Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Hidden diversity: genetic characterization of non-symbiotic rhizobium endophytes across agro- forestry legume species in Southern Ethiopia

Abraham Mikru
Hawasa University
Fassil Assefa
asefafasil2013@gmail.com

Addis Ababa University
Endalkachew Woldemeskel
World Agroforestry (ICRAF)
Kristina Lindström
University of Helsinki
Leena Räsänen
University of Helsinki
Belay Berza
Debre Markos University
Aneta Dresler-Nurmi
University of Helsinki

Research Article

Keywords: Agro-forestry legumes, Diversity, Endophytic bacteria, root nodules

Posted Date: June 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9663504/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Legumes often harbor a complex community of non-rhizobial endophytes whose roles and taxonomic identities remain poorly understood in diverse ecological zones. We investigated the diversity of endophytic bacteria associated with the root nodules of four woody legumes (*Acacia abyssinica*, *Albizia gummifera*, *Indigofera arrecta*, and *Millettia ferruginea*) and seven herbaceous species (*Arachis hypogaea*, *Cajanus cajan*, *Cicer arietenum*, *Glycine max*, *Phaseolus vulgaris, Vigna radiata*, and *Vigna unguiculata*) in southern Ethiopia. A total of 135 bacterial strains were isolated from surface-sterilized root nodules collected from 49 field plants and 62 trap hosts. Classification via partial 16S rDNA sequencing identified *Enterobacter* (31%), *Bacillus* (18%), and *Rhizobium* (15%) as the most prevalent genera. To further resolve the taxonomy of the *Rhizobium* isolates, we conducted a multilocus sequence analysis (MLSA) by using three housekeeping genes: 16S rDNA, *recA*, and *rpoB* genes. Maximum likelihood and concatenated phylogenies consistently grouped the test isolates together, showing high similarity to the type strains of *Rhizobium tropici* and *R. leucaenae*. Despite this genetic proximity, none of the isolates exhibited rhizobial function during nodulation tests or possessed key symbiotic genes (*nifH*, *nodA*, *nodC*). Furthermore, AFLP fingerprinting suggested these non-symbiotic *Rhizobium* isolates likely represent at least three distinct novel genospecies. We discuss the potential causes for these findings, including the inherent absence of symbiotic loci or the loss of symbiotic genes during repeated sub-culturing.

1. Introduction

Agro-forestry is a sustainable land-use strategy that integrates woody perennials with crops or livestock to foster agro-ecological succession and system resilience. In Ethiopia, smallholder farmers employ diverse agro-forestry practices tailored to specific socio-economic and bio-physical contexts [1]. Within these systems, retained woody species provide critical ecosystem services, including soil stabilization, energy, fodder, and timber, thereby bolstering household food security.

Legumes play a foundational role in these systems by enhancing biodiversity and driving soil fertility through atmospheric nitrogen fixation. Ethiopia serves as a major center of legume diversity, hosting over 600 species, 48 of which are endemic. While many indigenous legumes are integrated into traditional practices [2], the most recognized examples include scattered Acacia trees in croplands and the complex home-garden systems of Southern Ethiopia. These southern systems are typically defined by coffee and enset crops interspersed with multipurpose legumes such as *Acacia abyssinica, Albizia gummifera, Millettia ferruginea, and Indigofera arrecta* [1, 2]. Beyond these woody species, herbaceous legumes like *Arachis hypogaea* and *Vigna unguiculata* are widely grown, yet their full symbiotic potential remains under-exploited in the Ethiopian context.

The ecological success of these legumes is largely attributed to their symbiotic association with rhizobia, which can fix between 200 to 300 kg of nitrogen per hectare [3]. Historically, Ethiopian research has focused almost exclusively on these rhizobia [4–7] and mycorrhizal fungi [8–10]. However, the traditional view that non-rhizobial organisms within nodules were merely opportunistic invaders or contaminants has been overturned.

Contemporary research indicates that non-rhizobial endophytes (NREs) form stable, functional microbiomes alongside rhizobia, significantly influencing plant growth, nodulation, and protection [11–14]. These NREs comprising genera such as *Pseudomonas, Bacillus*, and *Enterobacter* provide advantageous traits including environmental stress resistance and disease suppression, even though they do not induce nodule formation by themselves [15]. Harnessing this microbial potential for bio-fertilizer development requires a comprehensive assessment of local strain diversity. The selection of effective inoculants depends heavily on a robust library of strains adapted to indigenous hosts. Consequently, this study investigated the diversity of endophytic bacteria within the root nodules of four woody agro-forestry legumes (*Acacia abyssinica, Albizia gummifera, Indigofera arrecta, and Millettia ferruginea*) across ten locations in Southern Ethiopia. We report on the taxonomy, genomic diversity, and phylogeny of these isolates using AFLP fingerprinting and multilocus sequence analysis of the 16S rRNA, *rpoB, and recA* genes.

2. Material and Methods

2.1 Sampling

Root nodules and soil samples were collected from ten locations in Southern Ethiopian regions (Fig. 1). These locations represent different agro-ecological zones in the region (Table 1).The natural vegetation of the locations comprised of bushes (Leku, Durame,

Shone and Wolyita-Sodo), open acacia woodlands (Arsi Negele) and upland rainforests (Wondo Genet, Yirga Alem, Dila, Wonago and Bule).

2.2 Collection of root nodules

We collected root nodule samples from 49 field-standing woody legumes distributed in the ten study locations (Table 1). The 49 woody legumes consisted of 2 *Acacia abyssinica*, 21 *Millettia ferruginea*, 20 *Albizia gummifera*, and 6 *Indigofera arrecta.* All the woody legumes were growing as natural stands, retained for their perceived values by the locals whereas, *Millettia ferruginea* was predominantly found interspersed in coffee-enset home gardens. *Indigofera arrecta* was found in sloppy scrubland in Durame and wooded grassland in Shone. *Acacia abyssinica* was found mostly on open grassland or boundaries along agricultural fields. *Albizia gummifera* was often found as part of homestead woodlot and field boundaries. On visit to each location, available target legumes were assessed for nodulation by excavation and follow up of lateral roots as described before [16]. In all cases, 10 to 20 healthy root nodules per plant were collected and maintained in silica gel vials [17]. Ethical clearance was obtained for the collection of plant materials. Plant seeds were obtained from wild as well as from farm owners with permission. The plants were identified by Dr. Nigussie Amsalu, a Botanist and voucher specimen of each plant was deposited at Debre Markos University Herbarium with collection numbers 001–011.

2.3 Trapping from field soil under tree canopy

Around the roots of the woody legumes from which multiple nodules were recovered, soil samples were collected and seeds of seven different grain legumes (*Aarachis hypogaea, Cajanus cajan, Cicer arietenum, Glycine max* L, *Phaseolus vulgaris, Vigna radiata* and *Vigna unguiculata*) were sawn on separate seedbeds. The purpose was to check if similar endophytic bacteria would occupy the root nodules of the herbaceous and woody legumes. Root nodules were finally harvested from nodulated seedlings after a month.

Table 1
Study locations geographic and Agro-ecological data for the target woody legumes (**numbers** 1- *Acacia, 2 -Albizia, 3 -Indigofera, 4 - Millettia*)

Location	Latitude/ Longitude	Altitude (m above sea level)	Mean ann. Rainfall (in mm)	Mean ann. temperature in°C	Agroecol. zone	Dominant vegetation under/near the woody legume	Woody legumes encountered
A/Ngele	7° 1' N/ 38° 35' E	2043	600–650	20–25	Semi-arid, warm	Wheat field	2
Bule	6°20' N/ 38°19' E	2800	900–1,200	9–21	Sub-humid highland	Enset-Coffee home garden, Faba bean field	4
Dilla	6° 45' 00" N/ 38°25' E	1570	1300–1350	16–30	Sub-humid, warm	Enset-Coffee home garden	4
Durame	7 ° 21’ N/ 38 ° 42' E	2101	1000–1400	13–28	Subhumid, warm	Coffee-Enset homegardens, Eucalyptus woodlots	1, 2, 3, 4
Leku	6° 54' N/ 37° 45' E	1830	750–1,200	18–24	Moist, cold	Homstead maize field, Coffee	2, 4
Shone	7° 14' N/ 37° 53' E	1500	1200–1300	15–25	Sub-humid, warm	Eucalyptus woodlots	2
W/Sodo	7° 10' N/ 37° 55' E	1600	1200–1300	13–17	Sub-humid warm	Eucalyptus and Albizia wood lots	2
Wonago	6°24'30" N/ 38°18' 30" E	1750	950–1300	16–30	Sub-humid, hot	Enset-Coffee home garden	4
W/genet	7° 1' N/ 38° 35' E	1723	950–1,350	10–27	Sub-humid, warm	Albizia woodlot	2
Yirga Alem	6° 50' N/ 38° 30' E	1776	950–1300	10–27	Moist, warm	Enset-Coffee home garden	2, 4

2.4 Trapping from soil samples

The purpose of this method was to address the legumes for which the direct excavation method was not successful. In addition, because, some target legumes have restricted distribution. For example, *Acacia* and *Indigofera* were encountered only in Durame (Table 1). Therefore, soil samples were brought to the National Soil Testing Center (NSTC), Addis Ababa, Ethiopia, for trapping experiment. The seeds of the woody legumes were treated by immersion in concentrated sulfuric acid for 15 minutes (45 minutes for *Acacia*) and rinsed in five changes of sterile water. For the grain legumes, the seeds were surface sterilized by immersion in 70% ethanol (1 minute) followed by 3% NaOCl for 3 minutes in each sequentially. The treated seeds were aseptically put into sterile Petri dish cushioned inside with sterile moistened filter paper and incubated at ambient conditions. The soil sample from each site was prepared in duplicates for each of the legumes. After three to four days, the germinated seedlings were aseptically transferred to plastic pots containing well-mixed soil. Two to three seedlings per pot were allowed to grow under glasshouse conditions for two months. Sterile water was used during the whole growth period.

2.5 Isolation and purification of endophytic bacteria

Endophytic bacteria were isolated from root nodules following standard procedure for the isolation of rhizobia [17]. Five to ten preserved nodules were taken from each vial and single nodules were separately surface-sterilized in 70% ethanol and 3% Sodium hypochlorite sequentially for 3 minutes and rinsed in 5 changes of sterile water. Surface sterile individual nodules were crushed separately in 2 mL sterile water. A loopful of the nodule suspension was streak plated on Yeast-Manitol Agar (YEMA) (DIFCO) supplemented with Congo red (0.2%; w/v) (YEMA-CR) and incubated at 28 ^{0}C for 5 to 10 days in dark [18]. Morphologically distinct and well-isolated colonies were picked and purified by repeated sub-culturing on YEMA-CR agar separately. For mixed cultures,

repeated sub-culturing was done on the same medium after dilution with Tween 80 buffer. The purified isolates were maintained and preserved at -80 ^{0}C in YEM broth with 20% glycerol.

2.6 Assessment of nodulation capacity

All isolates were tested for nodulation on original host plants following standard method [17]. Plastic growth pouches filled with sterile vermiculite were used for all legumes except *Millettia ferruginea*. Owing to their large size, the *Millettia* seeds were grown in plastic pots filled with sterile vermiculite. All seedlings were inoculated with 1 ml of log phase bacterial culture (10^8 CFU per ml). Non-inoculated seedlings were included as negative control. For all cases, sterile water and a quarter strength Jensen's medium were used on alternate days and the seedlings were allowed to grow for a month. For strains that were related to members of rhizobial genera (based on 16S rDNA sequence comparison), the nodulation test was repeated by growing seedlings in glass jars, filled with sterile mixtures of sand, Leca-gravel and vermiculite. The seedlings were allowed to grow for a month in a growth cabinet under controlled conditions.

2.7 DNA Isolation

Genomic DNA was extracted by following a modification of the hexadecyltrimethyl ammonium bromide (CTAB) method of Wilson [19].

2.8 AFLP fingerprinting

The isolates were analysed by using the amplified fragment length polymorphism (AFLP) fingerprinting method described before [20]. Double-stranded total DNA was restricted and half site-specific *Eco*R1 and *Mse*I adapters were ligated to the restriction fragments in the same reaction. The restricted and adapter tagged-DNA fragments were PCR amplified by using primers with two selective bases (bold) at the 3' ends: *Eco*R1-**AC** (5' FAM-GAC TGC GTA CCA ATT C**AC**-3') and *Mse*I-**GC** (5'-GAT GAG TCC TGA GTA A**GC**-3').

2.9 PCR amplification and sequencing of housekeeping and symbiotic genes

The primers used and the PCR programs followed for the different target genes are listed in Table 2. Conserved primers were used for all the respective genes as described in Asrese *et al*. [21]. For *rrn* amplification, a PTC-200 Peltier Thermal Cycler (MJ Research, Watertown, MA, U.S.A) was used. The *rrn* was amplified with 1µl of the DNA sample per 50 µl of master mix. The master mix consisted of 5 µl of 10 x Dynazyme buffer, 0.27 µl of 92.6 pmol/µl fD1 primer, 0.07 µl of 384.6 pmol/µl rD1 primer, 0.4 µl of 10 mM dNTP, 0.75 µl of 2 U/µl Dynazyme DNA polymerase II (Finnzymes) and pure water to make a reaction volume of 50 µl.

The *rpo*B and *rec*A genes were amplified by using Phusion polymerase in a gradient Master cycler (Eppendorf) following recommended protocol (Finnzymes). In both cases, 1 µl of the DNA sample per 50 µl of master mix was used. The PCR master mix consisted of 10 µl of 5x Phusion HF buffer, 0.25 µl of 100 pmol/µl each of the forward and reverse primers, 0.25 µl of 2 U/µl Phusion DNA polymerase (Finnzymes) and 1 µl of 10,000 pmol/µl dNTP and ddH_2O to make up a reaction volume of 50 µl. The presence of symbiotic genes (*nod*A, *nod*C and *nif* H) was probed using primers nodA-1, nodA-2 [22]; nodC-F, nodC-I [23]; nifH-f, and nifH-I [23] by Phusion DNA polymerase (Finnzymes). The phusion polymerase and the PCR master mix preparations were similar to the one described for *rpo*B and *rec*A.

The PCR products were verified by electrophoresis of 5 µl of product along with molecular mass markers, GeneRuler (Fermentas) in 1% agarose gel. Sequencing was done directly from the PCR products at the Institute of Biotechnology, University of Helsinki, Finland. For *rpo*B and *rec*A genes, sequencing was done from both DNA strands using their respective primers and the contiguous sequence was deduced using BioEdit software. For *rrn*, the hyper-variable region was sequenced first with rD1 primer (Table 2). Latter, to resolve the taxonomy of the strains that were related to rhizobial genera a larger stretch (ca 1350 bp) the *rrn* gene was re sequenced by using both fD1 and rD1 primers at the Norwegian University of Life Sciences, Norway.

Table 2
List of primers and PCR programs used in the study

Gene	Primers	5'-3' nucleotide sequence	Reference	PCR programs
16S rRNA *(srr)*	fD1	5'-AGA GTT TGA TCC TGG CTC AG-3'	Weisburg *et al*. 1991	– 98°C for 30 s initial denaturation – 35 rounds: – 98°C for 15 s denaturation – 60°C for 30 s annealing – 72°C for 30 s extension – 72°C for 5 min final extension
	rD1	5'-AAG GAG GTG ATC CAG CC-3'		
*rpo*B	rpoB-83F	5'-CCT SAT CGA GGT TCA CAG AAG GC-3'	Martens *et al*. 2008	– 98°C for 3 min initial denaturation – 3 rounds: – 98°C for 2 min denaturation – 60°C for 2 min annealing – 72°C for 1 min extension – 30 rounds: – 98°C for 30 s denaturation – 60°C for 1 min annealing – 72°C for 1 min extension – 72°C for 5 min final extension
	rpoB-1061R	5'-AGC GTG TTG CGG ATA TAG GCG-3'		
*rec*A	recA-6F	5'-CGK CTS GTA GAG GAY AAA TCG GTG GA-3'	Martens *et al*. 2008	– 98°C for 3 min initial denaturation – 32 rounds -98.0°C for 30 s denaturation -60.0°C for 1 min annealing – 74°C for 1 min extension -72°C for 5:00 min final extension
	recA-555R	5'-CGR ATC TGG TTG ATG AAG ATC ACC AT-3'		
*nod*A	nodA-1	5'-TGC RGT GGA ARN TRN NCT GGG AAA-3'	Haukka *et al*. 1998	– 98°C for 30 s initial denaturation – 35 rounds: – 98°C for 15 s denaturation – 60°C for 30 s annealing – 72°C for 30 s extension – 72°C for 5 min final extension
	nodA-2	5'-GGN CCG TCR TCR AAW GTC ARG TA-3' Symbols used: R = A or G, W = A or T, X or N= unknown		
*nod*C	nodC-F	5'-AYG THG TYG AYG ACG GTT C-3'	Laguerre *et al*. 2001	– 98°C for 30 s initial denaturation – 35 rounds: – 98°C for 15 s denaturation – 55°C for 30 s annealing – 72°C for 30 s extension – 72°C for 5 min final extension
	nodC-I	5'-CGY GAC AGC CAN TCK CTA TTG-3' Symbols used: Y = C or T; H = A, C or T; N = A, C, G or T; K = G or T		
*nif*H	nifH-f	5'-TAC GGN AAR GGS GGN ATC GGC AA-3'	Laguerre *et al*. 2001	– 98°C for 30 s initial denaturation – 35 rounds:

Gene	Primers	5'-3' nucleotide sequence	Reference	PCR programs
	nifH-I	5'-AGC ATG TCY TCS AGY TCN TCC A-3' Symbols used: N = A, C, G or T; R = A or G; S = C or G; Y = C or T		– 98°C for 10 s denaturation – 58°C for 10 s annealing – 72°C for 20 s extension – 72°C for 5 min final extension

2.10 Data analysis & phylogenetic reconstruction

AFLP Fingerprinting: AFLP data were processed via ABI GeneScan and converted to densitometry curves using Bionumerics v5. Fragments between 35–500 bp were compared using Dice coefficients (1% position tolerance; 0.5% optimization). Clusters were visualized using a UPGMA dendrogram. **Taxonomic Identification**: Initial genus-level identification was performed via NCBI BLAST (16S rDNA) and confirmed by constructing a Neighbor-Joining (NJ) tree using the SILVA database and ARB software. The 16S rRNA gene sequences were deposited to the NCBI Genbank with the following accession numbers: EU266588,HM137034,DQ536511,AB102735,AY972446,AB377521,EU399926,DQ849078,HQ164548,HM222691,FJ755919, AY461699,FJ561295,FJ784127,AB066339,HM584906, DQ207559,EU855197,GU252111,GU339273,FM164631,AB118223, AB330406,GU368375,EF114313,HQ171486,AB308444, FM164633,AY310732,AB675636,GQ478258,HM585374,EU304248,FJ190535,EU331414,AB435591,GQ451698,CP000653, FJ463411,EU855208,EU221358,EF175731,DQ122370,GQ259887,EU078621,FJ707292,HQ242723,AB495172,GU368371, GU186121,FU550141,AJ717361,EU704697,AB186915, AM421978,GU204957,CP002206,DQ122372,GU726581,GU358073,FJ527671,HM045838,AY569299,EU934229,GU475123, GU367870,FJ595998,EU074181,EF149005,HM582866,HM852118,EU637927,AF511524,CP000826 and AB334774.

Multi-Locus Sequence Analysis (MLSA): For Rhizobium-related isolates, taxonomic resolution was refined using 16S rDNA, recA, and rpoB genes: **Alignment**: Sequences were aligned via ClustalW (for NJ trees in MEGA11). **Processing**: Alignments were manually refined in BioEdit 7, trimmed to equal lengths, and protein-coding genes were verified via amino acid translation. **Concatenation**: A 1,608 bp concatenated dataset (recA + 16S rDNA + rpoB) was generated. DAMBE was used to filter the dataset to 27 unique sequences. **Phylogenetic Modeling**: Best-fit substitution models were determined online (LANL). Maximum Likelihood (ML) trees were then generated in PhyML 3.0 using: GTR + Gamma for recA and concatenated sequences. Tamura-Nei + Gamma for for rpoB.

3. Results

3.1 Isolation, identification and distribution of bacterial isolates

A total of 135 endophytic bacteria were isolated from the entire root nodule samples. None of the isolates were able to nodulate their legume host on authentication test. Partial 16S rDNA sequencing and comparison with NCBI GenBank allowed identification of the isolates into 20 different genera belonging to five bacterial classes. The majority of the isolates belonged to gamma-proteobacteria (67) followed by alpha-proteobacteria (34) and gram-positive firmicutes (25) (Table 3; Table 4).

Table 3
List of endophytic bacterial strains isolated from surface-sterilized root nodules of four woody and seven herbaceous legumes in southern Ethiopia.

Strain	Legume host	Sampling (associated legume)	Geographic Origin	Dominant vegetation	Bacterial Genus (class)	16S rDNA identity %	Accession No.
Ac111	*A.abyssinica*	Tg(If)	Durame	Eucalyptus	*Achromobacter* sp (β)	99	EU266588
AC113	*A.abyssinica*	Tg (Mf)	Leku	Cornfield	*Achromobacter* sp (β)	99	EU266588
Ag3b	*A.gummifera*	DE	Durame	Eucalyptus	*Achromobacter* sp (β)	99	HM137034
Mf36b	*M.ferruginea*	DE	Y/Alem	Coffee	*Acinetobacter* sp (γ)	98	DQ536511
Ac1g	*A.abyssinica*	DE	Durame	Enstet, Teff	*Agrobacterium* sp (α)	99	AB102735
Ag96	*A.gummifera*	DE	W/Genet	Albiziat	*Agrobacterium* sp (α)	99	AB102735
CW88b	*V.unguiculata*	Tf (Ag)	Shone	Eucalyptus	*Agrobacterium* sp (α)	99	AY972446
HB76	*P.vulgaris*	Tf (Ag)	W/Sodo	Eucalyptus	*Agrobacterium* sp (α)	99	AY972446
HB87a	*P.vulgaris*	Tf (Ag)	W/Genet	Albizia	*Agrobacterium* sp (α)	99	AB377521
Mf30c	*M.ferruginea*	DE	Dilla	Enset-coffee	*Agrobacterium* sp (α)	99	AY972446
PP143	*C.cajan*	Tg (Mf)	Leku	Cornfield	*Agrobacterium* sp (α)	100	EU399926
Ag14a	*A.gummifera*	DE	W/Sodo	Eucalyptus	*Arthrobacter* sp (δ)	97	EU399926
Ag126	*A.gummifera*	Tg (Mf)	Dilla	Enset-Coffee	*Arthrobacter* sp (δ)	96	DQ849078
Ag22b	*A.gummifera*	DE	Dilla	Coffee	*Bacillus* sp (Gm+)	99	HQ164548
GN79a	*A.hypogeae*	Tf (Mf)	Durame	Enset-Coffee	*Bacillus* sp (Gm+)	99	HM222691
PP46a	*C.cajan*	Tf (Mf)	Dilla	Enset-Coffee	*B.thuringiensis* (Gm+)	99	FJ755919
PP54a	*C.cajan*	Tf (Mf)	W/Genet	Albizia	*Bacillus sp* (Gm+)	98	AY461699
PP57a	*C.cajan*	Tf (Mf)	Y/Alem	Coffee	*Bacillus sp* (Gm+)	99	FJ561295
PP56a	*C.cajan*	Tf (Mf)	Leku	Corn field	*Bacillus sp* (Gm+)	99	FJ784127
CP64a	*C.arietenum*	Tf (Mf)	Durame	Enset-Coffee	*B.megaterium* (Gm+)	99	AB066339
CP78a	*C.arietenum*	Tf (Ag)	W/Sodo	Eucalyptus	*B.megaterium* (Gm+)	99	HM584906
SB83a	*G.max*	Tf (Ag)	A/Negele	Wheat field	*B.megaterium* (Gm+)	99	HM584906
SB131	*G.max*	Tg (Mf)	Leku	Cornfield	*B.pumilus* (Gm+)	99	DQ207559
IF124	*I.arrecta*	Tg (Mf)	Durame	Enset-Coffee	*B.pumilus* (Gm+)	99	HQ164548
IF115	*I.arrecta*	Tg (Mf)	Dilla	Enset-Coffee	*B.pumilus* (Gm+)	100	EU855197

Strain	Legume host	Sampling (associated legume)	Geographic Origin	Dominant vegetation	Bacterial Genus (class)	16S rDNA identity %	Accession No.
IF116	*I.arrecta*	Tg (Mf)	Bule	Faba bean field	*B.pumilus* (Gm+)	99	EU855197
HB52a	*P.vulgaris*	Tf (Mf)	Leku	Cofee	*B.thuringiensis* (Gm+)	99	FJ755919
HB58a	*P.vulgaris*	Tf (Mf)	Dilla	Enset-Coffee	*B.megaterium* (Gm+)	99	AB066339
HB82a	*P.vulgaris*	Tf (Ag)	W/Sodo	Eucalyptus	*Bacillus sp* (Gm+)	98	GU252111
HB84	*P.vulgaris*	Tf (Ag)	A/Negele	Wheat field	*Bacillus sp* (Gm+)	98	GU339273
MB50a	*V.radiata*	Tf (Mf)	Dilla	Cofee-Enset	*B.megaterium* (Gm+)	99	FM164631
MB70a	*V.radiata*	Tf (Ag)	W/Sodo	Eucalyptus	*Bacillus sp* (Gm+)	98	AB118223
CW51a	*V.unguiculata*	Tf (Mf)	Y/Alem	Enset-Coffee	*Bacillus sp* (Gm+)	98	AB330406
CW59a	*V.unguiculata*	Tf (Mf)	Dilla	Enset-Coffee	*Bacillus sp* (Gm+)	99	GU368375
CW60a	*V.unguiculata*	Tf (Mf)	Dilla	Enset-Coffee	*B.megaterium* (Gm+)	99	AB066339
CW66a	*V.unguiculata*	Tf(Mf)	Durame	Enset-Coffee	*B.megaterium* (Gm+)	99	GU368375
CW71a	*V.unguiculata*	Tf(Ag)	A/Negele	Wheat field	*B.aryabhattai* (Gm+)	99	EF114313
PP127	*C.cajan*	Tg(Mf)	Dilla	Enset-Coffee	*B.elkanii* (α)	99	HQ171486
PP141	*C.cajan*	Tg (Mf)	Leku	Cornfield	*B.elkanii* (α)	99	HQ171486
If106	*I.arrecta*	DE	Durame	Eucalyptus	*B.elkanii* (α)	100	HQ171486
Ac1a	*A.abyssinica*	DE	Durame	Enstet, Teff	*Enterobacter* sp (γ)	98	AB308444
AC114	*A.abyssinica*	T (Mf)	Bule	Faba bean field	*Enterobacter* sp (γ)	95	FM164633
AC120	*A.abyssinica*	T (Ag)	W/Sodo	Eucalyptus	*Enterobacter* sp (γ)	96	FM164633
Ag3c	*A.gummifera*	DE	Durame	Eucalyptus	*Enterobacter* sp (γ)	98	AY310732
Ag4c	*A.gummifera*	DE	Shone	Eucalyptus	*E.asburiae* (γ)	98	AY310732
Ag5a	*A.gummifera*	DE	W/Sodo	Grassland	*Enterobacter sp* (γ)	97	FM164633
Ag5b	*A.gummifera*	DE	W/Sodo	Grassland	*Enterobacter sp* (γ)	97	AB675636
Ag6a	*A.gummifera*	DE	Shone	Eucalyptus	*Enterobacter sp* (γ)	96	GQ478258
Ag9a	*A.gummifera*	DE	W/Sodo	Eucalyptus	*Enterobacter sp* (γ)	94	FM164633
Ag10b	*A.gummifera*	DE	W/Sodo	Grassland	*E.cloacae* (γ)	99	HM585374
Ag13a	*A.gummifera*	DE	Shone	Eucalyptus	*Enterobacter* sp (γ)	94	FM164633
Ag27a	*A.gummifera*	DE	W/Genet	Albizia	*Enterobacter* sp (γ)	98	EU304248
Ag33b	*A.gummifera*	DE	Leku	Eucalyptus	*Enterobacter* sp (γ)	97	FJ190535
Ag45a	*A.gummifera*	DE	W/Genet	Albizia	*Enterobacter* sp (γ)	94	EU331414
Ag45b	*A.gummifera*	DE	W/Genet	Albizia	*Enterobacter* sp (γ)	95	FM164633
GN62a	*A.hypogeae*	Tf (Ag)	W/Sodo	Eucalyptus	*Enterobacter* sp (γ)	97	EU331414

Strain	Legume host	Sampling (associated legume)	Geographic Origin	Dominant vegetation	Bacterial Genus (class)	16S rDNA identity %	Accession No.
GN69a	*A.hypogeae*	T f(Ag)	Shone	Eucalyptus	*Enterobacter* sp (γ)	97	EU331414
If103	*I.arrecta*	DE	Durame	Eucalyptus	*E.cloacae* (γ)	99	HM585374
If105	*I.arrecta*	DE	Durame	Eucalyptus	*Enterobacter* sp(γ)	98	HM585374
If107	*I.arrecta*	DE	Durame	Eucalyptus	*Enterobacter* sp(γ)	95	AY310732
IF118	*I.arrecta*	Tg (Mf)	Leku	Cornfield	*Enterobacter* sp(γ)	96	FM164633
IF119	*I.arrecta*	Tg (Mf)	Dilla	Enset-Coffee	*Enterobacter* sp(γ)	98	AB675636
IF129	*I.arrecta*	Tg (Mf)	Durame	Enset-Coffee	*Enterobacter* sp(γ)	97	EU331414
Mf2a	*M.ferruginea*	DE	Durame	Enset, coffee	*Enterobacter* sp(γ)	96	AB308444
Mf2b	*M.ferruginea*	DE	Durame	Enset, coffee	*Enterobacter* sp(γ)	97	AB435591
Mf29a	*M.ferruginea*	DE	Dilla	Enset, coffee	*Enterobacter* sp(γ)	94	GQ451698
Mf29b	*M.ferruginea*	DE	Dilla	Enset, coffee	*Enterobacter* sp(γ)	95	AB308444
Mf30b	*M.ferruginea*	DE	Dilla	Enset, coffee	*Enterobacter* sp(γ)	97	AY310732
Mf32a	*M.ferruginea*	DE	Dilla	Enset, coffee	*Enterobacter* sp(γ)	98	CP000653
Mf36a	*M.ferruginea*	DE	Y/Alem	Coffee	*Enterobacter* sp(γ)	96	FJ463411
Mf37a	*M.ferruginea*	DE	Dilla	Coffee	*Enterobacter* sp(γ)	95	FM164633
Mf38a	*M.ferruginea*	DE	Leku	Coffee	*Enterobacter* sp(γ)	98	AY310732
Mf38b	*M.ferruginea*	DE	Leku	Coffee	*Enterobacter* sp(γ)	99	AY310732
Mf39	*M.ferruginea*	DE	Leku	Corn field	*Enterobacter* sp(γ)	95	EU304248
Mf44	*M.ferruginea*	DE	Y/Alem	Coffee	*Enterobacter* sp(γ)	97	EU304248
MF121	*M.ferruginea*	Tg (Ac)	Durame	Eucalyptus	*Enterobacter* sp(γ)	98	FM164633
HB63a	*P.vulgaris*	Tf (Mf)	Durame	Enset-Coffee	*Enterobacter* sp(γ)	98	AB435591
HB63b	*P.vulgaris*	Tf (Mf)	Durame	Enset-Coffee	*Enterobacter* sp(γ)	98	AB435591
MB85a	*V.radiata*	Tf (Ag)	Shone	Eucalyptus	*Enterobacter* sp(γ)	96	EU855208
CW53a	*V.unguiculata*	Tf (Mf)	Dilla	Enset-Coffee	*Enterobacter* sp(γ)	97	AB308444
CW92a	*V.unguiculata*	Tf(Mf)	Durame	Enset-Coffee	*Enterobacter* sp(γ)	97	EU221358
CW95a	*V.unguiculata*	Tf(Ag)	Shone	Eucalyptus	*Enterobacter* sp(γ)	95	EF175731
Ac1b	*A.abyssinica*	DE	Durame	Enstet, Teff	*K.pneumoniae* (γ)	99	DQ122370
Mf24a	*M.ferruginea*	DE	Dilla	Enset, coffee	*Klebsiella* sp (γ)	98	GQ259887
HB48a	*P.vulgaris*	Tf(Ag)	W/Genet	woodlot/Albizia	*Klebsiella* sp (γ)	97	DQ122370
CW61a	*V.unguiculata*	Tf(Ag)	W/Sodo	Eucalyptus	*Klebsiella* sp (γ)	98	EU078621
AC112	*A.abyssinica*	Tg(Mf)	Durame	Enset-Coffee	*L.wisconsinensis* (α)	99	FJ707292
Mf34	*M.ferruginea*	DE	Dilla	Coffee	*L.adecarboxylata* (γ)	99	HQ242723
Mf21b	*M.ferruginea*	DE	Wonago	Enset, coffee	*L.aquatica* (δ)	99	AB495172

Strain	Legume host	Sampling (associated legume)	Geographic Origin	Dominant vegetation	Bacterial Genus (class)	16S rDNA identity %	Accession No.
Cp125	*C.arietenum*	Tg (Ag)	Durame	Eucalyptus	*Methylobacterium* (α)	88	GU368371
MB137	*V.radiata*	Tg(Mf)	Bule	Faba bean field	*Methylobacterium* (α)	92	GU186121
AC110	*A.abyssinica*	Tg(Ag)	Shone	Eucalyptus	*Microbacterium* sp (δ)	99	FN550141
IF137	*I.arrecta*	Tg(Ag)	A/Negele	Wheatfield	*Microbacterium* sp (δ)	99	AJ717361
Cp133	*C.arietenum*	Tg(Ag)	Shone	Eucalyptus	*Micrococcusluteus* (δ)	99	EU704697
MB136	*V.radiata*	Tg(Mf)	Wonago	Enset-Coffee	*Paenibacillus* sp (Gm+)	97	AB186915
Ac104	*A.abyssinica*	DE	Durame	Eucalyptus	*Pantoea* sp (γ)	99	AM421978
If12d	*I.arrecta*	DE	Durame	Eucalyptus	*Pantoea* sp (γ)	98	GU186121
Mf21a	*M.ferruginea*	DE	Wonago	Enset, coffee	*Pantoea* sp (γ)	98	GU186121
Mf30a	*M.ferruginea*	DE	Dilla	Enset, coffee	*Pantoea* sp (γ)	96	GU204957
Mf31a	*M.ferruginea*	DE	Dilla	Enset, coffee	*Pantoea* sp (γ)	97	CP002206
Mf40b	*M.ferruginea*	DE	Wonago	Enset, coffee	*Pantoea* sp (γ)	94	DQ122372
Mf89	*M.ferruginea*	DE	Dilla	Coffee	*Pantoea* sp (γ)	98	GU186121
MB102a	*V.radiata*	Tf (Mf)	Durame	Enset-Coffee	*Pantoea* sp (γ)	98	GU186121
Ac117	*A.abyssinica*	T g(Mf)	Dilla	Enset-Coffee	*Pseudomonas* sp (γ)	98	GU726581
Ag3a	*A.gummifera*	DE	Durame	Eucalyptus	*P.fluorescens* (γ)	99	GU358073
Ag27b	*A.gummifera*	DE	W/Genet	Albizia	*P.stutzeri* (γ)	99	FJ527671
Ag33a	*A.gummifera*	DE	Leku	Eucalyptus	*Pseudomonas* sp (γ)	98	HM045838
IF134	*I.arrecta*	Tg (Mf)	Y/Alem	Enset-Coffee	*Pseudomonas* sp (γ)	99	AY569299
Mf15b	*M.ferruginea*	DE	Durame	Enset, coffee	*P.fluorescens* (γ)	100	GU358073
Mf40a	*M.ferruginea*	DE	Wonago	Enset, coffee	*Pseudomonas* sp (γ)	97	EU934229
Mf43c	*M.ferruginea*	DE	Bule	Enset	*P.fluorescens* (γ)	100	GU475123
HB49a	*P.vulgaris*	Tf(Mf)	Y/Alem	Enset-Coffee	*P.fluorescens* (γ)	99	GU367870
AC139	*A.abyssinica*	Tg (Ag)	A/Negele	Wheatfield	*R.leguminosarum* (α)	99	FJ595998
Ag4a	*A.gummifera*	DE	Shone	Eucalyptus	*R.tropici* (α)	100	EU074181
Ag6b	*A.gummifera*	DE	Shone	Eucalyptus	*R.tropici* (α)	99	EU074181
Ag7a	*A.gummifera*	DE	W/Sodo	Eucalyptus	*Rhizobium* sp (α)	99	EF149005
Ag8a	*A.gummifera*	DE	A/Negele	Wheat field	*R.tropici (α)*	99	EU074181
Ag10a	*A.gummifera*	DE	W/Sodo	Grassland	*R.tropici (α)*	100	EU074181

Strain	Legume host	Sampling (associated legume)	Geographic Origin	Dominant vegetation	Bacterial Genus (class)	16S rDNA identity %	Accession No.
Ag11a	*A.gummifera*	DE	W/Sodo	Grassland	*R.rhizogenes* (α)	99	HM582866
Ag14b	*A.gummifera*	DE	W/Sodo	Eucalyptus	*R.tropici (α)*	99	EU074181
Ag17b	*A.gummifera*	DE	Durame	Eucalyptus	*R.tropici (α)*	99	EU074181
Ag19b	*A.gummifera*	DE	Shone	Eucalyptus	*R.tropici (α)*	100	EU074181
Ag20b	*A.gummifera*	DE	W/Sodo	Eucalyptus	*R.tropici (α)*	99	EU074181
Ag22a	*A.gummifera*	DE	Dilla	Coffee	*R.tropici (α)*	99	EU074181
Ag28a	*A.gummifera*	DE	Dilla	Coffee	*R.tropici (α)*	99	EU074181
If16a	*I.arrecta*	DE	Durame	Eucalyptus	*Rhizobium sp* (α)	98	HM852118
HB100a	*P.vulgaris*	Tf(Ag)	Shone	Eucalyptus	*R.leguminosarum* (α)	99	EU637927
Mf23b	*M.ferruginea*	DE	Durame	Eucalyptus	*R.tropici* (α)	100	EU074181
Mf25a	*M.ferruginea*	DE	Y/Alem	Enset, coffee	*R.tropici* (α)	99	EU074181
Mf26a	*M.ferruginea*	DE	Bule	Bean field	*R.tropici* (α)	99	EU074181
Mf30g	*M.ferruginea*	DE	Dilla	Enset, coffee	*Rhizobium sp* (α)	99	EF149005
CW55a	*V.unguiculata*	Tf (Ag)	W/Genet	woodlot/Albizia	*R.tropici* (α)	100	EU074181
Mf41	*M.ferruginea*	DE	Dilla	Enset, coffee	*Serattia* sp(γ)	96	AF511524
Mf43a	*M.ferruginea*	DE	Bule	Enset	*Serattia* sp(γ)	98	CP000826
IF132	*I.arrecta*	Tg (Ag)	W/Sodo	Eucalyptus	*S.melonis* (α)	99	AB334774

The principal geographic origins of the isolates were Durame, 31 (21.4%); Dilla, 27 (20%); and Wolayita Sodo, 18 (13.3%). Shone and Leku contributed 13 and 12 strains respectively, while less than 10 isolates came from each of the remaining five locations (Table 5). Wolayta Sodo (5), Shone (4), Dilla (3) and Durame (3) were the major origins of the isolates related to non symbiotic rhizobial gnera (Table 3).

The most frequently encountered genera in their order were *Enterobacter* 42 (31%), *Bacillus* 24 (18%), and *Rhizobium* 20 (15%). The rest of 49 isolates scattered among 17 bacterial genera (Table 4). Fourteen bacterial genera were represented by less than five isolates each. Seven genera including, *Acinetobacter, Labrys, Lecilericia, Leifsonia, Micrococcus, Paenibacillus* and *Sphingomonas* were represented by only one isolate each (Table 4). Twenty-five isolates were related to members of rhizobial genera consisting of 20 related to *Rhizobium*, 3 related to *Bradyrhizobium* and 2 related to *Methylobacterium* sp (Table 4). The majority of the isolates related to *Rhizobium* species were obtained from *Albizia gummifera* (12) and *Millettia ferruginea* (4); the remaining isolates were obtained one from each of *Acacia abyssinica, Indigofera arrecta, Phaseolus vulgaris* and *Vigna unguiculata* (Table 4). The three isolates related to *Bradyrhizobium* species were derived from the root nodules of *Cajanus cajan* (2) and *I. arrecta* (1), whereas the two isolates related to *Methylobacterium* came singly from the root nodules of *Cicer arietinum* and *Vigna radiate* (Table 4).

The majority (91 of 135 or 67%) of the bacteria were isolated from four woody legumes comprising of, 32 isolates from each of *Albizia gummifera* and *Millettia ferruginea*, 15 from *Indigofera arrecta* and 12 from *Acacia abyssinica* (Table 4). *Albizia gummifera* and *Millettia ferruginea* (31 isolates each) dominated the field-sanding legume sources whereas *Indigofera arrecta* (9) and *Acacia abyssinica* (8) dominated in the trapping method. Herbaceous legumes contributed 44 (33%) of the strains; *Phaseolus vulgar is* and *Vigna unguiculata* (11 or 8% each), being the major sources (Table 4). *Enterobacter* species were the most frequent (34) isolate from woody legumes, whereas *Bacillus* spp. (20 of 44) dominated the isolates from herbaceous legumes (Table 4). Although, *Glycine* and *Arachis* grew robustly in most of the locations, few seedlings were nodulated. *Cicer* exhibited the poorest growth, with scanty nodules in most of the sites.

Table 4
The number of endophytic bacteria isolated from surface-sterilized root nodules and their distribution by bacterial genera and host legumes growing in southern Ethiopia.

Bacterial Taxon	Ac	Ag	If	Mf	Ah	Ca	Cc	Gm	Pv	Vr	Vu	No. of isolates/r genus (%)
Alphaproteobacteria												
Agrobacterium sp	1	1	0	1	0	0	1	0	2	0	1	7 (5%)
Bradyrhizobium sp	0	0	1	0	0	0	2	0	0	0	0	3 (2%)
Labrys sp.	1	0	0	0	0	0	0	0	0	0	0	1 (0.7%)
Methylobacterium sp.	0	0	0	0	0	1	0	0	0	1	0	2 (1.5%)
Rhizobium sp	1	12	1	4	0	0	0	0	1	0	1	20 (15%)
Sphingomonas sp	0	0	1	0	0	0	0	0	0	0	0	1 (0.7%)
								Total (%) = 34 (25%)				
Betaproteobacteria												
Achromobacter sp.	2	1	0	0	0	0	0	0	0	0	0	3 (2%)
Gammaproteobacteria												
Acinetobacter sp	0	0	0	1	0	0	0	0	0	0	0	1 (0.7%)
Enterobacter sp	3	12	6	13	2	0	0	0	2	1	3	42 (31%)
Klebsiella sp.	1	0	0	1	0	0	0	0	1	0	1	4 (3%)
Leclericia sp	0	0	0	1	0	0	0	0	0	0	0	1 (0.7)
Pantoea sp.	1	0	1	5	0	0	0	0	0	1	0	8 (6%)
Pseudomonas sp	1	3	1	3	0	0	0	0	1	0	0	9 (7%)
Serattia sp	0	0	0	2	0	0	0	0	0	0	0	2 (1.5%)
								Total (%) = 67 (49.6%)				
Firmicutes												
Bacillus sp	0	1	3	0	1	2	4	2	4	2	5	24 (18%)
Paenibacillus sp	0	0	0	0	0	0	0	0	0	1	0	1 (0.7%)
								Total (%) = 25 (16%)				
Actinobacteridae												
Arthrobacter sp	0	2	0	0	0	0	0	0	0	0	0	2 (1.5%)
Leifsonia sp	0	0	0	1	0	0	0	0	0	0	0	1 (0.7%)
Microbacterium sp	1	0	1	0	0	0	0	0	0	0	0	2 (1.5%)
Micrococcus sp	0	0	0	0	0	1	0	0	0	0	0	1 (0.7%)
								Total (%) = 6 (4.4%)				
No. /legume host species	12	32	15	32	3	4	7	2	11	6	11	total = 135

Abrevations: Ac=Acacia abyssinica, Ag=Albizia gummifera, Ia=indigofera arrecta, Mf=Millettia ferruginea, Ca=Cicer arietenum, Vu=Vigna unguiculata, Ah=Arachis hypogeae, Pv=Phaseolus vulgaris, Vr=Vigna radiata, Cc=Cajanus cajan, Gm=Glycine max;

3.2 Phylogeny of the endophytic bacteria isolates related to *Rhizobium* species

In a preliminary assessment, partial 16S rDNA sequence phylogeny (Neighbour joining tree) of 32 non-symbiotic isolates related to rhizobial genera and those of reference strains retrieved from NCBI GenBank outlined the generic level taxonomic position of the test strains (Fig. 2). Accordingly, the majority (20 isolates) of the test isolates belonged to the genus *Rhizobium*, 7 to *Agrobacterium*, 3 to *Bradyrhizobium* and 2 to *Methylobacterium*. The 32 test isolates were clustered into six phylogenetic groups designated by capital letters A to F (Fig. 2). The 20 test isolates related to the genus *Rhizobium* nested in three 16S rRNA gene clads. Clad A included 18 test isolates derived mostly from *Albizia gummifera* (12) and *Millettia ferruginea* (4) trees, along with *Rhizobium leucaenae* LMG9517^T (0.001541), *R. lusitanum* P1-7^T (0.001540),*R. rhizogenes* IAM 1357^T (0.002311), *R. tropici* CCBAU 01483 (0.00077). Numbers in parenthesis are the pairwise 16S rRNA gene sequence distance values between the type strains and the test isolates which nested in clad A. Thus *R. tropici* CCBAU 01483 (0.00077) was the most closely related reference strain to the test isolates. The test isolates related to *R. radiobacter* LMG14^T (*Agrobacterium* complex), *Methylobacterium*, and *Bradyrhizobium* formed Clads D, E and F respectively (Fig. 2).

In order to resolve the taxonomy of the test isolates that were closely related to each other in clad A (Fig. 2), multilocus sequence analysis (MLSA) of three housekeeping genes including a longer stretch of the16S rRNA gene sequence (ca. 1300 bp) and two housekeeping genes including *rec*A (387 bp) and *rpo*B (401 bp) genes was performed for 16 representative isolates. The isolates were selected based on their AFLP fingerprinting pattern and legume hosts.

The maximum likelihood phylogenetic trees of individual genes as well as the concatenated sequences of the three housekeeping genes (*rec*A = 387 bp + *srr*=820 bp + *rpo*B = 401 bp) with size of 1608 bp, clearly resolved all the reference type strains and the test isolates related to *Rhizobium* species formed distinct clad (Fig. 3). The non-symbiotic test isolates remained closely related to each other with the closest sister clads being the type stains of *Rhizobium leucaenae* CFN299^T, *R. tropici* CIAT899^T, *R. multihospitium* CCBAU83401^T, *R. miluonense* CCBAU41251^T, and *R. hainanense* I66^T. A single test isolate (Mf30c) nested within the clad containing *R. radiobacter* LMG140^T (Agrobacterium complex) group.

The phylogenetic trees of the individual genes were more or less concordant with that of concatenated sequence where the closest sister clad to the test non-symbiotic *Rhizobium* species were *R. leucaenae* CFN299 ^{T} in *srr* (Fig S1) and *rec*A tree (Fig S2) and *R. miluonense* CCBAU41251^T in the *rpo*B tree (Fig S3).

3.3 Diversity of endophytic bacteria related to *Rhizobium* spp.

AFLP dendrogram constructed using UPGMA method segregated 28 isolates related to non-symbiotic *Rhizobium* spp. into six AFLP clusters at 50% similarity coefficient (Fig. 4). A single isolate (Mf30c) related to the type strain *Rhizobium radiobacter* LMG140 (*A. tumefacience)* in the housekeeping gene phylogenies remained un-clustered. AFLP1 encompassed 15 non-symbiotic *Rhizobium* species most of which were related to the “R. tropici group” based on multilocus sequence analysis of housekeeping genes. Two sub-clusters can be recognized within AFLP 1 at 60% similarity coefficient, namely: AFLP sub-cluster 1a consisting of 10 non-symbiotic *Rhizobium* species, derived mostly from *Albizia gumifera* (Ag20b, Ag22a, Ag14b, Ag17b, Ag19b, Ag11a,Ag4a, Ag28a, Ag6b, and Cw55a) and sub-cluster 1b encompassing 5 non-symbiotic *Rhizobium* species derived mostly from *Millettia ferruginea* (Mf23b, Mf26a, Mf30g, Mf25a, and If16a) (Fig. 4).

AFLP cluster 2 and 3 respectively contained 2 (Hb87a, Pp143a) and 3 (Hb76a, Ac1g, and Ag96a) isolates all of which were related to the *R. radiobacter* (*A. tumefacience* complex) group. AFLP cluster 4 contained two isolates related to *Bradyrhizobium elkani* (Pp127a, Pp141a), whereas AFLP cluster 5 contained two isolates one related to *R. leguminosarum* complex (Hb100a) and another one related to *Bradyrhizobium elkani* (If106a). AFLP cluster 6 encompassed 3 isolates of non-symbiotic *Rhizobium* species (Ag7a, Ag8a, and Ag10a) (Fig. 4).

4. Discussions

Research has extensively characterized the taxonomic identity of root-nodule-forming bacteria in tropical agricultural crops like mung bean, soybean, and cowpea [24–29]. While similar progress has been made for some tropical forest trees [30, 31], a significant knowledge gap remains regarding the microbiome of many ecologically vital tropical forest legumes. A defining feature of recent investigations, including the current study, is the high prevalence of non-rhizobial endophytes (NREs). While nitrogen-fixing bacteria (*Rhizobium*) typically dominate, metagenomic and culture-dependent analyses reveal they coexist with complex

communities of NREs. These "occupants" can comprise 7% to over 30% of the total nodule bacterial community, depending on host species and environmental conditions.

The presence of NREs is well-documented across both herbaceous and woody legumes. Studies on *Phaseolus vulgaris, Glycine max*, and *Arachis hypogaea* have frequently isolated NREs alongside traditional rhizobia. In *Acacia* species, NREs can account for over 25% of isolates [32]. In extreme cases, such as wild *Hedysarum* species in Algeria, NREs (specifically Gamma-proteobacteria) have been reported as the sole recoverable occupants, even when molecular methods like direct PCR detect the presence of rhizobial DNA [33]. In this study, 110 out of 135 isolates were identified as non-rhizobial genera. These isolates were unable to induce nodulation in their original hosts, echoing findings where NREs act as "opportunistic" residents that may enhance plant growth or yield without fixing nitrogen [34]. The failure to culture rhizobia from healthy nodules has led to the VBNC hypothesis. This state is a survival strategy where bacteria remain alive with low metabolic activity but cannot grow on standard media. Stress Induction: VBNC can be triggered by osmotic stress, oxygen limitation, or heavy metals (e.g., copper stress in R. leguminosarum). It is possible that the target legumes in this study harbor rhizobia in a VBNC state, which would explain why only NREs were recovered despite the presence of functional nodules.

A striking observation in this study was that even isolates related to known rhizobial genera lacked symbiotic features (the nodC and nifH genes). 1) This may be attributed to mobile genetic elements. Symbiotic genes are often located on symbiotic islands or mega-plasmids, which are inherently unstable [35]. 2) It could also be attributed to the spontaneous loss of symbiotic genes. Rhizobia can "jettison" these plasmids during repeated sub-culturing, frequent freeze-thaw cycles, or in a saprophytic soil state where a host is absent [36]. Extensive populations of non-symbiotic Rhizobium strains exist naturally in the rhizosphere, often outnumbering their symbiotic counterparts [37]. Identifying these diverse strains requires robust molecular tools. While excellent for differentiating closely related strains, AFLP can struggle with higher taxonomic levels due to high marker variability [38] and because 16S rRNA gene sequences can be too conserved to distinguish species, MLSA by using housekeeping genes (e.g., recA, atpD, rpoB) is now the gold standard. In this study, MLSA concordantly identified 17 isolates as relatives of *Rhizobium tropici.* The classification of *R. tropici* remains dynamic. Originally divided into types A and B, type A has since been reclassified as *Rhizobium leucaenae* based on polyphasic studies [39]. Despite the known environmental robustness of *R. tropici*, our isolates failed to nodulate either their original hosts, reinforcing the likelihood of symbiotic gene loss or inherent non-symbiotic status.

5. Conclusion

This study demonstrates that the root nodules of both woody and herbaceous legumes harbor a diverse plethora of endophytic bacteria beyond traditional nitrogen-fixing symbionts. While several isolates were taxonomically linked to rhizobial genera, none successfully induced nodulation in host plants. These findings underscore that the coexistence of non-rhizobial endophytes (NREs) and rhizobia is a widespread natural occurrence rather than an experimental anomaly. The absence of culturable, functional rhizobia in this study may be attributed to two key biological phenomena: the Viable but non-culturable state, induced by environmental or host-specific stressors, and the genetic instability of symbiotic plasmids. The spontaneous loss or rearrangement of mobile genetic elements (carrying nod and nif genes) appears to be a common trait among these isolates, particularly within the saprophytic soil phase or under lab frequent sub-culturing. Consequently, traditional culture-dependent methods may significantly underestimate actual rhizobial diversity.

Declarations

Funding source declaration

This was funded by collaboration of Academy of Finland, Helsinki University, Hawassa University, and Addis Ababa University.

Author Contribution

AM, data collection and analysis, methods, drat writing, revised manuscript; EW, KL, LA and AD supervised, resources, Methods, BB. Reviewed & revised the manuscript; FA, supervised, resources, data validation, reviewed and revised the manuscript.

Acknowledgement

Authors highly acknowledge the Academy of Finland, Helsinki University, Hawassa University, and Addis Ababa University for collaboratively funding this work.

Data Availability

16S rRNA sequenes were deposited in the NCBI genbank. accession numbers are: EU266588,HM137034,DQ536511,AB102735,AY972446,AB377521,EU399926,DQ849078,HQ164548,HM222691,FJ755919, AY461699,FJ561295,FJ784127,AB066339,HM584906,DQ207559,EU855197,GU252111,GU339273,FM164631,AB118223, AB330406,GU368375,EF114313,HQ171486,AB308444,FM164633,AY310732,AB675636,GQ478258,HM585374,EU304248, FJ190535,EU331414,AB435591,GQ451698,CP000653,FJ463411,EU855208,EU221358,EF175731,DQ122370,GQ259887,EU078621, FJ707292,HQ242723,AB495172,GU368371,GU186121,FU550141,AJ717361,EU704697,AB186915,AM421978,GU204957,CP002206, DQ122372,GU726581,GU358073,FJ527671,HM045838,AY569299,EU934229,GU475123,GU367870,FJ595998,EU074181,EF149005, HM582866,HM852118,EU637927,AF511524,CP000826 and AB334774.

References

1. Tesfaye Molla G, et al. Socioeconomic and biophysical drivers of agroforestry practices among smallholder farmers in Ethiopia. Land Use Policy. 2023;125:106–20.
2. Latamo Lameso L. Traditional agroforestry practices and the role of indigenous legumes in Southern Ethiopia. J Agric Rural Dev. 2021;14(2):88–102.
3. Dash D, Deole S. Role of Rhizobium in nitrogen fixation and plant growth promotion. Int J Chem Stud. 2019;7(4):223–7.
4. Fassil Assefa, Kleiner D. Rhizobium-legume symbiosis in Ethiopia: Effectiveness of indigenous isolates. Biol Fertil Soils. 1997;24:212–9.
5. Wolde-meskel E, et al. Genetic diversity and phylogeny of rhizobia isolating from Southern Ethiopia. Mol Ecol. 2005;14(6):1759–70.
6. Degefu T, et al. Phylogenetic diversity of rhizobia nodulating Acacia abyssinica and Acacia negrii in Ethiopia. Int J Syst Evol MicroBiol. 2011;61(1):71–80.
7. Aregu Amsalu G, Assefa F, Kleiner D. Characterization of rhizobia nodulating Millettia ferruginea (Hochst.) Baker in Southern Ethiopia. Afr J Biotechnol. 2013;12(15):1845–53.
8. Mulleta D. (2008) Microbial ecology of indigenous agroforestry systems in Southern Ethiopia. [Doctoral dissertation, Swedish University of Agricultural Sciences].
9. Beyene D, et al. Mycorrhizal associations in indigenous legumes of Ethiopia. Ethiop J Biol Sci. 2016;15(2):101–15.
10. Berza B, Pagano MC, Prabavathy VR, Belay Z, Assefa F. Arbuscular mycorrhizal status of *Erythrina brucei* in different land use types in Ethiopia. Appl Soil Ecol. 2021b;165:104018.
11. Berza B, Sekar J, Ramalingam PV, Pagano MC, Assefa F. Genetically and functionally diverse symbiotic and non-symbiotic native bacteria colonized root nodules of *Erythrina brucei* growing in different land use types in Ethiopia. Rhizosphere. 2021a;17:100301.
12. Debnath S, et al. Diversity and plant growth-promoting potential of non-rhizobial endophytes isolated from legume nodules. Curr Microbiol. 2023;80(5):154.
13. Ben Gaied S, et al. Stable and functional non-rhizobial nodule endophytes: Impact on plant growth and nodulation. Microb Ecol. 2024;87(1):12–25.
14. Ahmed Idris A, et al. The functional microbiome of legume nodules: Beyond the rhizobia-legume symbiosis. Journal of Applied Microbiology, Advance online publication; 2025.
15. Hnini M, Aurag J. Plant growth promotion and environmental stress resistance conferred by non-rhizobial endophytes in legumes. Symbiosis. 2024;92:45–58.

16. Wolde-meskel E, Terefework Z, Lindström K, Frostegård Å. Rhizobia nodulating African Acacia spp. and Sesbania sesban trees in southern Ethiopian soils are metabolically and genomically diverse. Soil Biol Biochem. 2004;36(12):2013–25.
17. Somasegaran P, Hoben HJ. Handbook for Rhizobia. Methods in Legume-Rhizobium Technology. New York: Springer-; 1994. pp. 240–58.
18. Vincent JM. A manual for the practical study of the root-nodule bacteria, International Biological Programme Handbook No. 15. Oxford: Blackwell scientific; 1970.
19. Wilson R, Ainscough R, Anderson K, Baynes C, Berks M, Bonfield J, Wohldman P. 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans. Nature. 1994;368(6466):32–8.
20. Liu L, Hao Z, Weng J, Li M, Zhang D, Bai L, Zhang S. Identification of drought-responsive genes by cDNA-amplified fragment length polymorphism in maize. Ann Appl Biol. 2012;161(3):203–13.
21. Asrese Amsalu A, Assefa F, Hailemariam A. Symbiotic and phenotypic characterization of Rhizobium isolates of field pea (Pisum sativum L.) Fabaceae, from central and southern Ethiopia. Ethiop J Biol Sci. 2012;11(2):163–79.
22. Haukka K, Lindström K, Young JPW. Three phylogenetic groups of nodA and nifH genes in Sinorhizobium and Mesorhizobium isolates from leguminous trees growing in Africa and Latin America. Appl Environ Microbiol. 1998;64(2):419–26.
23. Laguerre G, Bardin M, Amarger N. Isolation from soil of symbiotic and nonsymbiotic Rhizobium leguminosarum bv. viciae and characterization by DNA hybridization and PCR-RFLP of 16S rDNA. Can J Microbiol. 1993;39(12):1142–9.
24. Soe KM, Bhromsiri A, Yamakawa T. Differential responses of soybean cultivars to inoculation with Bradyrhizobium and Rhizobium strains in Thailand. J Fac Agric Kyushu Univ. 2012;57(1):1–7.
25. Chidebe IN, Jaiswal SK, Dakora FD. Distribution and phylogeny of Rhizobium and Bradyrhizobium symbionts nodulating cowpea [Vigna unguiculata (L.) Walp.] in different agro-ecologies of South Africa. Front Microbiol. 2018;9:3033.
26. Ramírez MA, Peix A, Rivas R, Mateos PF, Martínez-Molina E, Velázquez E. Diversity of Rhizobium and Bradyrhizobium strains nodulating beans in the Americas. Front Microbiol. 2019;10:2410.
27. Lopez SM, Jaiswal SK, Dakora FD. Legume-rhizobia symbiosis: Mechanisms and application in agricultural practice. South Afr J Bot. 2020;130:1–11.
28. Favero VO, Jaiswal SK, Dakora FD. The taxonomy and phylogeny of Rhizobium and Bradyrhizobium isolates nodulating soybean in Brazil. Sci Rep. 2021;11:14210.
29. Mayhood P, Mirza BS. Diversity and population structure of Rhizobium and Bradyrhizobium symbionts of mung bean (Vigna radiata) in Pakistan. Appl Soil Ecol. 2021;162:103893.
30. Silva K, de Meyer SE, Rouws LF, Farias SG, dos Santos JM, O'Hara GW, Zilli JE. Bradyrhizobium ingae sp. nov., isolated from effective nodules of Inga laurina in Brazil. Int J Syst Evol MicroBiol. 2014;64(10):3395–401.
31. Tian H, Chen G, Lu C, Xu X, Ren W, Zhang B, Wofsy S. Global methane and nitrous oxide emissions from terrestrial ecosystems due to multiple environmental changes. Ecosyst Health Sustain. 2015;1:1–20. https://doi.org/10.1890/EHS14-0015.1.
32. Hoque S, Broadhurst L, Thrall P. Genetic characterization of rhizobial populations associated with Acacia salicina and A. stenophylla in south-eastern Australia. J Appl Microbiol. 2011;111(1):163–75.
33. Muresu R, Sulas L, Squartini A. The dominance of Gammaproteobacteria in the nodules of Hedysarum species. Microb Ecol. 2010;60:455–61.
34. Ibanez F, Angelini J, Taurian T, Luduena L, Fabra A, Valetti L. Endophytic gammaproteobacteria and alphaproteobacteria isolated from Arachis hypogaea L. nodules. Arch Microbiol. 2009;191(10):731–9.
35. Young JPW, Crossman LC, Johnston AWB, Kanaeko T, Newcomb W, Roe BA. The genome of Rhizobium leguminosarum has a high capacity for symbiotic gene transfer. J Bacteriol. 2006;188(13):4879–90.
36. Sullivan JT, Patrick HN, Lowther WL, Scott DB, Ronson CW. (1995) Nodulating strains of Rhizobium loti arise through a horizontal transfer of symbiotic genes. Proceedings of the National Academy of Sciences, 92(19), 8985–8989.
37. Segovia L, Pinero D, Palacios R, Martinez-Romero E. Genetic structure of a Rhizobium leguminosarum biovar phaseoli population: Lack of correlation between effector and housekeeping genes. Appl Environ Microbiol. 1991;57(2):426–33.
38. Li YH, Wang ET, Kan FL, Wang DZ, Chen WX, Sui XH. AFLP and MLSA analysis of Rhizobium, Sinorhizobium and Mesorhizobium strains. Syst Appl Microbiol. 2009;32(4):233–42.

39. Ribeiro RA, Ormeño-Orrillo E, Indrasumunar A, Gigot K, Giraud E, Scott PT, Martinez-Romero E. Reclassification of Rhizobium tropici Type A strains as Rhizobium leucaenae sp. nov. Int J Syst Evol MicroBiol. 2011;61:1–10.

Figures

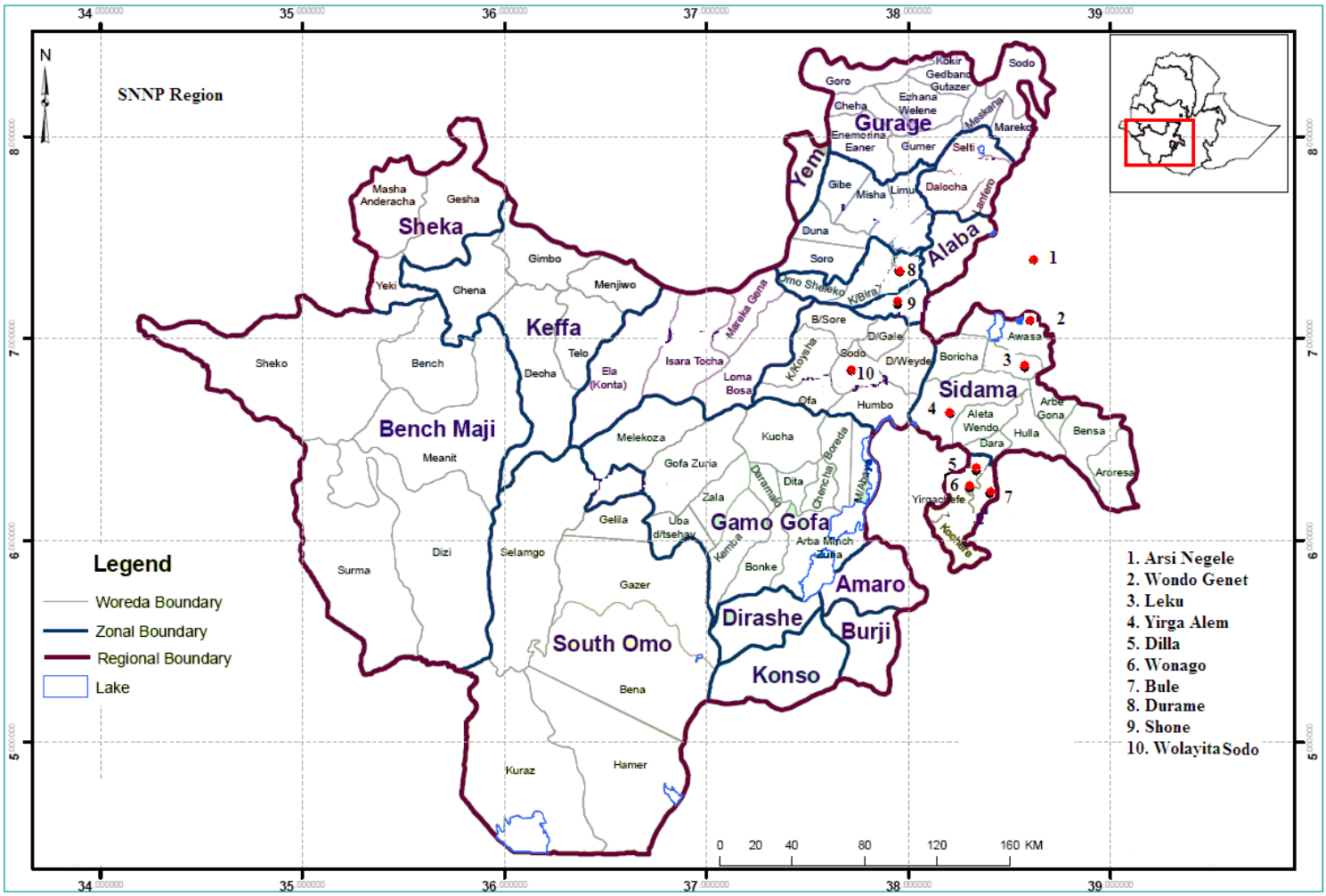


Figure 1

Geographic location of the ten sampling sites (pink dots). All the sampling sites are located in the Southern Nations and Nationalities Peoples (SNNP) Regional state of Ethiopia; except Arsi Negele, which is situated in the Oromiya National Regional state of Ethiopia.

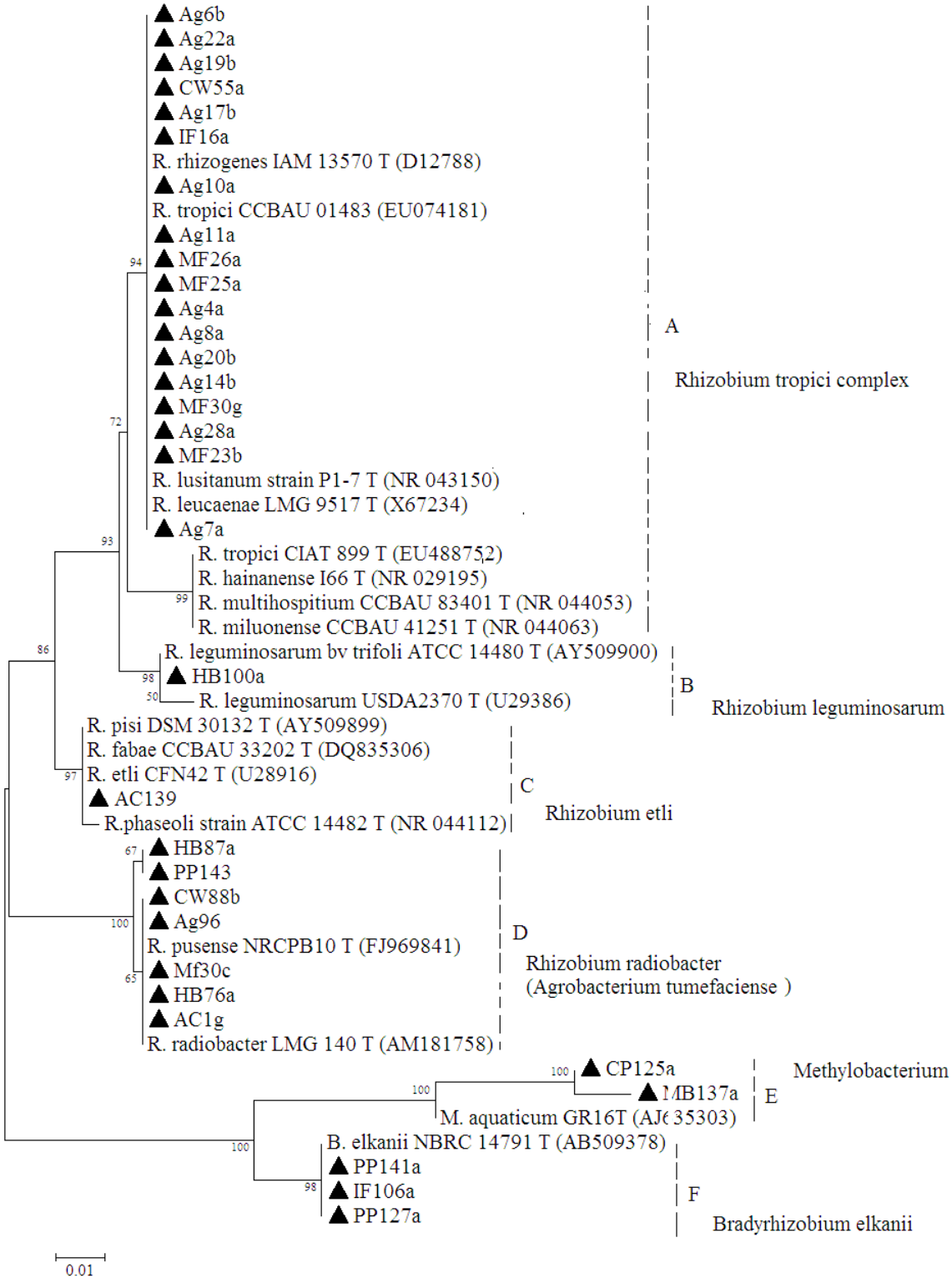


Figure 2

Phylogram (neighbor joining tree) based on partial 16S rDNA sequence (414bp) outlined the generic level relationship of the isolates related to rhizobial gnera (32 strains ▲) to reference type strains. Tree constructed using MEGA11 software.Percentages of bootstrap test values >50% of 500 replicates shown next to branches, T - after strain number indicates type strain, GeneBank accession numbers in brackets, bar scale represents estimate of % of nucleotide substitution rate.

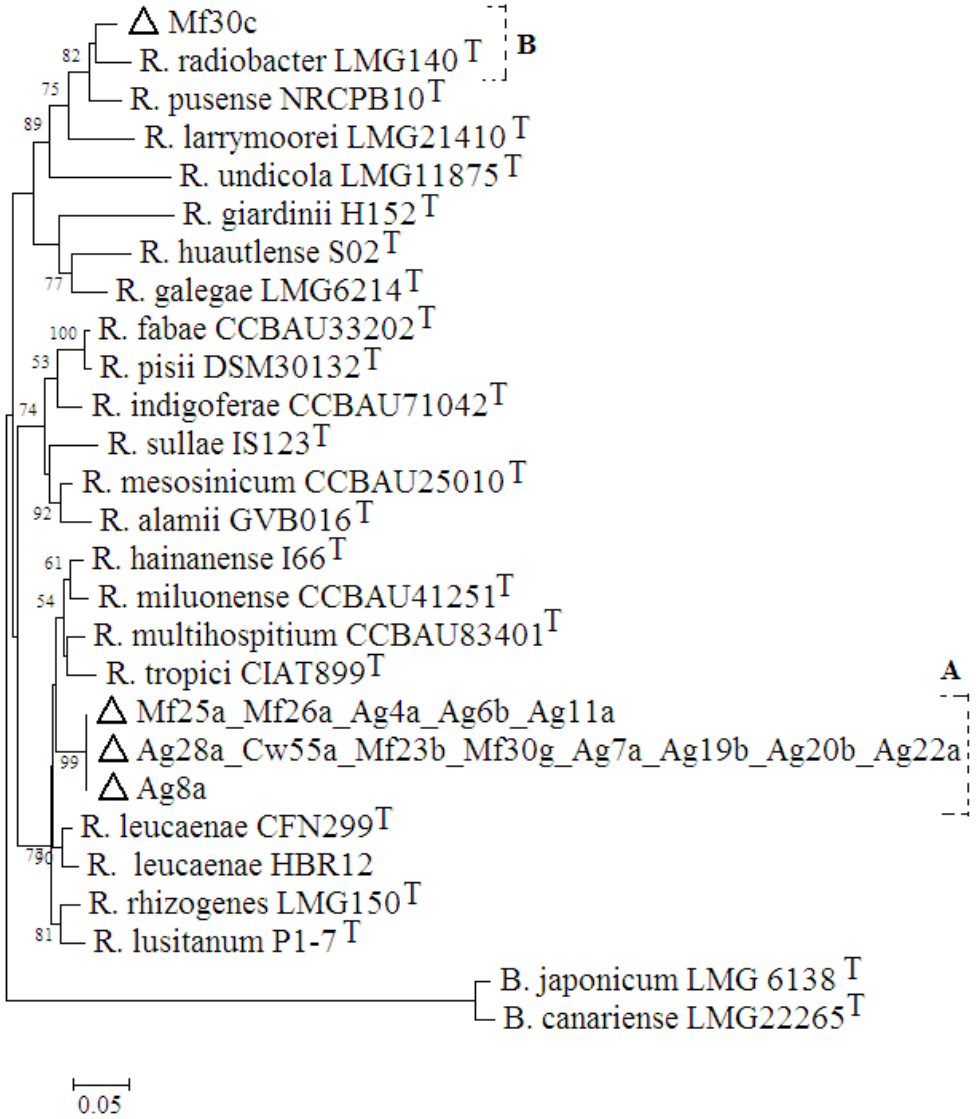


Figure 3

Maximum likelihood phylogenetic tree based on concatenated sequences of three housekeeping genes (*rec*A=387 bp + *srr*=820 bp + *rpo*B=401 bp with size of 1608 bp) revealed the relationship of non-symbiotic rhizobia (Δ 15 isolates) and reference type strains of rhizobia (superscript). The tree was constructed using Mega11 and bootstrap values >50% of 500 replicates are shown at branch points, scale bar represent estimated % nucleotide substitution rate, *Bradyrhizobia*used as out group.

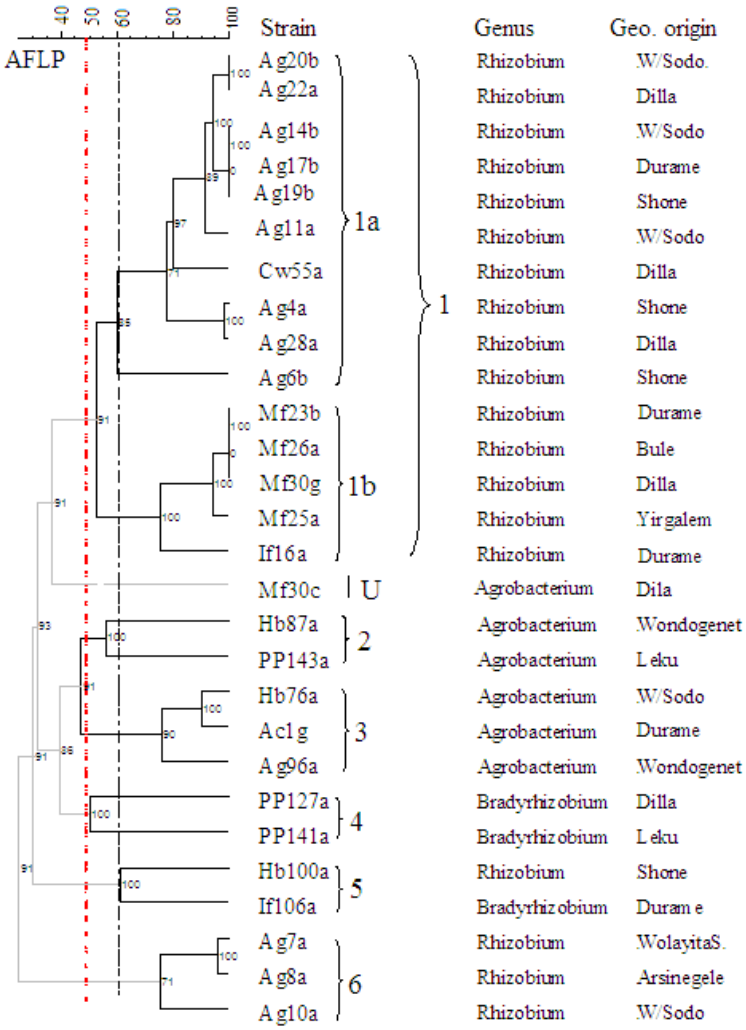


Figure 4

AFLP dendrogram of bacterial isolates (28) related to rhizobial genera (Band-based, Dice cof, 0.5% optimization and 1% position tolerance). The isolates separated at 50% similarity coefficient into their respective genera with few exceptions (AFLP 5).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.docx
- SupplementaryTables.docx