Presence of diverse rhizobial communities responsible for nodulation of common bean (Phaseolus vulgaris) in South African and Mozambican soils.

FEMS Microbiology Ecology, 93, 2017, fiw236

doi: 10.1093/femsec/fiw236
Advance Access Publication Date: 2 December 2016
Research Article

RESEARCH ARTICLE

Presence of diverse rhizobial communities responsible for nodulation of common bean (*Phaseolus vulgaris*) in South African and Mozambican soils

Mwajuma K. Zinga[1], Sanjay K. Jaiswal[2] and Felix D. Dakora[2,*]

[1]Department of Crop Sciences, Tshwane University of Technology, Pretoria 0001, South Africa and
[2]Department of Chemistry, Tshwane University of Technology, Arcadia Campus, Pretoria 0001, South Africa

*Corresponding author: Department of Chemistry, Tshwane University of Technology, Arcadia Campus, 175 Nelson Mandela Drive, Private Bag X680, Pretoria 0001, South Africa. E-mail: dakorafd@tut.ac.za

One sentence summary: Diverse communities of common bean rhizobia in South Africa and Mozambique.

Editor: Angela Sessitsch

ABSTRACT

The diversity and phylogeny of root-nodule bacteria isolated from common bean grown in Mozambique and different provinces of South Africa was studied by restriction fragment length polymorphism (RFLP) and phylogenetic analysis. The combined restriction banding pattern of 16S rRNA and *nifH* profile-generated dendrogram grouped all test isolates into four major clusters with XXI restriction groups and three clusters with VIII restriction groups. Location-based clustering was observed with the 16S rRNA RFLP analysis. Phylogenetic analysis of 16S rRNA, *glnII*, *gyrB* and *gltA* sequences showed that common bean was nodulated specifically by *Rhizobium etli* in Mozambican soils, and by a diverse group of *Rhizobium* species in South African soils (e.g. *R. etli, R. phaseoli, R. sophoriradicis, R. leucaenae* and novel group of *Rhizobium* spp.). Isolates from the Eastern Cape region of South Africa were dominated by *R. leucaenae*. Overall, the results suggested high nodulation promiscuity of common bean grown in Southern Africa. The *nifH* and *nodC* sequence analysis classified all the test isolates with *R. etli* group, except for isolates TUTPVSA117, TUTPVSA114 and TUTPVSA110 which delineated with *R. tropici* group. This finding was inconsistent with the phylogram of the housekeeping genes, and is probably an indication of horizontal gene transfer among the *Rhizobium* isolates tested.

Keywords: RFLP; housekeeping genes; 16S rRNA gene; nodulation; novel isolates; horizontal gene transfer

INTRODUCTION

Common bean is one of the most widely distributed food legumes in the tropics, sub tropics and temperate regions (Singh 1999). The domestication of common bean is believed to have occurred about 4000 years ago (Martinez-Romero 2003). The Mesoamerica and Andean region of South America are believed to be the centres of origin and domestication of the crop (Grange and Hungria 2004). Latin America has the largest land area (50%) under common bean cultivation in the world, followed by Africa with 25% (Beebe *et al.* 2013). In Africa, common bean was established as a pulse before the colonial era. It is now a major staple food legume in Eastern and Southern Africa. It is also recognised as the second most important source of human protein after soybean (Broughton *et al.* 2003), and the third most important source of calories after soybean and groundnut (Aserse *et al.* 2012). More than 100 million people in Eastern and Southern Africa depend on common bean diet (Aserse *et al.* 2012). The cultivation of common bean in West and North Africa is minor compared to its cultivation in the Eastern, Central and Southern Africa, where cultivation is concentrated at an altitude of

Received: 22 June 2016; **Accepted:** 21 November 2016

1000 m above sea level (Beebe *et al.* 2013). In Africa, Tanzania (676 000 MT) has been reported to be the largest bean producer, followed by Kenya (578 000 MT), Uganda (464 000 MT) and South Africa (42 000 MT) (USDA 2012).

Common bean is a promiscuous host that is nodulated by a wide variety of rhizobia for N_2 fixation (Kaschuk *et al.* 2006). The promiscuity involved in common bean symbiosis represents an interesting model for studies of rhizobial species. Cropping and/or nodulation history can have a significant impact on rhizobial diversity in soils (Nkot *et al.* 2008; Mothapo *et al.* 2013). The use of fertilisers, pesticides and herbicides can also alter rhizobial diversity in different agroecologies with different cropping history (Farooq and Vessey 2009)

Characterisation of rhizobia is increasingly becoming complex as strain identity has to be revised periodically due to the discovery of new genera and species. The Ad Hoc Committee for re-evaluation of species definition has recommended the use of 16S rRNA, housekeeping genes, symbiotic genes and DNA–DNA hybridisation as the molecular criteria for species delineation (Stackebrandt *et al.* 2002; Pongsilp 2012). Although sequence analysis of the 16S rRNA gene is most commonly used in bacterial taxonomy for measuring the degree of relatedness between organisms above the species level (Martens *et al.* 2008), the 16S rRNA gene sequence often lacks resolving power at and below the species level. To overcome these limitations, protein-encoding genes with higher level of sequence divergence than rRNA genes are used (Valverde *et al.* 2006; Martens *et al.* 2007, 2008; Aserse *et al.* 2012; Peix *et al.* 2015; Rouhrazi, Khodakaramian and Velázquez 2016). Combined data from these genes provide a global and more reliable overview of interorganismal relationships. Although it is agreed that phylogenetic analysis based on stable chromosomal genes is essential, studying the phylogeny of symbiotic genes provides information on the symbiotic properties of rhizobia. It is not included merely for the description of a new rhizobia. The identification and biodiversity of common bean-nodulating rhizobia from North African (Mnasri *et al.* 2007), East African (Anyango *et al.* 1995; Beyene *et al.* 2004; Aserse *et al.* 2012) and West African soils (Diouf *et al.* 2000) have been reported. However, to our knowledge, there is no information on bean-nodulating rhizobia from the Southern African region.

Therefore, the aim of this study was to characterise and evaluate the taxonomic position of common bean-nodulating rhizobial isolates from South Africa and Mozambique using restriction fragment length polymorphism (RFLP) analysis of 16S rRNA, *nifH* and phylogenetic analysis of 16S rRNA, *nifH* and *nodC* together with three housekeeping genes (*glnll*, *gyrB* and *gltA*).

MATERIALS AND METHODS

Nodule collection and isolation of rhizobia

Nodules and soil samples were collected from common bean plants grown on farmers' fields in Mozambique and different locations of Limpopo and Mpumalanga Provinces of South Africa (Table 1; Fig. 1). The soil samples were analysed at the Institute for Plant Production, Elsenburg, in the Western Cape. The nodules were sampled at the flowering stage of common bean plants. Rhizobia were isolated from surface-sterilised nodules, as described by Somasegaran and Hoben (1994).

Nodulation bioassay

The nodulation ability of isolates was tested by inoculating seedlings of commercial Brazilian black common bean (the homologous host) with the culture of each isolate in the glasshouse under strict microbiologically controlled conditions as described by Somasegaran and Hoben (1994). Seedlings were inoculated with 2 ml bacterial suspension ($\approx 10^7$–10^8 cells/ml). Three replicate pots were used per isolate, and plants irrigated three times a week using sterile (Broughton and Dilworth 1971) N-free nutrient solution. Three uninoculated pots were included as control. After 6 weeks, the plants were harvested and observed for presence or absence of root nodules.

Rhizobial DNA isolation and PCR amplification of 16S rRNA and *nifH* gene region

Rhizobial DNA was extracted using GenElute Bacterial Genomic DNA kit, according to the manufacturer instructions (Sigma Aldrich, USA). The 16S rRNA and *nifH* regions of bacterial genomic DNA was amplified in 25 μL reaction mixture containing 5x My Taq PCR buffer, Taq polymerase (5U) (Bioline, USA), 10 pM of each forward and reverse primers and 40–50 ng DNA as template. DNA amplifications were performed by Thermal cycle (T100 BIORAD, USA) with respective primers and standard temperature profile (Table 2).

RFLP analysis of PCR-amplified 16S rRNA and *nifH* regions

The PCR-amplified 16S rRNA region products were digested with different four-base cutting restriction enzymes (namely, *Hae*III, *Hpa*II and *Rsa*I), whereas *Msp*I and *Alu*I were used to digest *nifH*-PCR-amplified region, following the procedure recommended by the manufacturer (Thermo Scientific, Lithuania). The digested products were separated by electrophoresis in 2.5% agarose gel. Banding patterns were scored directly from gel photographs, and the isolates grouped through visual inspection of the banding.

RFLP cluster analysis of 16S rRNA, and *nifH* regions

Only distinct, well-resolved and unambiguous bands were scored, faint bands and band sizes $\leq$50 bp were not included for cluster analysis. The restriction enzyme-digested fragments were scored as (1) for the presence of and (0) for the absence of homologous bands. Thereafter, the similarity of test strains was evaluated by simple matching Jaccard's similarity coefficient with the help of NTSYSpc 2.1 software (Rohlf 2009), and a dendrogram constructed from the distance matrix by means of the unweighted pair group method with arithmetic mean algorithm (UPGMA).

Canonical correspondence analysis to understand relationships between environmental factors and *Rhizobium*

The canonical correspondence analysis (CCA) was used to assess the rhizobial distribution in response to environmental variables (temperature, rainfall and pH) using 'vegan' in R statistical package as described by Li *et al.* (2012).

PCR amplification of housekeeping genes (*glnII*, *gltA*, and *gyrB*) and nodulation gene (*nodC*)

The PCR amplification of *glnII*, *gltA*, *gyrB* and *nodC* genes of rhizobial genome was done, as described for 16S rRNA and *nifH* genes. The primers used and thermal cycle conditions are listed in

Table 1. Climate, cropping history and soil properties of common bean nodule collection sites.

Country	Province	Village	Geographical positions	Annual temp. (°C)	Annual rainfall (mm)	pH	Cropping history	Soil nutrient (mg/kg)										
								B	Ca	Cu	Fe	K	Mg	Mn	Na	P	S	Zn
South Africa	Limpopo	Ofcolaco	24° 06′ 00″ S 30° 23′ 00″ E	20–30	238	5.9	Maize	0.18	432	4.28	234.8	258	139	338	4	60	6.6	5.37
		Sekororo	24° 17′ 00″ S 30° 22′ 00″ E	20–30	238	6.4	Maize	0.19	1130	8.81	165.8	739	457	325	85	35	20	3.05
		Mankweng	23.886° S 29.718° E	17–28	206	7.1	Maize	0.26	444	2.69	139.3	127	315	138	39	24	3.1	1.42
		Vallis	24° 8′ 2.83″ S 30° 7′ 27.19″ E	10–26	82.6	6.0	Maize	0.20	1142	4.05	213	274	388	2648	8	27	1.4	2.74
		Fertilis	24.123° S 30.106° E	14–26	294	6.7	Maize	0.28	1634	9.31	228	404	425	3490	9	42	1.6	4.07
	Eastern Cape	Lujeciweni	32.153° S 26.447° E	11–25	226	4.2	Maize	0.17	282	2.74	171	171	74	211	9	25	3.1	1.76
		Tikitiki	31° 24′ 00″ S 28° 42′ 00″ E	11–25	226	4.1	Fallow	0.22	416	2.18	120	64	164	128	20	13	5.2	0.65
		Mfabantu	31° 35′ 00″ S 29° 00′ 00″ E	15–26	190	5.3	Maize	0.16	762	3.02	128	94	329	210	35	34	2.6	0.90
Mozambique	Gurue	Mutequelesse (Tissi)	15° 19′ 9.5″ S 36° 42′ 43.9″ E	20–27	584	6.5	–	–	–	–	–	–	–	–	–	–	–	–
		Lioma	15° 10′ 12″ S 36° 48′ 17″ E	11–25	622	6.5	–	–	–	–	–	–	–	–	–	–	–	–

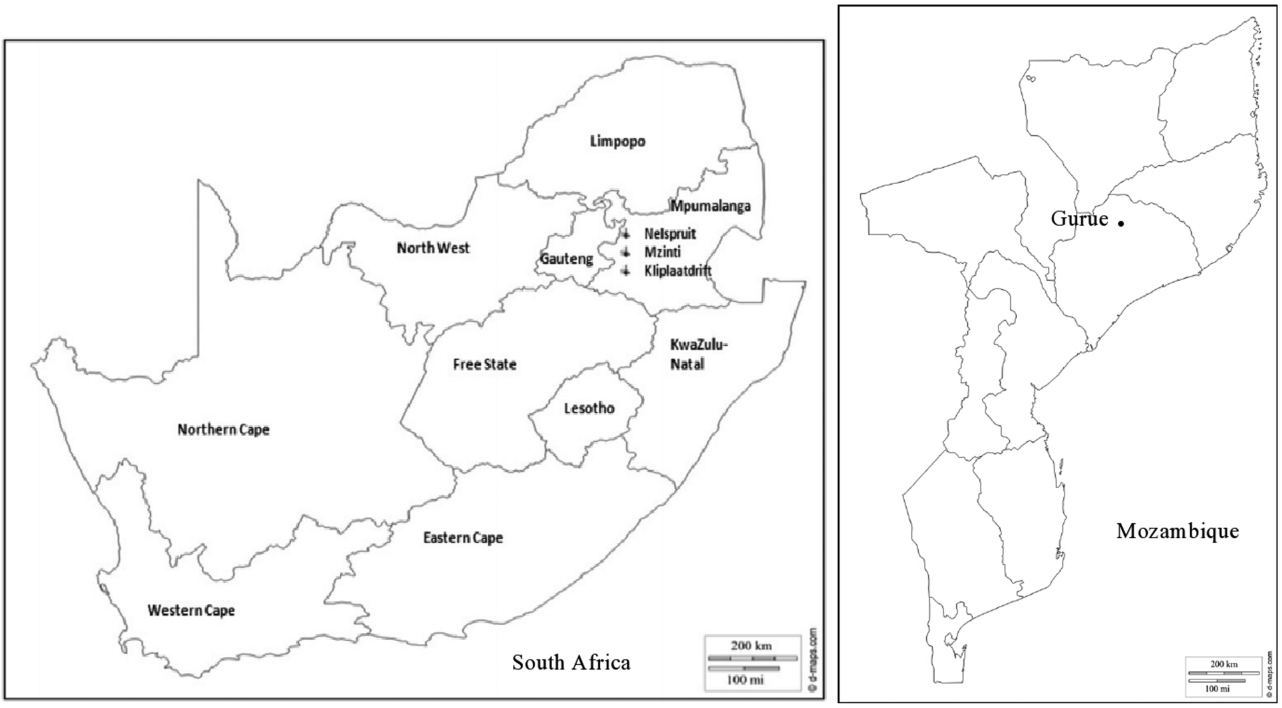

Figure 1. Map of study sites in South Africa (Limpopo and Eastern Cape), and Mozambique (Gurue).

Table 2. Primers and PCR temperature profiles used for DNA amplification.

Primers	Sequence	Temperature profile	References
16S rDNA	fD1 (9-29)AGAGTTTGATCCTGGCTCAG rD1 (1474-1494)CTTAAGGAGGTGATCCAGCC	5 min 95°C, 35 × (1 min 95°C, 1 min 55°C, 1 min 72°C), 10 min 72°C	Weisburg *et al.* (1991)
nifH F *nifH* R	5′TACGGNAARGGSGGNATCGGCAA3′ 5′AGCATGTCYTCSAGYYTCNTCCA3′	5 min 94°C, 20 × [30 s 94°C, 30 s 65°C (—0.5°C/cycle),90 s 72°C], 25 × (30 s 94°C, 30 s 55°C, 90 s 72°C), 10 min 72°C	Nzoué *et al.* (2009)
glnll F *glnll* R	5′AAGCTCGAGTACATCTGGCTCGACGG3′ 5′SGAGCCGTTCCAGTCGGTGGTGTCG3′	2 min 95°C, 35 × (45 s 95°C, 30 s 65°C, 90 s 72°C), 10 min 72°C	Stepkowski. *et al.* (2011)
gyrB F *gyrB* R	5′TTCGACCAGAAYTCCTAYAAGG3′ 5′AGCTTGTCCTTSGTCTGCG3′	10 min 95°C, 35 × (30 s 94°C, 30 s 58°C, 1 min 72°C), 10 min 72°C	Marek-Kozaczuk *et al.* (2013)
gltA F *gltAR*	5′CSGCCTTCTAYCAYGACTC3′ 5′GGGAGCCSAKCGCCTTCAG3′	5 min 94°C, 20 × [30 s 94°C, 30 s 63°C (-0.5°C\cycle), 90s 72°C], 25 X (30 s 94°C, 30 s 54.5°C, 90s 72°C), 10 min 72°C	Stepkowski *et al.* (2011)
nodC f *nodC* R	5′TGA TYG AYA TGG ART AYT GGC T3′ 5′CGY GAC ARC CAR TCG CTR TTG3′	30 s 94°C, 40 × [30 s 94°C, 1 min 53°C, 30 s 72°C], 5 min 72°C	Sarita *et al.* (2005)
nodC f *nodC* R	5′AYG THG TYG AYG ACG GTT C3′ 5′CGY GAC AGC CAN TCK CTA TTG3′	30 s 94°C, 40 × [30s 94°C, 1 min 55.4°C, 30s 72°C], 5 min 72°C	Laguerre *et al.* (2001)

Table 2. The same procedures used for 16S rRNA and *nifH* genes were followed to examine the PCR-amplified product on agarose gel.

Purification of PCR product, sequencing and phylogenetic analysis

The PCR-amplified products of 16S rRNA, symbiotic gene (*nifH* and *nodC*) and housekeeping loci (*glnll*, *gltA* and *gyrB*) were purified by GeneJET PCR purification kit (Thermo Scientific, Lithuania). The purified samples were sequenced (Macrogen, Netherlands), and the quality of all sequences checked using BioEdit 7.0.0 software (Hall 2004). NCBI GenBank databases were used to identify closely related species with test isolates using the BLASTn program. The sequences were deposited in the NCBI GenBank database to get accession numbers (Table S1, Supporting Information). Reference (type) sequences were selected from NCBI database to align with sequences of the test rhizobial isolates using MUSCLE (Edgar 2004) for construction of phylogenic tree by MEGA 6.0 program (Tamura *et al.* 2013). Phylogenetic trees were generated by the Kimura 2-parameter model to calculate evolutionary distances (Kimura 1980), and evolutionary history was inferred using the neighbour-joining method (Saitou and Nei 1987) algorithm with 1000 bootstraps

(Felsenstein 1985). The nucleotide information was obtained in terms of conserved, variable, parsimony-informative and singleton regions were analysed in consensus sequences. A parsimony-informative contains more than two types of nucleotide with at least two of them occurring with a minimum frequency, whereas singleton site contains at least two types of nucleotides with, at most, one occurring multiple times.

RESULTS

Soil characteristics

All the soils from the Eastern Cape Province were highly acidic (pH 4.1–5.3) in reaction, while those from Limpopo Province of South Africa and Gurue province of Mozambique were slightly acidic to neutral (Table 1). Soils analysed from the studied regions of South Africa showed that available phosphoros (P) content ranged from 13–60 mg/kg, calcium (Ca) 282–1634 mg/kg, copper (Cu) 2.18–9.31 mg/kg, iron (Fe) 120–234.8 mg/kg, potassium (K) 64–739 mg/kg, magnesium (Mg) 74–457 mg/kg, manganese (Mn) 128–3490 mg/kg, sodium (Na) 4–85 mg/kg, Sulphur (S) 1.4-20 mg/kg, zinc (Zn) 0.65-5.37 mg/ kg and boron (B) 0.16-0.28 mg/kg (Table 1).

PCR-RFLP analysis of 16S rRNA gene

A total of 69 samples taken for PCR amplification of full-length 16S rRNA gene from rhizobial genome yielded a single band of about 1500 bp as amplified product. The amplified products were digested with restriction enzymes *Hae*III, *Hpa*II and *Rsa*I, which produced polymorphic banding pattern on the gel. Restriction enzymes *Rsa*I and *Hae*III showed much greater ability to discriminate between isolates, and produced eight (A–H) restriction types (Table 3), followed by *Hpa*II which yielded five (A–E) (Table 3). The combined restriction banding pattern profile was used to generate a dendrogram for RFLP cluster analysis (Fig. 2). Four major clusters were generated with the XXI restriction groups (Fig. 2; Table 3) at 0.01 to 1.00 Jaccard's similarity coefficients. Clusters 1, 3 and 4 were very specific and consisted of isolates from Ofcolaco, Lujecweni and Mankweng villages in Limpopo. Cluster 2 was the biggest cluster formed from 8 (V–XII) restriction groups and they contained isolates from all sampled sites/villages. Isolates from Gurue in Mozambique were mostly grouped with those from Limpopo in South Africa. Restriction groups V, VI, VII, IX, XI and XII contained all the isolates from Limpopo (South Africa) and Gurue (Mozambique) (Fig. 2). Seven isolates (TUTPVSA1, TUTPVSA6, TUTPVSA9, TUTPVSA34, TUTPVSA40, TUTPVSA121 and TUTPVMZQ123) with *Hpa*II, four isolates (TUTPVSA25, TUTPVMZQ125, TUTPVMZQ126 and TUTPVMZQ127) with *Hae*III and ten isolates (TUTPVSA2, TUTPVSA4, TUTPVSA14, TUTPVSA18, TUTPVSA22, TUTPVSA24, TUTPVSA26, TUTPVSA38, TUTPVSA51 and TUTPVSA115) with *Rsa*I did not show any digestion and stood as a single band.

PCR-RFLP analysis of *nifH* symbiotic gene

The PCR-amplified 800 bp of *nifH* region of the test rhizobial isolates was digested with restriction enzymes *Msp*I and *Alu*I, and they revealed the presence of polymorphic bands. Five (A–E) and only two (A-B) restriction patterns formed respectively with *Msp*I and *Alu*I were found in all the rhizobial isolates tested (Table 3). A dendrogram was generated based on the combined *nifH* restriction pattern of bacterial isolates by the UPGMA algorithm (Fig. 3). All isolates were grouped into three clusters. While isolates TUTPVSA38 and TUTPVSA25 stood alone independently, three isolates (TUTPVSA1, TUTPVSA24 and TUTPVSA32) did not digest with *Msp*I restriction enzyme.

Canonical correspondence analysis

The 16S rRNA-RFLP of the isolates and environmental data (rain fall, temperature and pH) from the sites where nodules were collected were used to assess the relationship between environmental factors and rhizobia (Fig. 4). The CCA results showed that 33.3% were explainable variable, while about 66.6% of the total variation was unexplainable. The CCA biplot reflected that temperature showed the strongest correlation followed by rainfall and pH to the variation in 16S rRNA–RFLP pattern.

DNA sequencing and phylogenetic analysis of 16S rRNA

A total of 23 rhizobial isolates were selected from the results of 16S rDNA and *nifH* RFLP dendrogram for sequence analysis. Based on sequence quality, only 16 of the 23 isolates were included for 16S rRNA phylogenetic analysis. Sixteen isolates with 500–775 nucleotides were aligned, and used to build the phylogenetic tree together with reference type strains. The mean frequency of T, C, A and G in sequences was 20.7%, 25.2%, 23.3% and 30.8%, respectively. The consensus sequences of all *Rhizobium* strains contained 521 analysed sites which had 406 conserved, 112 variables, 53 parsimony-informative and 59 singleton sites (Table 4). The 16S rRNA phylogenetic tree branched into five clusters (I–V) under the genera *Rhizobium*. In cluster II, isolates TUTPVSA101 and TUTPVSA24 from Ofcolaco and Vallis villages respectively in Limpopo aligned with *Rhizobium leguminosarum* group and its allied species, while isolates TUTPVSA5, TUTPVSA16, TUTPVSA42, TUTPVSA104, TUTPVSA107, TUTPVSA121 and TUTPVSA6 from Limpopo and the Eastern Cape in cluster I formed a monophyletic group without any *Rhizobium* type strains. Isolate TUTPVSA81 from Mankweng village in Limpopo grouped with *R. sophoriradicis* in cluster III; isolates TUTPVMZQ124, TUTPVMZQ125 and TUTPVMZQ128 from Mozambique also grouped together in cluster IV with *R. binae*; and two isolates (TUTPVSA120 and TUTPVSA117) from the Eastern Cape grouped with *R. tropici* and its allied species in cluster V (Fig. 5).

Identification and phylogenetic analysis of housekeeping gene (*glnII*, *gyrB* and *gltA*) sequences

The phylogenetic relationship of *glnII*, *gyrB* and *gltA*, which respectively code for glutamine synthetase II, DNA gyrase and citrate synthase I, was determined. The PCR amplification of *glnII*, *gyrB* and *gltA* yielded about 600, 500 and 750 bp, respectively, each with a single band for the selected rhizobial population.

Based on partial gene sequences of housekeeping gene comparisons with the Genbank references, all the test isolates were identified as *Rhizobium* sp. These protein-encoding housekeeping loci showed variable informative positions (Table 4). Several housekeeping gene sequences obtained from GenBank were shorter than those of test isolate sequences. As a result, the final alignments of the new sequences of the test isolates were omitted, and the shorter lengths of the housekeeping gene nucleotide sequences of test isolates used for phylogenetic analysis. The lowest level of conserved sequences of only 49.5% was observed for *gyrB*. For the other sequences, the levels of conservation were 65.7% for *gltA* and 65.8% for *glnII* regions, while *gyrB*

Table 3. Restriction pattern of 16S rRNA and *nifH* genes of *P. vulgaris*-nodulating rhizobial isolates.

Isolate	16S rRNA restriction pattern type			Combined 16S rRNA restriction type	*nifH* restriction pattern type		Combined *nifH* restriction type	Location
	*Hae*III	*Hpa*II	*Rsa*I		*Msp*I	*Alu*I		
TUTPVSA 1	A	nd	A	I	nd	A	I	Ofcolaco
TUTPVSA 2	B	C	nd	V	A	A	I	Mankweng
TUTPVSA 3	C	C	H	XVI	B	A	III	Mankweng
TUTPVSA4	B	B	nd	VIII	A	A	I	Sekororo
TUTPVSA 5	D	C	B	VII	A	A	I	Mankweng
TUTPVSA 6	B	nd	B	V	A	A	I	Mankweng
TUTPVSA 7	B	C	B	V	A	A	I	Mankweng
TUTPVSA 9	B	nd	B	V	A	A	I	Ofcolaco
TUTPVSA 12	B	A	B	VI	A	A	I	Ofcolaco
TUTPVSA13	B	A	B	VI	A	A	I	Ofcolaco
TUTPVSA 14	A	A	nd	I	A	A	I	Ofcolaco
TUTPVSA 15	B	A	B	VI	A	A	I	Mankweng
TUTPVSA 16	B	A	A	IX	A	A	I	Ofcolaco
TUTPVSA 18	B	A	nd	VI	B	A	III	Sekororo
TUTPVSA 21	A	A	G	II	A	A	I	Ofcolaco
TUTPVSA22	B	B	nd	VIII	B	A	III	Mankweng
TUTPVSA 24	A	B	nd	III	nd	A	IV	Ofcolaco
TUTPVSA25	nd	C	A	XIII	A	B	VII	Ofcolaco
TUTPVSA 26	A	D	nd	VII	B	A	VIII	Ofcolaco
TUTPVSA 30	A	B	D	III	B	A	VIII	Ofcolaco
TUTPVSA 31	E	D	C	XVII	B	A	III	Mankweng
TUTPVSA 32	F	D	C	XVIII	nd	A	III	Mankweng
TUTPVSA 33	G	D	C	XIX	B	A	III	Mankweng
TUTPVSA 34	E	nd	C	XVII	A	A	I	Mankweng
TUTPVSA 36	E	D	C	XVII	B	A	III	Mankweng
TUTPVSA38	H	C	nd	XIII	C	A	VI	Mankweng
TUTPVSA40	A	nd	E	IV	A	A	I	Mankweng
TUTPVSA 41	B	E	G	XIV	B	A	III	Sekororo
TUTPVSA 42	A	C	E	IV	A	A	I	Mankweng
TUTPVSA 43	E	D	F	XX	B	A	III	Ofcolaco
TUTPVSA 51	B	C	nd	V	A	A	I	Mankweng
TUTPVSA 55	B	C	B	V	A	A	I	Mankweng
TUTPVSA 66	B	C	B	V	A	A	I	Ofcolaco
TUTPVSA 68	B	C	A	XII	A	A	I	Ofcolaco
TUTPVSA 69	B	C	B	V	A	A	I	Ofcolaco
TUTPVSA 70	B	C	A	XII	A	A	I	Ofcolaco
TUTPVSA 71	B	C	B	V	A	A	I	Ofcolaco
TUTPVSA 75	B	C	A	XII	A	A	I	Ofcolaco
TUTPVSA 81	B	C	A	XII	A	A	I	Mankweng
TUTPVSA 83	B	C	A	XII	A	A	I	Ofcolaco
TUTPVSA84	B	C	B	V	A	A	I	Mankweng
TUTPVSA 91	E	D	E	XXI	A	A	I	Ofcolaco
TUTPVSA 95	B	C	A	XII	A	A	I	Mankweng
TUTPVSA101	B	C	A	XII	A	A	I	Vallis
TUTPVSA102	B	C	A	XII	A	A	I	Vallis
TUTPVSA103	B	C	A	XII	A	A	I	Vallis
TUTPVSA104	B	C	B	V	A	A	I	Vallis
TUTPVSA105	B	C	B	V	A	A	I	Vallis
TUTPVSA106	B	C	B	V	A	A	I	Vallis
TUTPVSA107	B	C	B	V	A	A	I	Fertlis
TUTPVSA108	B	C	A	XII	A	A	I	Fertlis
TUTPVSA109	B	C	B	V	A	A	I	Fertlis
TUTPVSA110	B	C	B	V	A	A	I	Fertlis
TUTPVSA111	B	C	B	V	A	A	I	Fertlis
TUTPVSA112	B	C	A	XII	A	A	I	Fertlis
TUTPVSA113	B	C	B	V	A	A	I	Fertlis
TUTPVSA115	B	E	nd	XIV	B	A	III	Lujeciweni
TUTPVSA117	B	E	C	XV	D	A	IV	Lujeciweni
TUTPVSA118	B	E	C	XV	D	A	IV	Lujeciweni
TUTPVSA114	B	E	C	XV	D	A	IV	Lujeciweni
TUTPVSA120	B	E	A	X	E	A	V	Tikitiki

Table 3. (*Continued*).

Isolate	16S rRNA restriction pattern type			Combined 16S rRNA restriction type	*nifH* restriction pattern type		Combined *nifH* restriction type	Location
	*Hae*III	*Hpa*II	*Rsa*I		*Msp*I	*Alu*I		
TUTPVSA121	B	nd	A	VIII	B	A	III	Mfabantu
TUTPVMZQ123	B	nd	A	VIII	C	A	II	Muteguelesse
TUTPVMZQ124	B	C	A	XII	A	A	I	Lioma
TUTPVMZQ125	nd	C	A	XIII	B	A	III	Lioma
TUTPVMZQ126	nd	C	B	XI	A	A	I	Lioma
TUTPVMZQ127	nd	C	A	XII	A	A	I	Lioma
TUTPVMZQ128	B	C	A	XII	A	A	I	Muteguelesse
TUTPVMZQ129	B	C	A	XII	A	A	I	Lioma

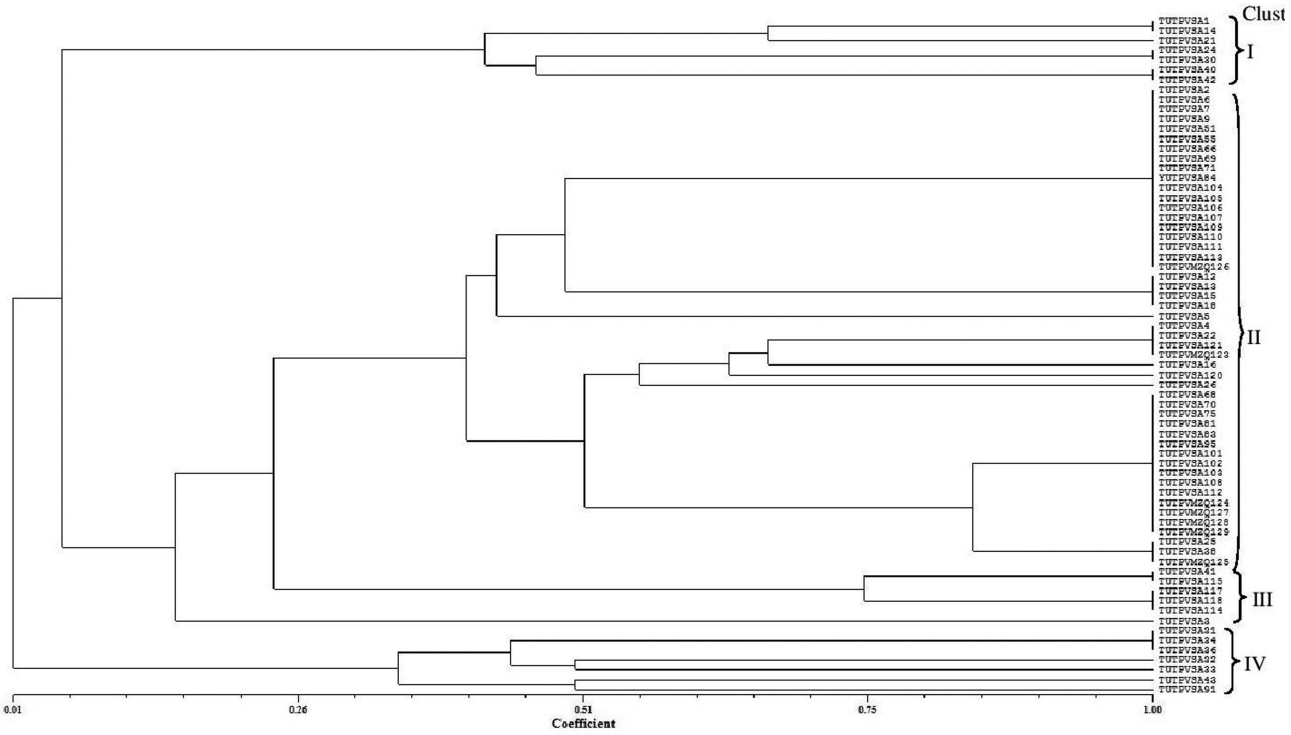

Figure 2. Dendrogram constructed from 16S rRNA-RFLP analysis of common bean-nodulating rhizobial isolates.

recorded the highest variable (50.5%), parsimony-informative (32.1%) and singleton (18.5%) sequence information (Table 4).

The phylogenies of the three housekeeping genes placed the test isolates into five clusters. The *glnII* phylogenetic tree grouped the isolates with *R. etli, R. phaseoli* and *R. leucaenae* (Fig. 6a). Isolates in cluster I, IV and III were proximally related to *R. etli* group. However, isolates TUTPVSA6, TUTPVSA16, TUTPVSA107 TUTPVSA104, TUTPVSA5 and TUTPVSA42 failed to group closely with any reference type strain in cluster I. But in cluster II, isolates TUTPVSA24 and TUTPVSA21 grouped with *R. phaseoli*, while in cluster V, isolates TUTPVSA117, TUTPVSA114 and TUTPVSA120 grouped with *R. leucaenae* (Fig. 6a).

The phylogenetic tree based on *gyrB* and *gltA* gene sequences was in agreement with the delineation by *glnII* gene phylogeny (Fig. 6b and c). However, a minority group represented by isolates TUTPVSA104 and TUTPVSA121 showed some discrepancies, as the closeness of these isolate was different in *gltA* phylogeny.

The aligned sequences of 16S rRNA, *gyrB*, *glnII* and *gltA* housekeeping genes were concatenated, and 2077 sites were used to construct a tree. But a number of type strain sequences and test isolate sequences were excluded from the combined analysis due to the absence of sequences for some housekeeping genes. A neighbour-joining tree based on combined housekeeping partial gene sequences was constructed (Fig. 7). The generated concatenated tree was congruent with the results of individual housekeeping gene phylogenies, and placed the isolates into five clusters (I—V). Cluster I consisted of four isolates (TUTPVSA107, TUTPVSA16, TUTPVSA104, TUTPVSA42) which formed an independent lineage different from *R. etli*. As with the individual gene phylogenies, isolates TUTPVMZQ 124 and TUTPVMZQ 125 from Mozambique grouped with *R. etli* in cluster II with 97% bootstrap support and 96.7%–97.2% sequence identity. Isolate TUTPVSA81 was positioned independently and stood as an outgroup of cluster I, but showed closer relatedness to *R. sophoriradicis* in the 16S rRNA and *glnII* phylograms with 99.5%–100% sequence similarity. But due to the absence of *gyrB* and *gltA* nucleotide sequences of *R. sophoriradicis* in the GenBnak database, we could not include it in the concatenated sequence

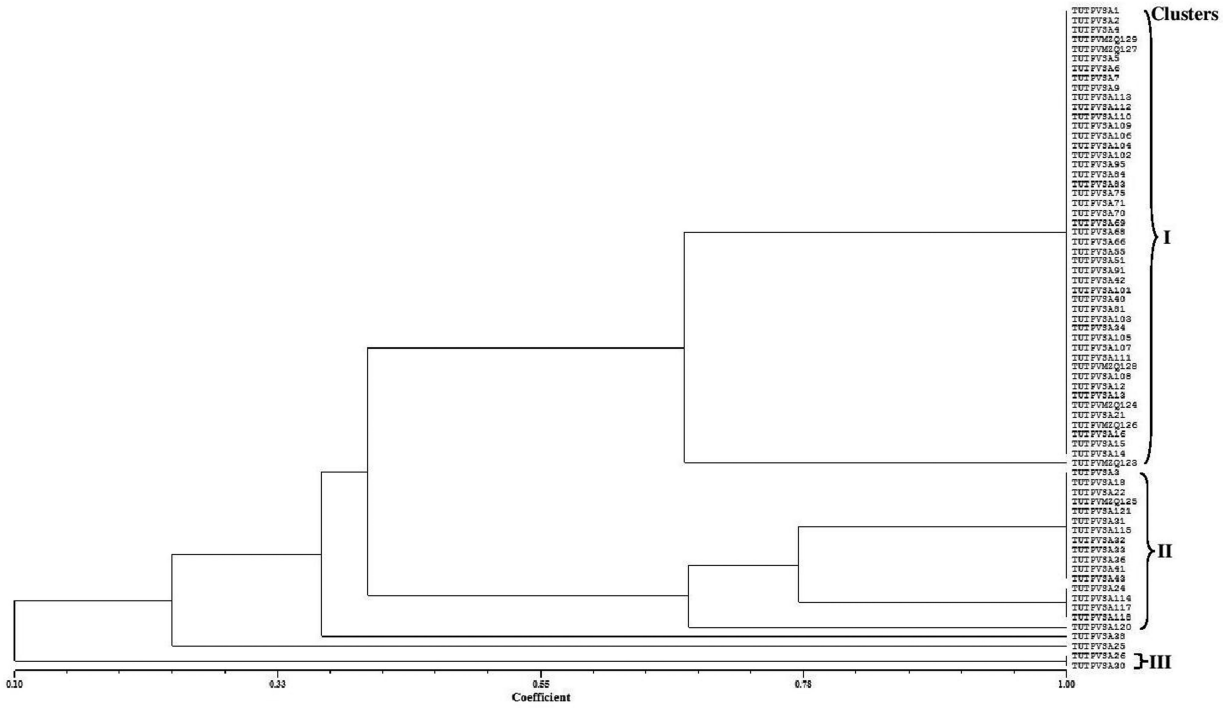

Figure 3. Dendrogram constructed from *nifH*-RFLP analysis of common bean-nodulating rhizobial isolates.

Table 4. Information on nucleotide sequences used in phylogenetic analysis.

Locus	No. of strains used for tree construction	Nucleotide sequence information					
		Cnoserved (C)	Variables (V)	Parsimony-informative (Pi)	Singleton (S)	Total[a]	Frequency T/C/A/G (%)
16S rDNA	50	406 (77.93)	112 (21.50)	53 (10.17)	59 (11.32)	521	20.7/25.2/23.3/30.8
gyrB	33	284 (49.5)	290 (50.5)	184 (32.1)	106 (18.5)	574	22.4/29.5/17.1/31.0
gltA	27	369 (65.7)	193 (34.3)	98 (17.4)	95 (16.9)	562	20.2/27.4/19.0/33.4
glnII	35	276 (65.1)	148 (34.9)	112 (26.4)	36 (8.5)	424	20.0/28.9/17.5/33.6
Concatenated (16S rDNA+*gyrB*+*glnII*+*gltA*)	17	1435 (69.1)	639 (30.8)	327 (15.7)	312 (15.0)	2077	20.8/27.8/19.2/32.2
nifH	38	217(63.8)	120(35.3)	79(23.2)	41(12.1)	340	21.2/30.3/17.8/30.6
nodC	35	223 (44.95)	272 (54.83)	235 (47.37)	37 (6.65)	496	24.4/30.2/18.1/27.2

[a]Mean number of PCR amplified sequences/number of sites used for tree construction.

phylogeny. In cluster IV, isolates TUTPVSA 24 and TUTPVSA21 from Limpopo Province formed another lineage without any close type strains, even though they were very close to *R. phaseoli* in individual gene trees. In cluster V, isolates TUTPVSA117 and TUTPVSA120 from the Eastern Cape Province grouped with *R. leucaenae* with high 100% bootstrap support and sequence similarity.

DNA sequencing and phylogenetic analysis of *nifH* and *nodC* symbiotic gene

The selected isolates produced 600 bp and 900 bp of *nifH* and *nodC* PCR amplified product, respectively, and yielded good quality sequences that were used for the phylogenetic analysis. The sequence alignment results for the both genes are presented in Table 4. Sequence similarity of *nifH* and *nodC* with reference type sequences varied from 98.1% to 100%. The phylogenetic trees of both symbiotic genes categorised the isolates into two monophyletic groups (Figs 8 and 9). From the *nifH* tree, most of the test isolates (TUTPVMZQ128, TUTPVMZQ125, TUTPVSA121, TUTPVSA107, TUTPVSA104, TUTPVSA101, TUTPVSA42, TUTPVSA30, TUTPVSA24, TUTPVSA21, TUTPVSA18, TUTPVSA16, TUTPVSA5, TUTPVMZQ124 and TUTPVSA41) aligned with *R. etli*, *R. vallis* and *R. phaseoli* with 100% sequence similarity and 65% bootstrap support in a major cluster I, while in *nodC*, those isolates did not show any close relationship with reference type strains (Figs 8 and 9). Isolates TUTPVSA114 and TUTPVSA6 stood with *R. sophoriradicis* as an outgroup of cluster I with sequence similarity 99.7%–100% in *nifH* phylogeny. As with the concatenated phylogram, isolate TUTPVSA81 was positioned independently in

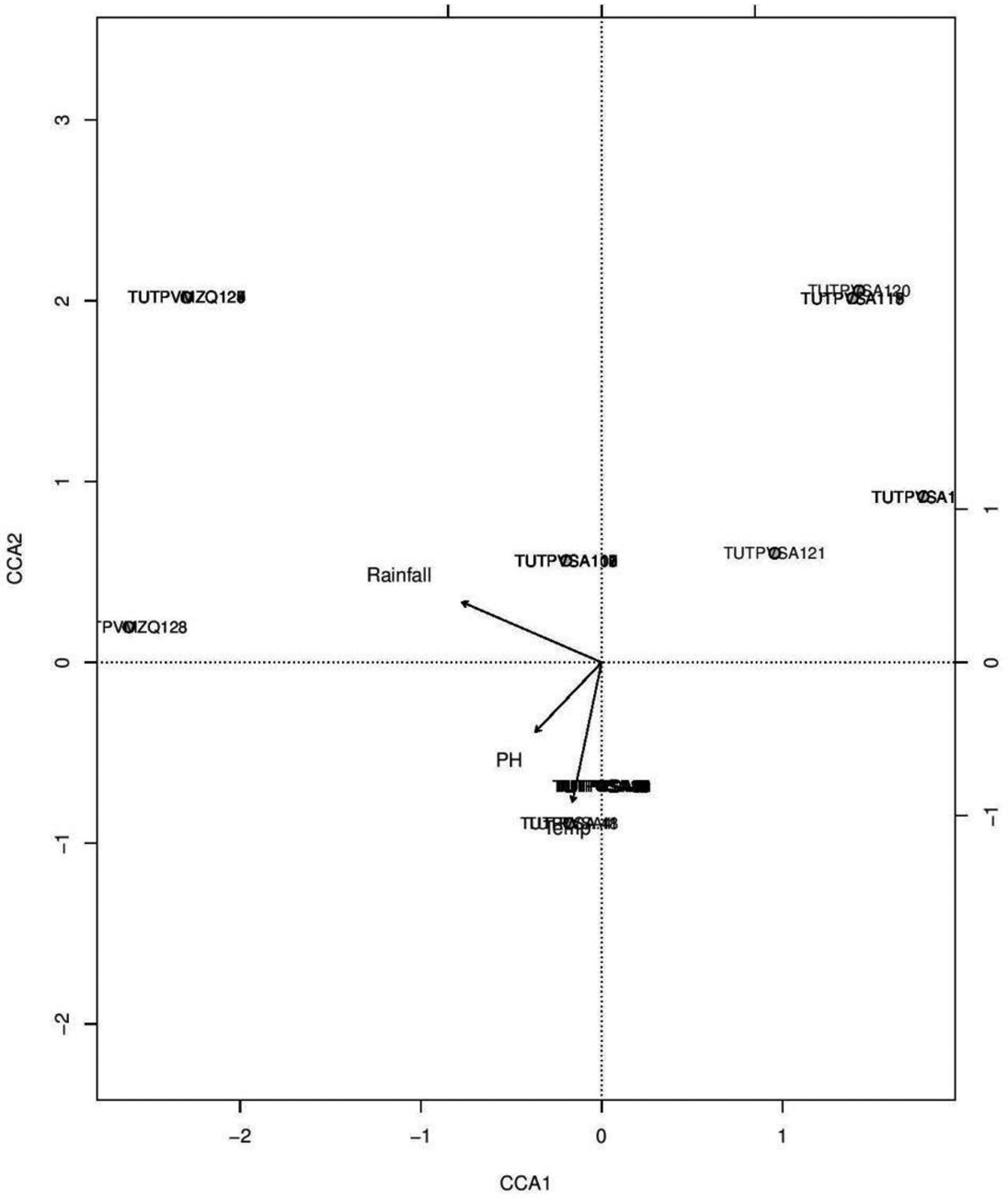

Figure 4. CCA plot representation of association between environmental variables and 16S rRNA-RFLP data of rhizobial isolates.

the *nifH* and *nodC* phylogenies. In cluster II, isolates TUTPVSA117 and TUTPVSA110 shared their sequences with *R. miluonense* and *R. multihospitium* in the *R. tropici* group with 100% sequence similarity and bootstrap support (Fig. 8). However, compared to concatenated and *nodC* phylogenies, isolates TUTPVSA114 and TUTPVSA110 showed discrepancies in *nifH* phylogeny.

DISCUSSION

The diversity of rhizobia nodulating common bean has been widely studied at both the centre of origin and new areas of introduction (Laguerre *et al.* 1996; Pinto, Hungria and Martins Mercante 2007; Martens *et al.* 2008). In this study, the use of restriction endonucleases successfully discriminated between rhizobial isolates based on combined restriction banding patterns of the 16S rRNA region. This indicates the presence of polymorphism, a finding consistent with the results of Pongsilp (2012). The 16S rRNA-RFLP analysis has been used to study the diversity of rhizobial species (Laguerre *et al.* 1996; Odee *et al.* 2002; Shamseldin *et al.* 2005, 2014). The appearance of polymorphism was probably due to gain or loss of restriction sites (Laguerre *et al.* 2001). The dendrogram generated from restriction banding pattern clustered according to location. However, the RFLP technique proved unsuitable for studies of the *nifH* region of bean rhizobial isolates due to few restriction sites conserved among species and/or biovars (Laguerre *et al.* 2001). The *nifH*-RFLP of common bean isolates from farmers' fields and

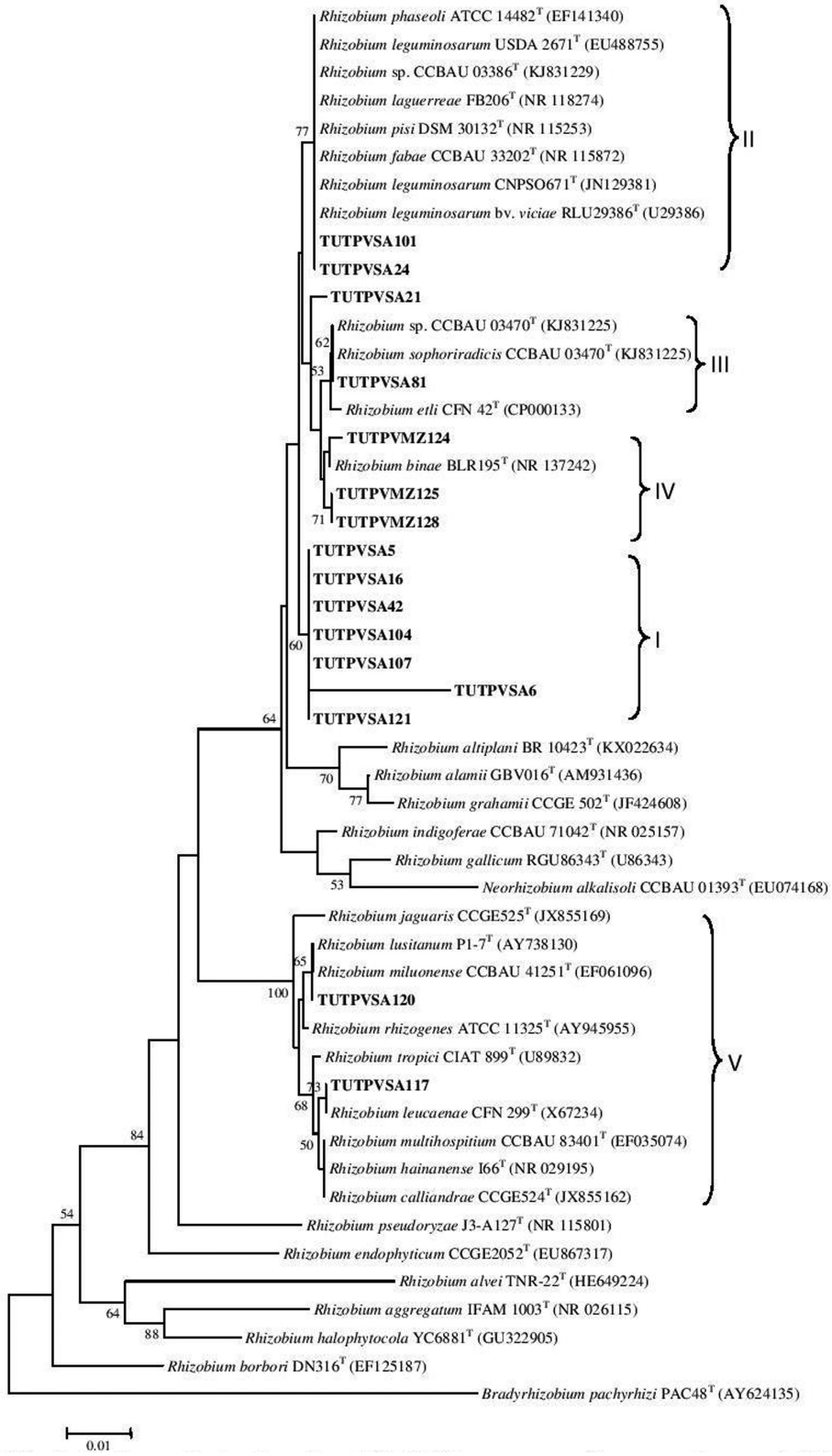

Figure 5. Phylogenetic tree based on 16S rRNA sequences of common bean-nodulating rhizobial isolates. Bootstraps values are indicated on the internal branches of the tree. The scale bar represents 1% nucleotide substitution.

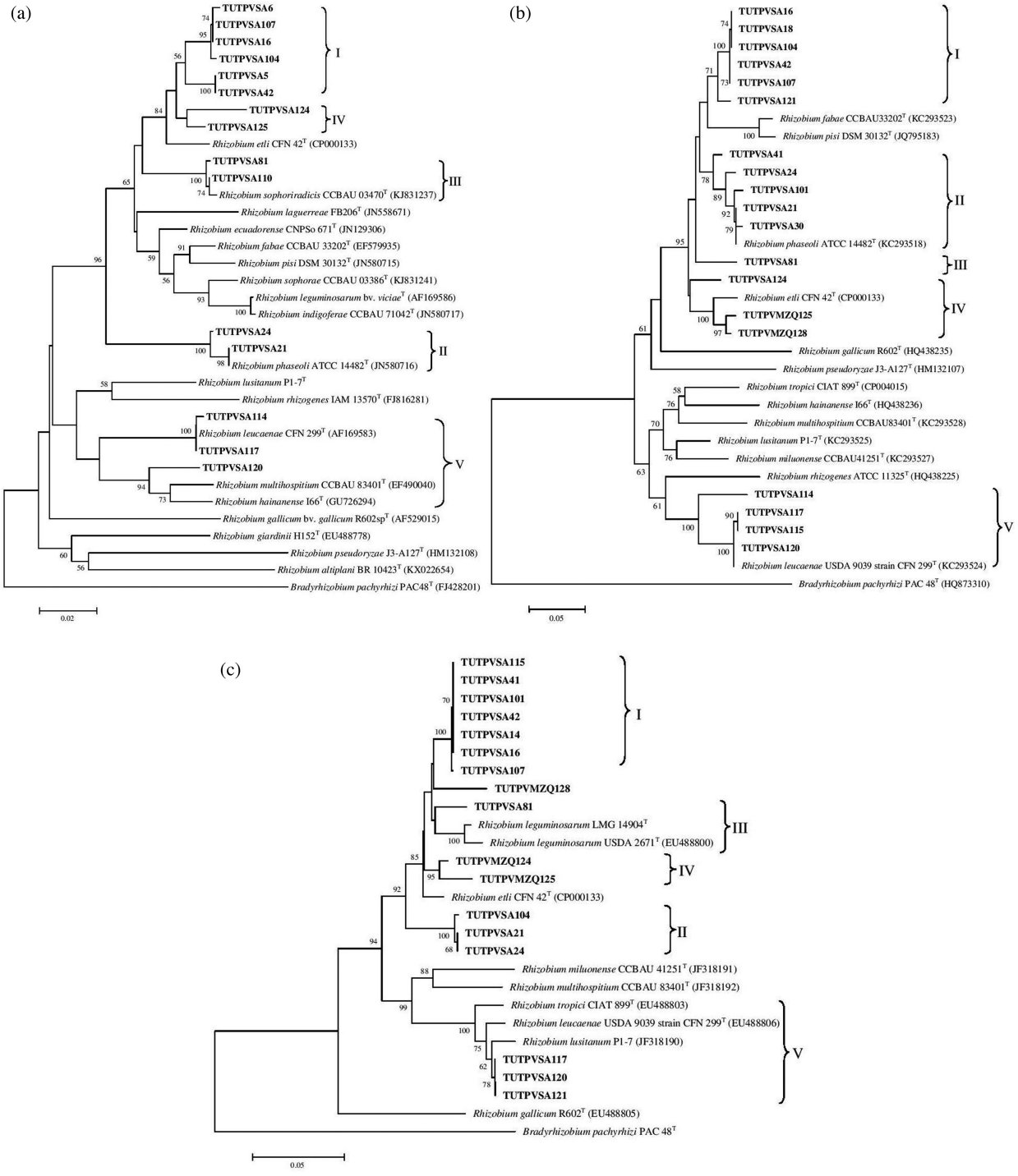

Figure 6. **(a)** Phylogenetic tree based on *glnII* gene sequences of common bean-nodulating rhizobial isolates. Bootstraps values are indicated on the internal branches of the tree. The scale bar represents 2% nucleotide substitution. **(b)** Phylogenetic tree based on *gyrB* sequences of common bean-nodulating rhizobial isolates. Bootstraps values are indicated on the internal branches of the tree. The scale bar represents 5% nucleotide substitution. **(c)** Phylogenetic tree based on *gltA* sequences of common bean-nodulating rhizobial isolates. Bootstraps values are indicated on the internal branches of the tree. The scale bar represents 5% nucleotide substitution.

experimental plots in Limpopo was highly conserved due to the presence of very low polymorphism (Fig. 3). The *nifH*-RFLP analysis showed the presence of polymorphic bands with *Msp*I (Table 3) which supports the results of Laguerre *et al.* (2001). Most of the isolates produced similar monomorphic restriction-digested bands with the two restriction enzymes. This finding is consistent with the report of Amarger, Macheret and Laguerre (1997) and Martinez *et al.* (1985) which showed that common bean-nodulating rhizobia such as *Rhizobium giardinii*, *R. etli* and *R. phaseoli* revealed similar RFLP profiles and contained three copies of *nifH* gene in their symbiotic plasmids. The presence of low polymorphism in all isolates could suggest that the

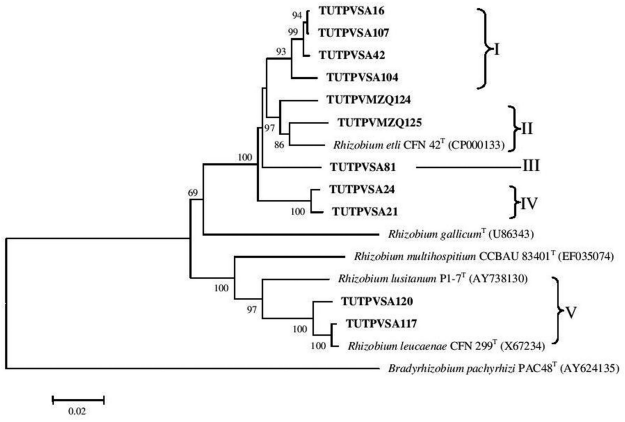

Figure 7. Phylogenetic tree based on 16S rRNA+*glnII*+*gyrB*+*gltA* concatenated sequences showing relationship with reference bacterial strains obtained from NCBI GeneBank. Bootstraps values are indicated on the internal branches of the tree. The scale bar represents 2% nucleotide substitutions.

genes are closely related. It was therefore very difficult to construct restriction maps for the *nifH* region of the test isolates. Nucleotide sequences are required for a more rigorous phylogenetic analysis.

The 16S rRNA sequence data showed that the bean-nodulating microsymbionts formed many groups with *R. leguminosarum*, *R. etli* and *R. tropici* (Fig. 5). The results from sequencing of housekeeping genes further confirmed that common bean is indeed a promiscuous legume nodulated by different rhizobial types (Kaschuk *et al.* 2006). The rhizobial species naturally associated with *Phaseolus vulgaris* for N_2 fixation have been identified as *R. leguminosarum* sv. *phaseoli*, *R. phaseoli*, *R. tropici*, *R. etli*, *R. leucaenae*, *R. giardinii* sv. *phaseoli*, *R. gallicum*, *R. lusitanum*, *R. pisi*, *R. freirei*, *R. mesoamericanum*, *R. paranaense*, *R. giardinii* sv. *giardinii*, *R. miluonense* and *R. ecuadorense* (Ribeiro *et al.* 2015). Other unnamed species-level *Rhizobium* lineages have been reported elsewhere (Aserse *et al.* 2012; Lopez-Guerrero *et al.* 2012).

In this study, the groupings obtained with individual gene phylogenies were congruent with those of the concatenated tree. Furthermore, the constructed concatenated phylognetic tree was supported by high bootstrap values which resolved the aberrant clustering found in individual housekeeping gene phylogenies. The isolates in cluster I of each housekeeping gene tree formed a separate sister clade with *R. etli* with 97% sequence similarity (Fig. 7). Although isolate TUTPVSA81 stood alone in the concatenated tree, it was closely related to *R. sophoriradicis* in the 16S rRNA and *glnII* gene phylograms (Jiao *et al.* 2015). *Rhizobium shophoriradicis* isolated from *Sophora flavescens* nodules has been identified as a widely distributed species responsible for nodulation of *P. vulgaris* (Rouhrazi, Khodakaramian and Velázquez 2016). Anyango *et al.* (1995) and Diouf *et al.* (2000) also detected the presence of *R. etli* and *R. leguminosarum* in African soils, and suggested that *R. etli* might have been introduced to Africa together with common bean seed from the Americas in the early 16th century. The alignment of some rhizobial isolates such as TUTPVMZQ124, TUTPVMZQ125 and TUTPVMZQ128 with *R. etli* is consistent with the results of Pinto, Hungria and Martins Mercante (2007) and Aserse *et al.* (2012), which showed that *R. etli* is the true microsymbiont of common bean. Even *R. etli* has been reported to be the main common bean-nodulating rhizobia in the centre of host origin and diversity in Latin America (Young and Martinez-Romero 1993). In concert with our results, Rouhrazi, Khodakaramian and Velázquez (2016) recently reported the presence of diverse common bean-nodulating rhizobial species (*R. sophoriradicis*, *R. leguminosarum*, *R. tropici*) and novel lineages in Iranian soil.

The two symbiotic genes (*nifH* and *nodC*) studied showed incongruency with the phylogenies of core genes phylogenies. Although the isolates were grouped in clusters I–IV in the concatenated tree, they formed only one cluster in *nifH* and *nodC* phylogenies. It could be due to the location of *nifH* and *nodC* genes in transmissible plasmids in some *Rhizobium* species (Martinez-Romero and Palacios 1990). Aserse *et al.* (2012) also found that Ethiopian bean-nodulating rhizobia exhibited *nifH* and *nodC* phylogenies, which were incongruent with the phylogenies of core genes. The discordance of *nifH* and *nodC* phylogenies with 16S rRNA was also reported by Laguerre *et al.* (2001). The detected inconsistency when comparing *nifH* and *nodC* with 16S rRNA and housekeeping gene sequences might be due to inter strain gene transfer and recombination of symbiotic nucleotide sequences. This would suggest differences in evolutionary history between chromosomal and symbiotic genes. The grouping of test isolates with *R. etli*, *R. vallis* and *R. phaseoli* in one big cluster supports the presence of interspecific symbiotic plasmid exchange and common evolutionary history of *nifH* and *nodC* genes among these species (Martinez *et al.* 1985; Amarger, Macheret and Laguerre 1997; Laguerre *et al.* 2001; Aserse *et al.* 2012). According to Eardly *et al.* (1995), *R. leguminosarum* bv. *phaseoli*, *R. leguminosarum* bv. *viciae* and *R. leguminosarum* bv. *trifolii* were traditionally classified as biovars of a single species; however, through the use of *nifH* and *nodC* gene sequencing two subdivisions have been distinguished based on differences in symbiotic plasmid and host nodulation range. Isolates TUTPVSA110 and TUTPVSA117, which occupied a different cluster in *nifH*-RFLP analysis, probably had different restriction sites for the restriction enzymes used. These isolates showed identical (100%) nucleotide sequences with strains of the *R. tropici* group (e.g. *R. miluonense*, *R. multihospitium*, *R. lusitanum*) and formed a cluster, which suggests that the *nifH* gene of this cluster probably has the same origin.

Based on the sequences of two housekeeping genes (*glnII* and *gyrB*) and 16S rRNA, rhizobial isolate TUTPVSA101 was found to be a species of *R. phaseoli*, and TUTPVSA104 a novel *Rhizobium* species, in contrast to the *gltA* gene sequence phylogeny, where isolates TUTPVSA101 and TUTPVSA104 were not linked to any type strains of *Rhizobium* (Fig. 6c). Even isolate TUTPVSA115 showed inconsistent result in *gyrB* and *gltA* phylogenies. Isolates with phylogenetically inconsistent behaviour were discrepant in their placement within the different housekeeping gene trees, and therefore grouped in different clusters. The discrepant clustering of these isolates could be attributed to differences in the evolutionary history of the genes, horizontal gene transfer and/or subsequent recombination events (Degefu, Wolde-meskel and Frostegård 2012; Naamala, Jaiswal and Dakora 2016).

Some African soils (e.g. in Senegal and Gambia in West Africa, or Kenya in East Africa) harbour bean-nodulating rhizobia related to *R. etli* and *R. tropici* (Anyango *et al.* 1995; Diouf *et al.* 2000). Aserse *et al.* (2012) have identified *R. etli*, *R. phaseoli* and a novel group of *Rhizobium* spp. to be the common bean-nodulating rhizobia in Ethiopian soils. In this study, all the isolates from Mozambique and the Limpopo Province of South Africa were closely associated with *R. etli*, but isolates from the strongly acidic soils of the Eastern Cape Province delineated with *R. leucaenae* and *R. miluonense* within the *R. tropici* group in the phylogenetic analysis. There is thus a clear and striking existence

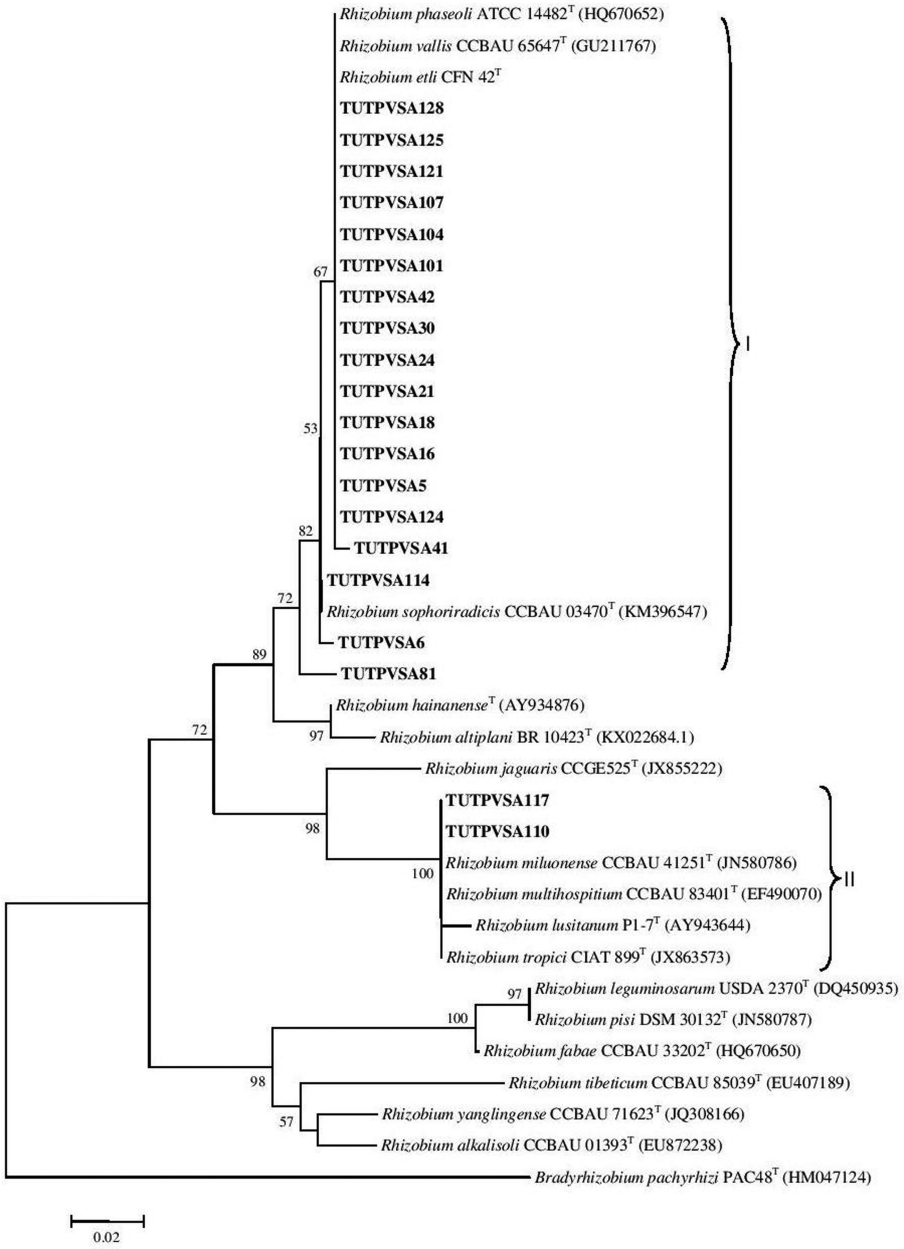

Figure 8. Phylogenetic tree based on *nifH* sequences of common bean-nodulating rhizobial isolates. Bootstraps values are indicated on the internal branches of the tree. The scale bar represents 2% nucleotide substitution.

of geographical distribution of different bean-nodulating rhizobial species in soils of Southern Africa. Although the common bean rhizobial isolates exhibited different RFLP patterns and occupied different RFLP clusters in this study (Figs 2 and 3), they revealed identical housekeeping and symbiotic gene sequences (see Figs 5–9). This result underscores the importance of RFLP as a tool for delineating strain-level diversity among common bean-nodulating rhizobial populations.

In this study, isolate identification to species level showed that bean rhizobia in Mozambican soils belong to *R. etli* appeared to be the predominant species-nodulating common bean. In contrast, South African soils harboured to *R. etli, R. phaseoli, R. sophoriradicis* and *R. leucaenae,* together with a unique cluster containing isolates TUTPVSA16, TUTPVSA107, TUTPVSA42 and TUTPVSA104 that related to *R. etli* but seemingly forming a putative new species. *Rhizobium phaseoli* was first isolated in Europe, and it was suggested that the native *R. leguminosarum* obtained a symbiotic plasmid from introduced South American bean seed-containing *R. etli* sv. *phaseoli* (Segovia, Young and Martínez-Romero 1993; Pérez-Ramírez *et al.* 1998). Therefore, the presence of *R. etli* in Southern African soils could also be due to the introduction of this bacterium with bean seed from South America. *Rhizobium leucaenae* present in Brazilian soils was probably derived from *R. tropici* type A strain due to its genetic instability (Ribeiro *et al.* 2012). Like Cerrados savannah region of Brazil, the Eastern Cape Province of South Africa is probably a hub of *R. leucaenae* within the *R. tropici* group since it was isolated from very stressful acidic conditions (soil pH 4.1–5.3), similar those reported by Anyango *et al.* (1995) and Hungria *et al.* (2000).

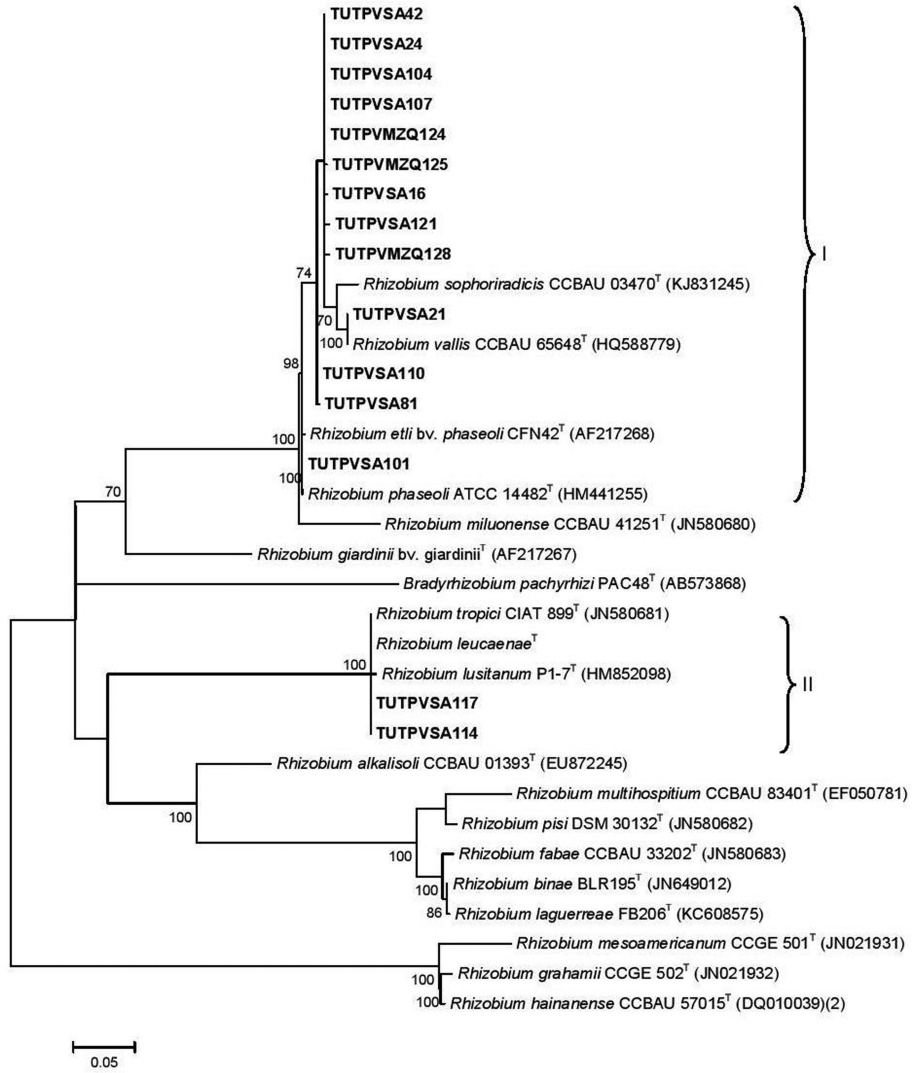

Figure 9. Phylogenetic tree based on *nodC* sequences of common bean-nodulating rhizobial isolates. Bootstraps values are indicated on the internal branches of the tree. The scale bar represents 5% nucleotide substitution.

The constructed phylogenies from gene sequences have confirmed that the *P. vulgaris* strains from Southern Africa are closely related to symbionts from common bean in its native neotropics. The *R. etli*, *R. tropici* and *R. leucaenae* are species that effectively nodulate common bean even though they are the natural symbionts of *Leucaena leucocephala*, *Mimosa caesalpiniifolia*, *Gliricidia sepiumhas* and *Mimosa affinis* in central Mexico (Bontemps *et al.* 2016). In fact, this part of the neotropics is home to many related rhizobial species such as *R. calliandrae*, *R. grahamii* and *R. mesoamericanum*. Common bean therefore shows promiscuity with a wide range of rhizobia for its nodulation. A diverse group of rhizobia present in soils of the centre of origin of common bean is responsible for its nodulation. In other words, there is no 'genuine' symbiont of *P. vulgaris*. *Rhizobium etli* is reported to be carried on the testa of bean seeds (Pérez-Ramírez *et al.* 1998); therefore, a selection of these microsymbionts might have been brought to Africa with bean seeds during colonial times, and their descendants (both vertical and horizontal) are now the principal symbionts of common bean plants in Africa.

In conclusion, this study exhibited a geographical distribution of diverse bean-nodulating rhizobial populations in soils of Southern African. Phylogenetic analysis showed that some of rhizobial isolates are unique and novel; thus, further studies are required to unravel their identity. The diverse rhizobial isolates should be further evaluated for N_2-fixing efficacy and symbiotic performance in order to select superstrains for inoculant production and increased common bean yield in Africa.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSEC* online.

FUNDING

This work was supported with grants from the Bill and Melinda Gates Foundation Project on Capacity Building in Legume Sciences in Africa, the South African Department of Science and Technology, the Tshwane University of Technology, the National

Research Foundation in Pretoria, and the South African Research Chair in Agrochemurgy and Plant Symbioses. MKZ is grateful for a competitive fellowship from the Bill and Melinda Gates Foundation Project on Capacity Building in Legume Sciences in Africa.

Conflict of interest. None declared.

REFERENCES

Amarger N, Macheret V, Laguerre G. *Rhizobium gallicum* sp. nov. and *Rhizobium giardinii* sp. nov., from *Phaseolus vulgaris* nodules. *Int J Syst Evol Micr* 1997;**47**:996–1006.

Anyango B, Wilson KJ, Beynon JL *et al*. Diversity of rhizobia nodulating *Phaseolus vulgaris* L. in two Kenyan soils with contrasting pHs. *Appl Environ Microb* 1995;**61**:4016–21.

Aserse AA, Räsänen LA, Assefa F *et al*. Phylogeny and genetic diversity of native rhizobia nodulating common bean (*Phaseolus vulgaris* L.) in Ethiopia. *Syst Appl Microbiol* 2012;**35**:120–31.

Beebe SE, Rao IM, Blair MW *et al*. Phenotyping common beans for adaptation to drought. In: Monneveux P, Ribaut JM, Okono A (eds). *Drought Phenotyping in Crops: From Theory to Practice*, Vol. 4, SA: Frontier Media. 2013, 1–20.

Beyene D, Kassa S, Ampy F *et al*. Ethiopian soils harbor natural populations of rhizobia that form symbioses with common bean (*Phaseolus vulgaris* L.). *Arch Microbiol* 2004;**181**:129–36.

Broughton W, Dilworth M. Control of leg *Haemoglobin* synthesis in snake beans. *Biochem J* 1971;**125**:1075–80

Broughton WJ, Hernandez G, Blair M *et al*. Beans (Phaseolus spp.): model food legumes. *Plant Soil* 2003;**252**:55–128.

Bontemps C, Rogel MA, Wiechmann A *et al*. Endemic Mimosa species from Mexico prefer alphaproteobacterial rhizobial symbionts. *New Phytol* 2016;**209**:319–33.

Degefu T, Wolde-meskel E, Frostegård A. Phylogenetic multilocus sequence analysis identifies seven novel Ensifer genospecies isolated from a less-well-explored biogeographical region in East Africa. *Int J Syst Evol Micr* 2012;**62**: 2286–95.

Diouf A, de Lajudie P, Neyra M *et al*. Polyphasic characterization of rhizobia that nodulate *Phaseolus vulgaris* in West Africa (Senegal and Gambia). *Int J Syst Evol Micr* 2000;**50**:159–70.

Eardly BD, Wang FS, Whittam TS *et al*. Species limits in *Rhizobium* populations that nodulate the common bean (*Phaseolus vulgaris*). *Appl Environ Microb* 1995;**61**:507–12.

Edgar RC. MUSCLE; Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acid Res* 2004;**31**:1792–7.

Farooq FT, Vessey JK. Genetic diversity of *Bradyrhizobium japonicum* within soybean growing regions of the north-eastern Great Plains of North America as determined by REP-PCR and ERIC-PCR profiling. *Symbiosis* 2009;**48**:131–42.

Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 1985;**39**:783–91.

Grange L, Hungria M. Genetic diversity of indigenous common bean (*Phaseolus vulgaris*) rhizobia in two Brazilian ecosystems. *Soil Biol Biochem* 2004;**36**:1389–98.

Hall T. *BioEdit version 7.0. 0. Distributed by the author*. 2004, www.mbio. ncsu. edu/BioEdit/bioedit. html (24 November 2016, date last accessed).

Hungria M, de S Andrade D, de O Chueire LM *et al*. Isolation and characterization of new efficient and competitive bean (*Phaseolus vulgaris* L.) rhizobia from Brazil. *Soil Biol Biochem* 2000;**32**:1515–28.

Jiao YS, Yan H, Ji ZJ *et al*. *Rhizobium sophorae* sp. nov. and *Rhizobium sophoriradicis* sp. nov., nitrogen-fixing rhizobial symbionts of the medicinal legume *Sophora flavescens*. *Int J Syst Evol Micr* 2015;**65**:497–503.

Kaschuk G, Hungria M, Andrade DS *et al*. Genetic diversity of rhizobia associated with common bean (*Phaseolus vulgaris* L.) grown under no-tillage and conventional systems in Southern Brazil. *Appl Soil Ecol* 2006;**32**:210–20.

Kimura M. A simple method for estimating evolutionary rate of base substi- tutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980;**16**:111–20.

Laguerre G, Mavingui P, Allard MR *et al*. Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: application to *Rhizobium leguminosarum* and its different biovars. *Appl Environ Microb* 1996;**62**:2029–36.

Laguerre G, Nour SM, Macheret V *et al*. Classification of rhizobia based on *nod*C and *nif*H gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology* 2001;**147**:981–93.

Li L, Sinkko H, Montonen L *et al*. Biogeography of symbiotic and other endophytic bacteria isolated from medicinal Glycyrrhiza species in China. *FEMS Microbiol Ecol* 2012;**79**:46–68.

Lopez-Guerrero MG, Ormeno-Orrillo E, Velazquez E *et al*. *Rhizobium etli* taxon- omy revised with novel genomic data and analyses. *Syst Appl Microbiol* 2012;**35**:353–8.

Marek-Kozaczuk M, Leszcz A, Wielbo J *et al*. *Rhizobium pisi* sv. *trifolii* K3. 22 harboring nod genes of the *Rhizobium leguminosarum* sv. *trifolii* cluster. *Syst Appl Microbiol* 2013;**36**: 252–8

Martens M, Dawyndt P, Coopman R *et al*. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus Ensifer (including former Sinorhizobium). *Int J Syst Evol Micr* 2008;**58**:200–14.

Martens M, Delaere M, Coopman R *et al*. Multilocus sequence analysis of Ensifer and related taxa. *Int J Syst Evol* **Micr** 2007;**57**:489–503.

Martinez ES, Pardo MA, Palacios R *et al*. Reiteration of nitrogen fixation gene sequences and specificity of *Rhizobium* in nodulation and nitrogen fixation in *Phaseolus vulgaris*. *Microbiology* 1985;**131**:1779–86.

Martinez-Romero E. Diversity of *Rhizobium*–Phaseolus vulgaris symbiosis: overview and perspectives. *Plant Soil* 2003;**252**: 11–23.

Martinez-Romero E, Palacios R. The *Rhizobium* genome. *Crit Rev Plant Sci* 1990;**9**:17–29.

Mnasri B, Mrabet M, Laguerre G *et al*. Salt-tolerant rhizobia isolated from a Tunisian oasis that are highly effective for symbiotic N2-fixation with *Phaseolus vulgaris* constitute a novel biovar (bv. *mediterranense*) of *Sinorhizobium meliloti*. *Arch Microbiol* 2007;**187**:79–85.

Mothapo NV, Grossman JM, Maul JE *et al*. Genetic diversity of resident soil rhizobia isolated from nodules of distinct hairy vetch (*Vicia villosa* Roth) genotypes. *Appl Soil Ecol* 2013;**64**: 201–13.

Naamala J, Jaiswal SK, Dakora FD. Microsymbiont diversity and phylogeny of native bradyrhizobia associated with soybean (*Glycine max* L. Merr.) nodulation in South African soils. *Syst Appl Microbiol* 2016;**39**:336–44

Nkot NL, Krasova-Wade T, Etoa FX *et al*. Genetic diver- sity of rhizobia nodulating *Arachis hypogaea* L. in diverse land use systems of humid forest zone in Cameroon. *Appl Soil Ecol* 2008;**40**:411–6.

Nzoué A, Miché L, Klonowska A *et al*. Multilocus sequence analysis of bradyrhizobia isolated from Aeschynomene species in Senegal. *Syst Appl Microbiol* 2009;**32**:400–12.

Odee DW, Haukka K, McInroy SG *et al.* Genetic and symbiotic characterization of rhizobia isolated from tree and herbaceous legumes grown in soils from ecologically diverse sites in Kenya. *Soil Biol Biochem* 2002;**34**: 801–11.

Pérez-Ramírez NO, Rogel MA, Wang E *et al.* Seeds of *Phaseolus vulgaris* bean carry *Rhizobium etli*. *FEMS Microbiol Ecol* 1998;**26**:289–96.

Pinto FGS, Hungria M, Martins Mercante F. Polyphasic characterization of Brazilian *Rhizobium tropici* strains effective in fixing N2 with common bean (*Phaseolus vulgaris* L.). *Soil Biol Biochem* 2007;**39**:1851–64.

Pongsilp N. *Phenotypic and Genotypic Diversity of Rhizobia*. Thailand: Bentham Science Publishers 2012.

Peix A, Ramírez-Bahena MH, Velázquez E *et al.* Bacterial associations with legumes. *Crit Rev Plant Sci* 2015;**34**:17–42.

Ribeiro RA, Martins TB, Ormeño-Orrillo E *et al.* *Rhizobium ecuadorense* sp. nov., an indigenous N_2-fixing symbiont of the Ecuadorian common bean (*Phaseolus vulgaris* L.) genetic pool. *Int J Syst Evol Micr* 2015;**65**:3162–9.

Ribeiro RA, Rogel MA, López-López A *et al.* Reclassification of *Rhizobium tropici* type A strains as *Rhizobium leucaenae* sp. nov. *Int J Syst Evol Micr* 2012;**62**:1179–84.

Rohlf FJ. NTSYSpc: numerical taxonomy system. ver. 2.21c. Exeter Software: Setauket: New York, 2009.

Rouhrazi K, Khodakaramian G, Velázquez E. Phylogenetic diversity of rhizobial species and symbiovars nodulating Phaseolus vulgaris in Iran. *FEMS Microbiol Lett* 2016;**363**, DOI: 10.1093/femsle/fnw024.

Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987;**4**:406–25.

Sarita S, Sharma PK, Priefer UB *et al.* Direct amplification of rhizobial nodC sequences from soil total DNA and comparison to nodC diversity of root nodule isolates. *FEMS Microbiol Ecol* 2005;**54**:1–11.

Segovia L, Young JP, Martínez-Romero E. Reclassification of American *Rhizobium leguminosarum* biovar *phaseoli* type I strains as *Rhizobium etli* sp. nov. *Int J Syst Evol Micr* 1993; **43**:374–7.

Shamseldin A, Moawad H, El-Rahim WMA *et al.* Near-full length sequencing of 16S rDNA and RFLP indicates that *Rhizobium etli* is the dominant species nodulating Egyptian winter Berseem clover (*Trifolium alexandrinum* L.). *Syst Appl Microbiol* 2014;**37**:121–8.

Shamseldin AAY, Vinuesa P, Thierfelder H*et al.* *Rhizobium etli* and *Rhizobium gallicum* nodulate *Phaseolus vulgaris* in Egyptian soils and display cultivar-dependent symbiotic efficiency. *Symbiosis* 2005;**38**:145–61.

Singh SP. Integrated genetic improvement, In: *Com-mon Bean Improvement in the Twenty First Century. Developments in Plant Breeding*, Vol. **7**. Dordrecht: Kluwer Academic Publishers, 1999, 133–65.

Somasegaran P, Hoben HJ. *Handbook for Rhizobia*. New York: Springer 1994.

Stackebrandt E, Frederiksen W, Garrity GM *et al.* Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Micr* 2002;**52**:1043–7.

Stępkowski T, Żak M, Moulin L *et al.* *Bradyrhizobium canariense* and *Bradyrhizobium japonicum* are the two dominant *Rhizobium* species in root nodules of lupin and serradella plants growing in Europe. *Syst Appl Microbiol* 2011;**34**:368–75.

Tamura K, Stecher G, Peterson D *et al.* MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 2013;**30**:2725–9.

USDA. *Agriculture Marketing Services, Livestock and Grain Market News*. Greeley, CO, 2012.

Valverde A, Igual JM, Peix A *et al.* *Rhizobium lusitanum* sp. nov. a bacterium that nodulates *Phaseolus vulgarus*. *Int J Syst Evol Micr* 2006;**56**:2631–7.

Weisburg WG, Barns SM, Pelletier DA *et al.* 16S ribossomal DNA amplification for phylogenetic study. *J Bacteriol* 1991;**173**:697–703.

Young JPW, Martinez-Romero EM. Reclassification of American *Rhizobium leguminosarum* biovar phaseoli type I strains as *Rhizobium etli* sp . nov. *Int J Syst Evol Micr* 1993;**43**:374–7.