998). It is quite clear from the *nodA* tree that all the *Cyclopia* isolates contain the same conserved *nodA* gene. As was previously found by other authors, it is clear that different chromosomal genotypes harbour the same symbiotic genotype (Guo *et al.*, 1999; Zhang *et al.*, 2000; Laguerre *et al.*, 2001). The *Cyclopia* isolates probably acquired the symbiotic genes through horizontal gene transfer as proposed by other researchers (Haukka *et al.*, 1998). All of the five *Cyclopia* isolates have been collected from different geographical positions in the fynbos distribution pattern.

CHAPTER 9

CONCLUSIONS

In the study of the root-nodulating bacteria associated with the indigenous South African scrubby legume genus *Cyclopia*, several novel genotypes have been found. Approximately 42% of the Cyclopia isolates belong to one of the new root-nodulating species in the β-*Proteobacteria*, *Burkholderia tuberum* described by Vandamme *et al.* (2002). The collection of *B. tuberum* strains characterised in this study is the largest reported to date.

Members of the *Burkholderia* genus are highly related based on 16S rDNA sequence analysis, since several species share more than 97% sequence homology with each other. DNA-DNA hybridisation analysis is necessary to delineate species due to the inadequacy of 16S rDNA sequence analysis for this purpose (Stackebrandt and Goebel, 1994). Approximately 52% of the isolates are clearly new *Burkholderia* spp. Based on IGS-RFLP and partial 16S rDNA sequence analyses previously found similar strains will be included in the description of new species. This would be done after DNA-DNA hybridisation analysis. Unfortunately, our laboratory is not equipped to do this study. Substrate utilisation patterns of several representative strains of each possible novel species should also be done. The inclusion of several strains in the description of a new species is of utmost importance to refrain from describing species based on a single strain.

The finding that most of the root-nodulating isolates associated with the *Cyclopia* host plants belong to the β-*Proteobacteria* implies that co-evolution of the *Cyclopia* host plants and *Burkholderia* spp. has taken place for quite while.

All of the isolates, which belonged to the α-*Proteobacteria* with the exception of one strain (Cpub 4), belonged to the acid-tolerant *Bradyrhizobium* genus and the acid-tolerant *Rhizobium* species, *R. tropici*. The soils from the sites were mostly acidic and the finding of these rhizobia was thus not surprising.

One isolate (Cpub 4) and two other isolates, whose sequences were obtained from the GenBank database, might be members of a novel genus in the α-*Proteobacteria*. Future work should include phenotypic and more genotypic analyses of all three isolates and if possible the inclusion of similar strains to satisfy the requirements for the description of new species and genera.

The symbiotic genes of all the isolates in this study have been found to be highly conserved. Different chromosomal genotypes harbour the same symbiotic genotype, which suggests that horizontal gene transfer occurred between these root-nodulating organisms.

The focus shift from studying only rhizobia associated with agricultural crops to the characterisation of root-nodulating rhizobia of legumes indigenous to a given geographical region, led to the description of novel rhizobial and seemingly unrelated isolates. This would undoubtedly continue and would help to gain more information on the legume-rhizobium symbiosis and the evolutionary mechanisms involved.