Molecular Phylogenetics and Evolution 214 (2026) 108471

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Argyrolobium legumes from an African centre of endemism associate with novel *Bradyrhizobium* species harbouring unique sets of symbiosis genes

Mabodiba M. Maake [a,*], Chrizelle W. Beukes [a,b], Magriet A. van der Nest [a], Juanita R. Avontuur [c], Esther K. Muema [d], Tomasz Stępkowski [e], Stephanus N. Venter [a], Emma T. Steenkamp [a,*]

[a] Department of Biochemistry, Genetics, and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa
[b] Department of Ecological Sciences, The James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK
[c] Marine Microbiology Group (MMG), Department of Animal and Microbial Biodiversity, Mediterranean Institute for Advanced Studies (IMEDEA, CSIC-UIB), Esporles, Spain
[d] Department of Soil Science, Faculty of AgriSciences, Stellenbosch University, South Africa
[e] Department of Biochemistry and Microbiology, Institute of Biology, Warsaw University of Life Sciences (SGGW), Poland

ARTICLE INFO

Keywords:
Phylogeography
Grassland biome
Great Escarpment
Nod factor
Sulfation
Fucosylation

ABSTRACT

Given that several, mainly endemic South African Genisteae genera occupy basal positions in legume phylogenetic trees, this region of Africa is considered a primaeval centre of diversification of this legume tribe. Despite the importance of South Africa in Genisteae evolution, almost all studies have focused on rhizobia nodulating Genisteae in their centres of diversity in either the Mediterranean Basin or the Americas. Therefore, this study aimed to identify and characterize rhizobial strains associated with *Argyrolobium* species native to areas of the Grassland biome associated with the Great Escarpment, which dominates the subcontinent's eastern landscape, and compare these to bradyrhizobia nodulating Genisteae in other centres of diversity. Phylogenetic analyses of five housekeeping genes (*dnaK, glnII, gyrB, recA,* and *rpoB*) separated the 18 *Bradyrhizobium* strains examined into five well-supported groups. Three of these were conspecific with *B. arachidis*, *B. brasilense/B. australafricanum* and *B. ivorense*, while the remaining two appeared to be new to science. After confirming their novelty using Average Nucleotide Identity, a metric for genome relatedness, and certain phenotypic traits, we recognized them as novel species for which we proposed the names *B. spitzkopense* sp. nov. (Arg816^{Ts}) and *B. mpumalangense* sp. nov. (Arg237L^{Ts}). Phylogenetic analyses of *nodA* gene sequences showed that about half of the strains examined, irrespective of their species identity, harboured alleles known only from the Grassland biome along the Great Escarpment that were previously detected in *Bradyrhizobium* strains nodulating Crotalarieae endemic to this region. Genome-based analyses of data from this and previous studies further showed that strains with these unique *nodA* alleles typically encode the *nodH* gene, the product of which adds a sulfate moiety to the Nod factor (the signalling molecule for establishing the nitrogen-fixing symbiosis). The remaining strains had *nodA* alleles commonly encountered elsewhere in South Africa and other tropical regions of the world. Also, the genomes of these other strains lacked *nodH* but encoded *nodZ*, the gene involved in the fucosylation of the Nod factor. Our findings, therefore, showed that the root nodules of Genisteae (and its sister tribe Crotalarieae) native to the Grassland biome along the Great Escarpment are often related *Bradyrhizobium* strains that are distinct from bradyrhizobia nodulating Genisteae in the Mediterranean and the Americas.

1. Introduction

In recent years, the symbiosis between rhizobia and leguminous plants (family Leguminosae) has emerged as an interesting model for exploring the biogeography of plant–microbe interactions (Beukes et al., 2016; Han et al.; Sprent, 2007; Stepkowski et al., 2018; Tian et al., 2010; Vinuesa et al., 2008). Rhizobia are soil bacteria capable of infecting legume roots/stems, resulting in the formation of nodules within which the bacteria convert inert atmospheric nitrogen to plant-accessible ammonia in exchange for plant photosynthates as nutrients (Held

* Corresponding authors.
E-mail addresses: mabodibamaake@gmail.com (M.M. Maake), emma.steenkamp@up.ac.za (E.T. Steenkamp).

https://doi.org/10.1016/j.ympev.2025.108471
Received 14 November 2024; Received in revised form 23 March 2025; Accepted 12 September 2025
Available online 14 September 2025
1055-7903/© 2025 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

et al., 2010; Kereszt et al., 2011; Santi et al., 2013). In rhizobia, the genetic determinants for this nitrogen-fixing symbiosis forms part of the highly variable accessory genome, that governs adaptive or niche-associated traits (Aoki et al., 2013; Ormeno-Orrillo et al., 2013; Barcellos et al., 2007). Within this sub-genomic compartment, nodulation is determined by products of the *nod*, *nol* and *noe* genes, while products of the *nif* and *fix* genes mediate nitrogen fixation (Kaneko et al., 2011; Kaneko et al., 2002).

Unlike the core genome that is conserved among all members of a species and stably passed on from parent to offspring, the accessory genome with its symbiosis-associated genes is prone to horizontal gene transfer (HGT) and frequently exchanged among individuals of a species and even between species and genera (Barcellos et al., 2007; Ormeno-Orrillo and Martinez-Romero, 2019; Parker et al., 2002; Vinuesa et al., 2005a,b). Accordingly, these two sub-genomic compartments have different evolutionary trajectories (Stepkowski et al., 2007). While the core genome reflects species ancestries, the accessory genome reflects the adaptation of symbiotic properties across geographic space and through geologic time (Han et al.; Tian et al., 2010; Vinuesa et al., 2008). Indeed, studies on Leguminosae typically suggest that various rhizobial species may nodulate a particular legume species, but that their symbiotic abilities coevolved with rhizobial ancestors present in the native habitat of the legume hosts (Beukes et al., 2016; Parker, 2004; Steenkamp et al., 2008; Stepkowski et al., 2012; Tian et al., 2010).

Among rhizobia, those in the genus *Bradyrhizobium* (family *Nitrobacteraceae*, class *Alphaproteobacteria*, phylum *Pseudomonadota*) are the most widely distributed and have the broadest host range (Urquiaga et al., 2019; Parker, 2004, 2015; Stepkowski et al., 2018). More than 90 *Bradyrhizobium* species have been named (as of October 2024) according to the List of Prokaryotic names with Standing in Nomenclature (LPSN) (Parte et al., 2020), with numerous known strains still awaiting formal species description (e.g., Bouznif et al., 2019; Costa et al., 2023; Stepkowski et al., 2012). Although no clear biogeographic patterns have been reported for *Bradyrhizobium* species, the phylogenetic signal contained within sequences of certain symbiosis genes broadly correlates with the geographic origins of the legume hosts with which the rhizobia interact (e.g., Banasiewicz et al; 2021; Costa et al., 2023; Stepkowski et al., 2018). This has been well illustrated using sequences for the *nodA* gene (Moulin et al., 2004; Stepkowski et al., 2005; Steenkamp et al., 2008; Beukes et al., 2016; Stepkowski et al., 2018). This gene is found in most rhizobia and is needed for producing the Nod factor signalling molecule essential for establishing the symbiosis (the *nodA* gene product transfers an acyl group to the non-reducing end of the lipochitooligosaccharide backbone of the Nod factor) (Atkinson et al., 1994; Röhrig et al., 1994). Phylogenies based on *nodA* sequences separate *Bradyrhizobium* into clades that correspond to particular geographic regions (Moulin et al., 2004; Stepkowski et al., 2005; Steenkamp et al., 2008; Beukes et al., 2016; Stepkowski et al., 2018). The *nodA* clades often also correlate with a subset of nodulation genes (or their alleles) determining the rhizobial symbiont's ability to interact with certain legume species (Stepkowski et al., 2007). In other words, sequences for genes (when present) encoding enzymes involved in host-specific ornamentation of the Nod factor (e.g., *nodZ*, *nolL* and *noeI*) generally support the *nodA* clades (Moulin et al., 2004; Stepkowski et al., 2007). Information regarding these host-specific nodulation (*hsn*) genes is thus invaluable for investigating the evolution and dispersal of *Bradyrhizobium* and its legume hosts (Parker, 2004; Rodríguez-Echeverría, 2010; Stepkowski et al., 2018).

In the current study, we explored the diversity and evolution of rhizobial strains isolated from the root nodules of indigenous South African *Argyrolobium* species (sub-family Papilionoideae, tribe Genisteae). This large genus of legumes had its origins in South Africa (SA) (Cardoso et al., 2013; Polhill et al., 2005; Polhill, 1968) from where it spread widely, with extant species occurring in Madagascar and across Africa, southern Europe, the Mediterranean region, into middle-eastern Asia and the Indian subcontinent (POWO, 2019). In SA, *Argyrolobium* is particularly common in the Grassland biome, which spans an area that overlaps the Great Escarpment, a topographical feature dominating the country's eastern landscape (Edwards and van Wyk, 2005; Rutherford and Mucina, 2006). The subtropical climate with low annual minimum temperatures, combined with high altitudes and low soil pH likely shaped the unique assemblage of *Argyrolobium* and other plant species abounding in this biome (Trytsman et al., 2016). However, not much is known about the rhizobial symbionts of legumes associated with the Great Escarpment and its surroundings. This is despite the edaphic, topographic, and climatic heterogeneity linked to this ancient geological formation that has been in its current position since the late Cretaceous period (Clark et al., 2011; Thuiller et al., 2006).

Previous work has suggested that legume-nodulating *Bradyrhizobium* strains from the Great Escarpment likely represent a unique centre of diversity for the genus (Beukes et al., 2016). Although work in the region focused mostly on species from the tribe Crotalarieae, it revealed that Grassland legumes are nodulated by novel *Bradyrhizobium* species (Ardley et al., 2013; Beukes et al., 2016), some of which have since been formally described (Avontuur et al., 2021). The novelty of these bacteria also extended to their symbiotic genes by encoding unique alleles of the *nodA* and *nifD* genes (*nifD* encodes the α subunit of dinitrogenase) (Avontuur et al., 2021; Beukes et al., 2016). This was evident from their clustering in *nodA* Clades XIII, XIV and XV (Beukes et al., 2016) that have so far only been detected in biomes along the Great Escarpment (Avontuur et al., 2021; Beukes et al., 2016). Legumes from other parts of SA and elsewhere in the world are nodulated by *Bradyrhizobium* strains with *nodA* sequences from clades I to XII and XVI (Beukes et al., 2016, Stepkowski et al., 2018, Steenkamp et al., 2008). Also, the *Bradyrhizobium* symbionts of native Genisteae sampled elsewhere in the world mostly harbour *nodA* alleles forming part of Clades I, III, IV, VII, and II in particular (Beukes et al., 2016; Stepkowski et al., 2018). In fact, Clade II's predominance amongst Genisteae symbionts from the Mediterranean, Europe, western North America and the Andes is thought to reflect the enduring coevolutionary relationship between the nodulation genes of these microsymbionts and their legume hosts (Stepkowski et al., 2007, 2018).

Here our overall goal was to determine whether the previously observed patterns inferred from the rhizobia of indigenous Crotalarieae and a small number of *Argyrolobium* species hold true when examining a more extensive collection of rhizobia obtained from *Argyrolobium* sampled in the Grassland biome along the Great Escarpment. For this purpose, we targeted a collection of *Bradyrhizobium* strains originating from the root nodules of *Argyrolobium* collected at sites in the Mpumalanga (MP) and KwaZulu-Natal (KZN) Provinces of SA. The first objective was to identify and delineate the bacteria to species level using a taxonomic workflow that starts with multigene-based genealogical concordance analysis followed by determination of a broad range of phenotypic and genomic characters (Venter et al., 2017). The second objective was to provide formal descriptions for any new species delineated, following the rules of the Code of Nomenclature of Prokaryotes described from Sequence Data (SeqCode) (Van Lill et al., 2024; Hedlund et al., 2022). The latter utilizes genome sequences as nomenclatural type, thereby providing an avenue for continued taxonomic exploration while taking into consideration SA's legislative complexities regarding the distribution of indigenous biological material outside the country (da Silva et al., 2023; Hamer et al., 2021; Rahi, 2021). The third objective was to determine the evolutionary history of the *nodA* gene in the sampled strains and whether the pattern observed is reflected in the histories of other symbiotic genes, especially *hsn* genes (Lerouge et al., 1990; Roche et al., 1991; Garcia et al., 1996; Sadowsky et al., 1991; Stepkowski et al., 2007; Mergaert et al., 1997). This would allow us to evaluate two competing hypotheses regarding the coevolution of *Bradyrhizobium* and its symbiosis genes with Genisteae: (i) diversification of *nodA* Clade II occurred in SA during a process linked to the early stages of Genisteae diversification (Martínez-Romero, 2009; Stepkowski et al., 2018); and (ii) *nodA* Clade II emerged elsewhere in the world, while

Genisteae from SA coevolved with strains harbouring a unique version (s) of the gene (Beukes et al., 2016). Our findings would thus contribute towards the description of SA's rhizobial diversity, as well as provide valuable insights into the evolutionary histories of genes involved in Nod factor modification and host range.

2. Materials and methods

2.1. Sampling, DNA isolation and 16S rRNA-based identification

The 18 *Bradyrhizobium* strains included in this study (Table 1) were obtained from the root nodules of an unidentified *Argyrolobium* species, *A. harveyanum* (a slender herb with small tuberous roots), *A. rupestre* (a low-growing herb with a woody taproot), and *A. robustum* (a large herb colloquially known as liquorice bean) (Edwards and van Wyk, 2005). Five of these strains were isolated in August 2010 and May 2011 from nodules of *A. rupestre* collected in the Mesic Highveld Grassland bioregion (Rutherford and Mucina, 2006) near Barberton (MP) and six strains came from the nodules of the unidentified *Argyrolobium* sp. sampled in August 2010 from the Mesic Highveld Grassland bioregion near eManzana (MP). The remaining strains were obtained during October 2019 from root nodules of *A. harveyanum* (n = 6) sampled in the Sub-Escarpment Grassland bioregion near Spitzkop (KZN) and *A. robustum* (n = 1) sampled from the Mesic Highveld Grassland bioregion in the Mkhondo (MP) area (Fig. 1A). Root nodules were collected in the field with landowner permission and stored in screwcap vails containing silica gel until processing in the laboratory, 3 to 5 days later.

Bacterial isolation from root nodules were performed as previously described (Howieson and Dilworth, 2016). In short, individual nodules were surface sterilized by soaking in 3.5 % (m/v) sodium hypochlorite for three minutes, followed by immersion in 70 % (v/v) ethanol for 30 s. After five rounds of rinsing with sterile water, the nodule was crushed onto tryptone yeast agar (TYA) containing 5 g/L tryptone (Oxoid, Basingstoke Hampshire, UK), 3 g/L yeast extract (VWR chemicals, Leuven, Belgium) and 15 g/L agar (VWR chemicals), supplemented with 4.4 g/L calcium chloride dihydrate (Howieson and Dilworth, 2016). The resulting plates were incubated at 28 °C until growth was observed. Pure cultures of the isolates are also preserved at −80 °C in 20 % (v/v) glycerol and on Microbank™ storage beads (Pro-lab Diagnostics, Richmond Hill, Canada). Isolates Arg816^{Ts}, 31Argb, and Arg237L^{Ts} were submitted to the South African Rhizobium Culture Collection (SARCC, Pretoria, SA; with numbers SARCC-3537, SARCC-3536, and SARCC-3538) and the culture collection of the Forestry and Agricultural Biotechnology Institute (FABI) at the University of Pretoria with numbers CMW65401, CMW65403, and CMW65402, respectively.

DNA was extracted from fresh TYA cultures using the Quick-gDNA™ MiniPrep kit according to the manufacturer's instructions (Zymo Research, Freiburg, Germany) and quantified using a NanoDrop™ spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The genus-level identity of these strains was then confirmed using sequences of the 16S ribosomal RNA (rRNA) gene. This was achieved by amplifying and sequencing the the gene as described previously (Beukes et al., 2016) and then using BLASTN (Altschul et al., 1990) searches to confirm that they share > 95 % 16S rRNA gene sequences similarity to those of known *Bradyrhizobium* strains in the nucleotide database of the National Centre for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov).

2.2. Genealogical concordance analysis

The 18 strains were separated into putative species using genealogical concordance analysis as described by Venter et al. (2017). For this purpose, we utilized the nucleotide sequences for five protein-coding genes (*dnaK* encoding chaperone DnaK, *glnII* encoding glutamine synthetase isoform II, *gyrB* encoding DNA gyrase subunit B, *recA* encoding recombinase A, and *rpoB* encoding the beta subunit of RNA polymerase) routinely used for the delineation *Bradyrhizobium* species. Portions of these genes were amplified using primers and reaction conditions previously reported (Beukes et al., 2016). See Supplementary Table S1 for primers and PCR cycling conditions. Amplicons were purified using 2 U μl^{-1} Exo I and 0.1 U μl^{-1} FastAP (Thermo Fisher Scientific) according to the manufacturer's protocol and sequenced using the original PCR primers, the BigDye Terminator v3.1 Cycle Sequencing kit (Thermo Fisher Scientific) and an ABI3100 Automated Capillary DNA sequencer (Thermo Fisher Scientific). Forward and reverse sequences were edited and trimmed with ChromasLite v2.01 (Technelysium, Brisbane, Australia) and assembled into consensus sequences using BioEdit v7.05 (Hall, 1999) and submitted to NCBI (https://submit.ncbi.nlm.nih.gov/) under accession numbers OP888049-OP888057 (*dnaK*), OP897847-OP897855 (*glnII*), OP709859-OP709867 (*gyrB*), OP880329-OP880337

Table 1
Strain numbers and *Argyrolobium* hosts of *Bradyrhizobium* species examined in this study, as well as the collection sites of root nodules in the Mpumalanga (MP) and Kwazulu-Natal (KZN) provinces in South Africa. Their nodulation ability on cowpea and siratro, and *nodA* clade designations are also shown.

Isolate	*Bradyrhizobium* species	*Argyrolobium* host	Geographic location	Grassland bioregion[a]	Nodulation[b]	*nodA* Clade[c]
Arg237L^{Ts}	*B. mpumalangense* sp. nov.	*Argyrolobium robustum*	Mkhondo area, MP	Mesic Highveld	Nod-	–
31Arga	*B. mpumalangense* sp. nov.	*A. rupestre*	Barberton area, MP	Mesic Highveld	Nod-	XV
31Argb	*B. mpumalangense* sp. nov.	*A. rupestre*	Barberton area, MP	Mesic Highveld	Nod-	XV
Arg816^{Ts}	*B. spitzkopense* sp. nov.	*A. harveyanum*	Spitzkop area, KZN	Sub-Escarpment	C, S	III
Arg826	*B. spitzkopense* sp. nov.	*A. harveyanum*	Spitzkop area, KZN	Sub-Escarpment	C, S	III
Arg836	*B. spitzkopense* sp. nov.	*A. harveyanum*	Spitzkop area, KZN	Sub-Escarpment	C, S	III
Arg846	*B. spitzkopense* sp. nov.	*A. harveyanum*	Spitzkop area, KZN	Sub-Escarpment	C, S	III
Arg856	*B. spitzkopense* sp. nov.	*A. harveyanum*	Spitzkop area, KZN	Sub-Escarpment	C, S	III
Arg866	*B. spitzkopense* sp. nov.	*A. harveyanum*	Spitzkop area, KZN	Sub-Escarpment	C, S	III
62Argb22	*B. brasilense/australafricanum*	*Argyrolobium* sp.	eManzana area, MP	Mesic Highveld	C, S	XV
62Argb07	*B. brasilense/australafricanum*	*Argyrolobium* sp.	eManzana area, MP	Mesic Highveld	C, S	XV
62Arga22	*B. brasilense/australafricanum*	*Argyrolobium* sp.	eManzana area, MP	Mesic Highveld	C, S	XV
68Arga19	*B. ivorense*	*Argyrolobium* sp.	eManzana area, MP	Mesic Highveld	C	XV
68Argb19	*B. ivorense*	*Argyrolobium* sp.	eManzana area, MP	Mesic Highveld	C	XV
68Argb17	*B. ivorense*	*Argyrolobium* sp.	eManzana area, MP	Mesic Highveld	C	XV
105Arga19	*B. arachidis*	*A. rupestre*	Barberton area, MP	Mesic Highveld	C, S	III
105Argb19	*B. arachidis*	*A. rupestre*	Barberton area, MP	Mesic Highveld	C, S	III
105Argb17	*B. arachidis*	*A. rupestre*	Barberton area, MP	Mesic Highveld	C, S	III

[a] Bioregions as defined by Rutherford and Mucina (2006).
[b] C = cowpea, S = siratro, and Nod- = no nodules induced on either cowpea or siratro.
[c] *nodA* clades inferred from *nodA* phylogeny (Fig. 2 and S8). The Arg237L^{Ts} strain lacked most nodulation genes, including *nodA*.

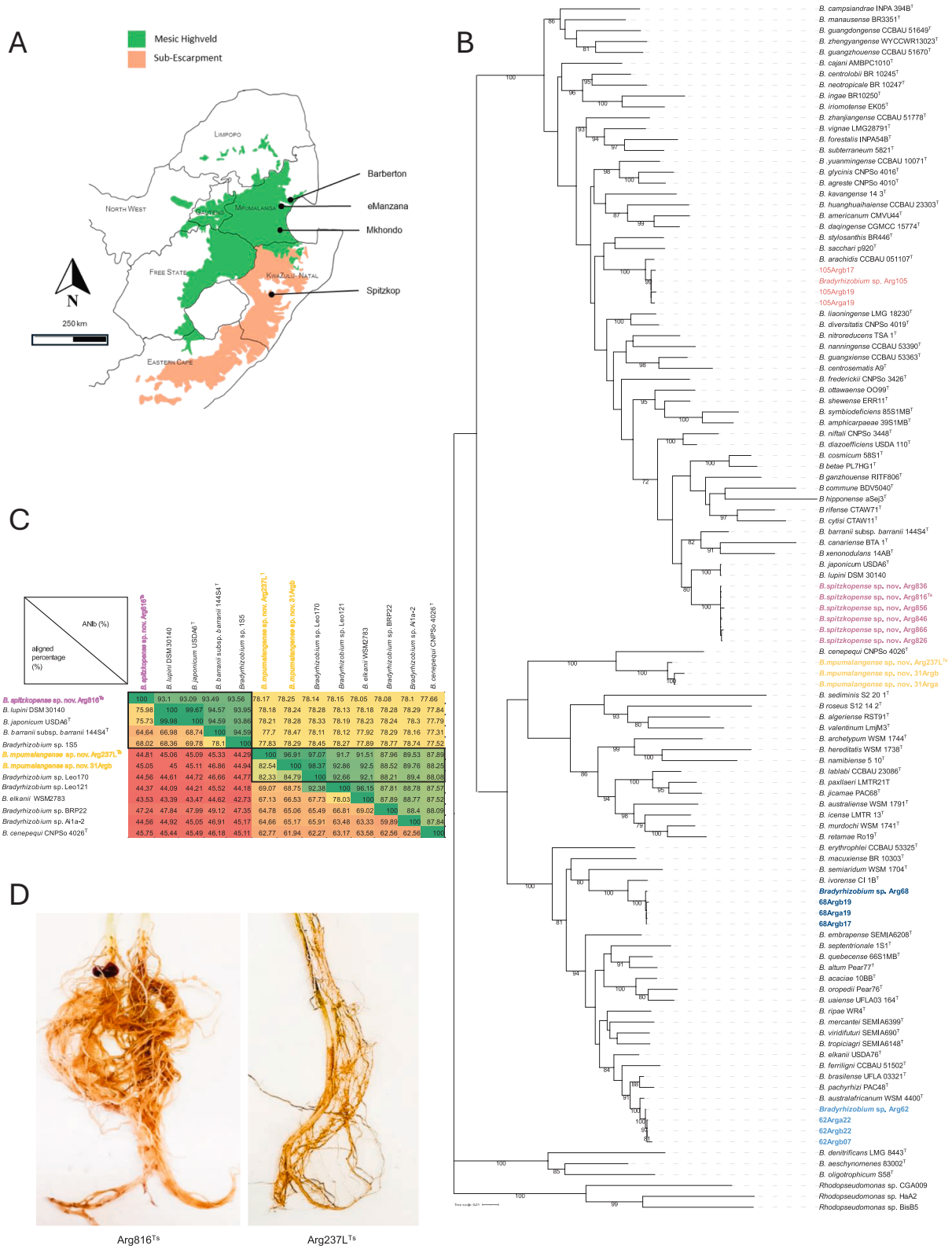

Fig. 1. Summary of the geographic origin, species-level identity and nodulation ability of the strains examined in this study. (A) Map of South Africa indicating the sampling localities and bioregions from which strains originate. (B) Maximum likelihood phylogeny inferred from the concatenated *dnaK, glnII, gyrB, recA* and *rpoB* dataset. Strains from this study are coloured and in bold font. Only bootstrap support values of ≥ 70 % are indicated, and *Rhodopseudomonas* strains were used as the outgroup. The scale bar represents the number of nucleotide changes per site. Supergroups were assigned following Avontuur et al. (2019) and are indicated as in the key provided. Information regarding strains and sequence accession are listed in Suppl. Tables S2 and S3. (C) Results of ANIb-based JSpecies analyses using the genomes of our new species and those of close relatives according to the Genome Taxonomy Database (GTDB). Mean Average Nucleotide Identity (ANI) values are provided in the top-right half of the matrix, while the average proportion of the genomes included in the respective pair-wise comparisons (i.e., the proportion of two genomes that are homologous to one another) are indicated in the left-bottom half of the matrix. (D) Results of nodulation tests on cowpea for *B. spitzkopense* Arg816^{Ts} and *B. mpumalangense* Arg237L^{Ts}. Other members of the two species behaved similarly to these representatives.

(*recA*), OP888040-OP888048 (*rpoB*) for the strains 31Arga, 31Argb, Arg237L^{Ts}, Arg816^{Ts}, Arg826, Arg836, Arg846, Arg856, and Arg866, respectively. The sequences for the remaining strains (62Argb07, 62Arga22, 62Argb22, 68Arga19, 68Argb17, 68Argb19, 105Arga19, 105Arg17, and 105Argb19) have been submitted under accession numbers PV236179-PV236187 (*dnaK*), PV236170-PV236178 (*glnII*), PV236161-PV236169 (*gyrB*), PV236152-PV236160 (*recA*), PV236188-PV236196 (*rpoB*).

For each of the five protein-coding genes, individual datasets were compiled. These contained the sequences generated in the current study, together with those for *Bradyrhizobium* type strains according to the LPSN (https://lpsn.dsmz.de/; Accessed February 2023; Parte et al., 2020). The datasets also included sequences for *Bradyrhizobium* strains reported in the literature that are yet to be named, especially those associated with Genisteae (Suppl. Tables S2 and S3). There were therefore differences in the number of taxa included across the different datasets, with some of the genes (e.g. *glnII* and *recA*) having been sequenced more frequently. Sequences of the five genes for three *Rhodopseudomonas* (family *Nitrobacteraceae*) strains (CGA009, HaA2, and BisB5) obtained from NCBI were included as the outgroup.

The nucleotide datasets were aligned manually using BioEdit based on each gene's inferred amino acid sequences. These single-gene datasets were partitioned and concatenated using FASconCAT-G v1.02 (Kück and Longo, 2014) and subjected to maximum likelihood (ML) phylogenetic analyses using the General Time Reversible (GTR) substitution model (Tavaré, 1986) with parameter optimization in RAxML v8.2.12 (Stamatakis, 2014). The concatenated datasets were analysed in the same way, except that independent parameter optimization for each gene partition was used. In all cases, branch support was estimated using the original model parameters and the standard/traditional nonparametric bootstrap approach based on 1000 pseudoreplicates. The single-gene genealogies were then compared to one another and to the multigene tree to identify the putative novel species, which manifested as consistent and exclusive groups among the taxa included (Venter et al., 2017).

2.3. Whole genome-based analyses

The genomes for representative strains Arg816^{Ts}, Arg237L^{Ts}, and 31Argb were subjected to 150 base pair (bp), paired-end genome sequencing using the Illumina HiSeq platform at the Biotechnology Platform of the Agricultural Research Council (ARC; Onderstepoort, Pretoria, SA). For this purpose, high-quality genomic DNA was extracted using hexadecyltrimethylammonium bromide as described previously (Cleenwerck, 2002). Raw sequence data were quality filtered with FastQC version 0.11.9 (Andrews, et al., 2017), trimmed with Trim Galore (Krueger, 2015) and assembled *de novo* using SPADes v3.12 (Nurk et al., 2013). The genome assemblies for isolates Arg816^{Ts}, Arg237L^{Ts}, and 31Argb were submitted to the NCBI database under the accession numbers JAPRAQ000000000, JAQMIL000000000, and JBEGDB000000000, respectively.

A 92-gene tree was constructed using the Up-to-date Bacterial Core Gene (UBCG) pipeline (Na et al., 2018). The 92-gene dataset included the genome sequences for Arg816^{Ts}, Arg237L^{Ts}, 31Argb, as well as the previously sequenced genomes for strains Arg62 and Arg68 (Avontuur et al., 2021). In addition, this dataset also included all the described *Bradyrhizobium* species for which genome data were available, as well as representative genomes of unnamed *Bradyrhizobium* species according to the Genome Taxonomy Database (GTDB) (Suppl. Table S4) (Parks et al., 2022). Individual gene alignments were constructed with MAFFT v.7.310, and then concatenated and partitioned as described above. The final 92-gene dataset was subjected to ML analysis with IQ-TREE v 1.6.6 (Nguyen et al., 2015) using the best-fit models according to ModelFinder (Kalyaanamoorthy et al., 2017), again with independent parameter optimization for each gene partition, with branch support estimated using 1000 pseudoreplicates for the parametric ultrafast bootstrap approximation (UFBoot) (Minh et al., 2013) and Shimodaira–Hasegawa approximate likelihood ratio test (SH-aLRT) (Guindon et al., 2010).

Additionally, Average Nucleotide Identity (ANI) between strains Arg816^{Ts} and Arg237L^{Ts} and their close relatives (see below), as determined on the basis of the UBCG phylogeny were determined with the online tool JSpeciesWS (Richter et al., 2016). The ANIb option was used to determine the average BLAST-based pairwise identity shared across genomes, where ANI was calculated from 1020 bp homologous regions showing $\geq$ 30 % identity and $\geq$ 70 % alignment length (Richter et al., 2016). The proportion of the respective genomes that were homologous were examined by recording the alignable fragment percentage from which JSpeciesWS calculated the ANIs.

2.4. Phenotypic characterization

A wide range of phenotypic traits were investigated for six representative strains for the proposed species (Arg816^{Ts}, Arg846, Arg866, Arg237L^{Ts}, 31Arga, and 31Argb). Strains were grown on either TYA or yeast mannitol agar (YMA) (Somasegaran and Hoben, 2012) at 28 °C for 7–14 days. Cell morphology and size were determined using a Zeiss Stereo phase-contrast microscope (Microscope Central, Feasterville-Trevose, PA, USA) with associated Auxiovision software v4.8.

Growth was analysed at different temperatures, NaCl concentrations and pH. For this purpose, strains were incubated on YMA at a range of temperatures (4 °C, 10 °C, 15 °C, 20 °C, 25 °C, 28 °C, 30 °C, 35 °C, 37 °C and 40 °C) for 7–10 days, while different NaCl concentrations (0.3, 0.5, 1.0, 1.5, 2.0, and 2.5 % w/v) were tested on YMA with incubation at 28 °C for 7–10 days. To determine the pH range at which each strain can grow, yeast mannitol broth (YMB) was prepared to varying pH levels (4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0) with hydrochloric acid or sodium hydroxide. The resulting flasks were incubated in a shaking incubator at 28 °C for 7–10 days and then evaluated for growth (Gerhardt et al., 1994).

To test the ability of each isolate to produce either an acidic or alkaline environment during growth, YMA was amended with 0.125 % (w/v) bromothymol blue pH indicator (Howieson and Dilworth, 2016). Growth was observed after 3–5 days of incubation at 28 °C, with a yellow colour change recorded as acidic and malachite green or dark blue as basic/alkaline. Oxidase and catalase activity were evaluated as described before, based upon a colour change observed after 20 s after exposure of isolates to 1 % (w/v) N, N, N', N'-tetramethyl ρ-phenylenediamine (DMSO) solution on sterile filter paper (Gerhardt et al., 1994; Zilli et al., 2014).

The BIOLOG GEN III ID MicroPlate (Biolog, Hayward, CA, USA) system was used to evaluate the strains' metabolic capabilities on the basis of 94 phenotypic tests (utilization of 71 carbon sources and sensitivity to 23 chemicals). Additionally, to evaluate the utilization of additional carbon sources as well as nitrate reduction, biochemical tests were performed by inoculating API 20E strips (bioMerieux, Craponne, France). As *Bradyrhizobium* is typically slow growing, the manufacturers' protocols were modified to extend incubation to 7 days.

The nodulation ability of the six representative strains was tested on *Vigna unguiculata* (cowpea) and *Macroptilium atropurpureum* (siratro), which are both known to interact with a relatively wide range of rhizobial symbionts (Howieson and Dilworth, 2016; Lewin et al., 1987). Seeds were soaked in 95 % (v/v) ethanol for 30 s, then in 3.5 % sodium hypochlorite for 6 min. Following rinsing with sterile water, seeds were soaked in sterile water for 3 h and then incubated on water agar (1.5 % w/v) at 28 °C for 7 days in the dark. Two germinated seeds per host were then planted in 1.9L pots containing sterilized sand. Inoculum was prepared by incubating the cultures in Yeast Mannitol Broth (YMB) for 5 days at 28 °C, after which volumes of 2 ml were taken and applied directly onto each of the two seeds per pot, with each inoculum consisting of only one *Bradyrhizobium* strain. Negative controls, representing seedlings that were not inoculated, were included for each legume

species in the experiment. Plants were maintained in a glasshouse (set at 28 °C for 14 h daytime and 15 °C for 10 h nighttime) and watered three times a week using nitrogen free Hoagland solution (Howieson and Dilworth, 2016). The seedlings were grown for six weeks after which nodules were harvested. At the end of the nodulation experiment, the respective root-nodule bacteria were isolated from surface-sterilized nodules, grown on TYA medium and subjected to *rpoB* gene sequence analysis to confirm their identity (by comparison with the *rpoB* sequence of the original strain).

2.5. Analysis of symbiotic genes

Amplification of the *nodA* gene and the *nifH* gene (encoding dinitrogenase reductase) were done as described previously (Beukes et al., 2016; Gronemeyer et al., 2017), for all of the strains being investigated here. Amplicons were purified as described above, or extracted from gel slices following agarose gel electrophoresis, using the Zymoclean Gel DNA Recovery Kit (Zymo Research). All purified amplicons were sequenced as described above, using the same primers as for PCR, after which the sequences were used to compile *nodA* and *nifH* datasets (see Suppl. Tables S5 and S6 for NCBI accession numbers).

For building the *nodA* dataset, data published previously (Avontuur et al., 2021; Beukes et al., 2016) were supplemented with sequences of type strains published since 2022 (Suppl. Table S5). The *nifH* dataset comprised all *Bradyrhizobium nifH* sequences deposited in the NCBI nucleotide database (accessed 1 June 2023) (Suppl. Table S6). For the known species recovered in this study, a relevant representative was included in these datasets. For outgroup purposes, the *nodA* sequence for *Methylobacterium nodulans* ORS2060 was included, and for the *nifH* dataset sequences of *Rhodopseudomonas* strains CGA009, BisB5, and HaA2 were incorporated. The nucleotide sequences for both datasets were aligned using the online version of MAFFT with the auto strategy (Katoh et al., 2019). Both alignments were parsed into haplotypes using DnaSP v5.5 (Librado and Rozas, 2009) to produce datasets in which taxa with identical sequences were represented by a single sequence/haplotype. During the generation of the haplotype files, gaps were treated as fifth characters and invariable sites were retained. The two haplotype nucleotide datasets were then converted to amino acid sequences and subjected to ML-based phylogenetic analysis in RAxML v8.2 using the best-fit model parameters as indicated by ProtTest 3.4 (Abascal et al., 2005) and 1000 pseudoreplicates of non-parametric bootstrap for estimating branch support.

The presence of known *hsn* genes were explored in the genomes of isolates Arg816^{Ts}, Arg237L^{Ts}, 31Argb and other southern African strains of *Bradyrhizobium* available from NCBI (Perret et al., 2000; Shang et al., 2022). This included *nodZ* that codes for a fucosyltransferase that adds a fucose (6-Deoxy-L-galactopyranose) moiety to the terminal reducing N-acetylglucosamine end of the Nod factor, as well as the *nodH* gene which encodes for a sulfotransferase that specifically adds a sulfate moiety to the terminal reducing N-acetylglucosamine end of the Nod factor. We also included *noeI* and *noeE,* of which the methyltransferase and sulfotransferase products are fucose-specific and add a methyl group and a sulfate group to the fucose, respectively.

The *hsn* genes were identified using amino acid sequences for known *Bradyrhizobium* homologues (Leo170 for *nodH*, USDA110 for *nodZ*, and CB756 for *noeE* and *noeI*) as the query in TBLASTN v.2.9.0 searches performed in CLC Workbench v20.0.4 against translated genome sequences included in a locally created database containing the genomes for Arg816^{Ts}, Arg237L^{Ts}, 31Argb, as well as 870 *Bradyrhizobium* genomes obtained from the GTDB. Additionally, these genes were searched for in NCBI's non-redundant nucleotide database using BLASTn with a percentage identity threshold of 60 %, an E-value cut-off of $1x10^{-5}$, and a minimum alignment length of 60 % of the query sequence. The respective lists of best-match sequences were subsequently cross-verified manually against the initial TBLASTN search results. Together, the identified sequences were used to compile comprehensive datasets containing amino acid sequences inferred for the *nodH, nodZ, noeE, noeI* genes. Following alignment with MAFFT, the individual datasets were subjected to ML analysis on the IQ-TREE web server (http://iqtree.cibiv.univie.ac.at/) using the best-fit models according to ModelFinder. Branch support was estimated using 1000 UFBoot and SH-aLRT branch test. All phylogenetic trees were visualized with Mega v7.0.26 (Kumar et al., 2016) and edited in iTOL v6.9 (Letunic and Bork, 2024) or Inkscape v.1.2 (https://www.inkscape.org).

3. Results

3.1. Delineation of putative species

Genealogical concordance analysis based on five, protein-encoding housekeeping genes was used for identifying the 18 *Bradyrhizobium* strains to species level. However, the number of taxa available for the five genes differed, especially regarding strains associated with Genisteae legumes (see Suppl. Table S3). The aligned *dnaK* dataset comprised 794 nucleotides for 295 taxa, the *glnII* dataset contained 339 taxa and 540 nucleotides, while the *gyrB* dataset included 274 taxa spanning 673 nucleotides. The *recA* alignment contained 541 nucleotides for 434 taxa, while the *rpoB* alignment comprised 303 taxa and 607 nucleotides.

All of the single gene phylogenies (Suppl. Figs. S2 to S6), together with the multigene tree (Fig. 1B) separated the strains into five clades. Based on the grouping patterns observed, nine strains were identified as conspecific to *B. arachidis* (the species belonging to the *B. japonicum* supergroup), and *B. brasilense/B. australafricanum* and *B. ivorense* (all three represent the *B. elkanii* supergroup). Strains 105Arga19, 105Argb19, 105Argb17 obtained from *A. rupestre* collected in the Mesic Highveld Grassland bioregion near Barberton as well as strain Arg105 previously isolated from *A. rupestre* in the same region (Beukes et al., 2016) formed part of the group containing *B. arachidis* CCBAU 051107^{T}. Bootstrap support (Bs) for this exclusive group in the various phylogenies was 71–100 % (Fig. 1B, Suppl. Figs. S2-S6), except in the *recA* tree (Suppl. Fig. S5), which showed only moderate support (70 % Bs). The latter group included strain ISLU9, previously isolated from *Lupinus cosentinii* (tribe Genisteae) in Spain (Sánchez-Cañizares et al., 2011). However, strain ISLU9 might not be a member of *B. arachidis* as it grouped with other undescribed strains from Genisteae in the *glnII* phylogeny (Suppl. Fig. S3), and to therefore conclusively determine its species affiliation would require more sequence data.

Strains 68Arga19, 68Argb19, 68Argb17, and the previously described strain Arg68 (Beukes et al., 2016; Avontuur et al., 2021) formed part of a clade containing the type strain (Cl-1B^{T}) of *B. ivorense*. All four strains originate from *Argyrolobium* spp. collected in the Mesic Highveld Grassland bioregion near eManzana. The *B. ivorense* clade was well supported (94–100 % Bs) across all the phylogenies, except for the *rpoB* tree (Suppl. Fig. S6) in which it had only moderate support (≥70 % Bs).

Strains 62Arga22, 62Argb22 and 62Argb07 (Suppl. Figs. S2-S6), originating from the eManzana area, formed part of a clade containing strain Arg62 previously isolated in the region (Beukes et al., 2016) and later identified as belonging to *B. brasilense* (Avontuur et al., 2021). However, the four eManzana strains grouped closely with the type strains of *B. brasilense* (UFLA 03321^{T}), *B. australafricanum* (WSM 4400^{T}) and *B. pachyrhizi* (PAC48^{T}) in an unresolved clade in the *glnII* (74 % Bs) and *rpoB* (75 % Bs) trees. Similarly, our four eManzana strains formed part of unresolved clades in the *dnaK* tree (containing only WSM 4400^{T} and UFLA 03321^{T}), the *gyrB* tree (containing PAC48^{T}, UFLA 03321^{T}, *B. elkanii* USDA76^{T}), and in the *recA* tree (containing only WSM 4400^{T}). Because of these uncertainties and the close phylogenetic relationship between the various taxa, we opted to treat the *B. brasilense/B. australafricanum* assemblage as a single taxonomic unit, and designated strains 62Arga22, 62Argb22, 62Argb07 and Arg62 as such.

The remaining two groups delineated by the five protein-coding genes included strains from the current study only and likely

represented two new species (Fig. 1B and Suppl. Figs. S2 to S6). The one group consist of six strains (Arg816^{Ts}, Arg826, Arg836, Arg846, Arg856 and Arg866) that were isolated from the nodules of *A. harveyanum* collected in the Spitzkop area in the Sub-Escarpment bioregion. This group was consistently recovered with high support (94–100 % Bs) and is an exclusive group not containing any type strains and no other strains. However, the only exception was in the *rpoB* phylogeny in which our strains grouped (< 70 % Bs) with strains zar1, cspf19 and csnc4 previously isolated from *Cytisus scoparius* (tribe Genisteae) in the United States (Horn et al., 2014) (Suppl. Fig. S6). Due to this lack of genealogical concordance, the three *C. scoparius* strains were not regarded as being part of our putative new species. Nevertheless, based on our concatenated five-gene phylogeny (Fig. 1B), its closest relatives are *B. japonicum* USDA 6^{T} and strain DSM 30140, which is listed in the NCBI database as *B. lupini* despite having been reported as a member *B. japonicum* (Peix et al., 2015).

The second putative species include three strains, with one (Arg237L^{Ts}) originating from *A. robustum* sampled in the Mesic Highveld bioregion near Mkhondo and the remaining two (31Arga and 31Argb) obtained from *A. rupestre* sampled in the Mesic Highveld Grassland bioregion near Barberton. These strains form an exclusive group (86–100 % Bs) across all of the phylogenies. According to the concatenated gene tree, the most closely related species is *B. cenepequi*, represented by the strain CNPSo 4026^{T} (Fig. 1B), which was also evident in the single-gene phylogenies (Suppl. Figs. S2-S6).

3.2. Whole genome-based support for putative new species

To further interrogate the novelty of the two putative species, a representative of each (Arg816^{Ts} and Arg237L^{Ts}) was sequenced and then subjected to whole genome-based phylogenetic analysis. The draft genomes for these strains, consisted of 9,704,017 and 9,335,704 nucleotides, respectively, with GC content of 63.2 % and 62.5 % (Suppl. Table S7), all of which are in the same range as the values for these metrics reported for other *Bradyrhizobium* genomes (Avontuur et al., 2019; Ormeno-Orrillo and Martinez-Romero, 2019). The two genome assemblies were, respectively represented by 277 and 289 contigs at an average coverage of 82.48x and 165.04x, suggesting that they are of sufficient quality for taxonomic studies (Field et al., 2008; Chun et al., 2018).

The UBCG dataset, containing 92 core gene sequences for all *Bradyrhizobium* strains with publicly available genome sequences, including representatives for unnamed species recognized in the GTDB, contained 250 taxa. The phylogeny inferred from this dataset (Suppl. Fig. S7) grouped our strain Arg816^{Ts} (representing the six-strain group) most closely with *B. japonicum* USDA 6^{T}, "*B. lupini*" DSM 30140, and *B. barranii* (strains 1S5 and 144S4^{T}). Strains Arg237L^{Ts} and 31Argb representing the second putative new species formed part of a clade containing *B. cenepequi* CNPSo 4026^{T}, and strains representing four unnamed species in the GTDB. These included strain Ai1a-2 (GTDB species *Bradyrhizobium* sp000426245) isolated from *Andira inermis* in Costa Rica (Parker, 2004), strain BRP22 (GTDB species *Bradyrhizobium* sp020194175) isolated from *Cajanus cajan* in India (Jorrin et al., 2021), as well as strains Leo170 (GTDB species *Bradyrhizobium* sp004296405) and WSM2783 (GTDB species *Bradyrhizobium elkanii* A), respectively obtained from *Leobordea lanceolata* and *Leobordea carinata* (tribe Crotalarieae) isolated from the Mesic Highveld Grassland bioregion in SA (Ardley et al., 2013; Beukes et al., 2016). The unnamed species represented by GTDB species *Bradyrhizobium elkanii* A also included strain Leo121 from *Leobordea divaricata* sampled in this region and a strain originating from *C. cajan* in India (NCBI Biosample SAMN13634622). However, strains Arg237L^{Ts} and 31Argb were most closely related to strain Leo170.

Whole genome-based sequence comparisons using JSpeciesWS yielded ANI values supporting the novelty of the two putative new species. Results for the pair-wise ANI comparisons between Arg816^{Ts} and its closest known relatives yielded values < 94 %, which is below those typically observed for members of the same species (Richter and Rosselló-Móra, 2009) (Fig. 1C). It shared 93.1–93.56 % ANI with *B. japonicum* USDA 6^{T}, *B. lupini* DSM 30140, and *B. barranii* (1S5 and 144S4^{T}), with only 64–75 % of the respective genomes being homologous. In the case of strains Arg237L^{Ts} and 31Argb, an ANI of 96.91 % was obtained, supporting the notion that they are conspecific. However, they also shared 97.07 % and 98.37 % ANI with the unnamed GTDB species represented by strain Leo170 (*Bradyrhizobium* sp004296405), with > 82 % of their genomes being homologous. By contrast, comparisons of Arg237L^{Ts}, 31Argb and Leo170 with any of the members from the same clade in the 92-gene tree (*B. cenepequi* CNPSo 4026^{T}, *Bradyrhizobium* spp. WSM2783, Leo121, BRP22, and Ai1a-2), yielded < 92 % ANI values with < 70 % of the respective genomes being homologous.

3.3. Phenotypic support for new species

Strains representing the two species shared a range of biological features (Supp. Tables S8-S11), some of which are commonly reported for rhizobia. Strains Arg816^{Ts}, Arg846, and Arg866, (representing the six-strain group) and strains Arg237L^{Ts}, 31Arga, and 31Argb (representing the three-strain group) all grew optimally at 28 °C and could routinely grow on a medium containing up to 0.5 % (w/v) NaCl. Interestingly, strain 31Arga could also grow in a medium containing 1 % (w/v) NaCl, similar to *B. cenepequi* CNPSo 4026^{T}, which, however, is not common among *Bradyrhizobium* spp. (Klepa et al., 2022). Strains Arg816^{Ts}, Arg846, and Arg866 could all grow at pH 5–10, while strains Arg237L^{Ts}, 31Arga, and 31Argb could grow at pH 7–10. All strains had a positive reaction for alkaline, catalase and oxidase tests.

Based on the API and Biolog results, none of the strains could produce indole, ferment glucose, and hydrolyse gelatin, but they could all assimilate urea, esculin ferric citrate, β-galactosidase, D-glucose, L-arabinose, 4-nitrophenyl-βd-galactopyranoside (β-galactosidase), adipic acid, D-fucose, L-fucose, fusidic acid, D-galacturonic acid, L-galactonic acid lactone, glucuronamide, L-lactic acid (Suppl. Table S11). However, many traits common to one group were variable in the other group. For example, tests using representatives of the five-strain group showed that they could reduce nitrates to nitrites and to nitrogen, and according to the Biolog results, they could grow in the presence of D-serine, D-fructose6-PO_4, D-glucuronic acid, mucic acid, D-saccharic acid, p-hydroxyphenylacetic acid, D-malic acid, L-malic acid, lithium chloride, potassium tellurite, acetoacetic acid, acetic acid, formic acid, α-D-glucose, d-mannose and sodium butyrate (but not in the presence of β-methyl-D-glucoside, D-salicin, N-acetyl-β-D-mannosamine, N-acetyl-D-galactosamine, N-acetyl neuraminic acid, myo-inositol, gelatin, glycyl-L-proline, L-arginine and pectin). For all of these traits, the three-strain group was variable, although they were consistently capable of assimilating D-mannitol, N-acetyl-glucosamine and phenylacetic acid, as well as grow in the presence of sodium lactate and methyl pyruvate (traits that were all variable in the five-strain group).

The two groups of strains further behaved differently in the nodulation tests on cowpea and siratro under glasshouse conditions. Six weeks after inoculation with representatives of the six-strain group, all the plants had root nodules with pink interiors (Fig. 1D, Suppl. Fig. S1), indicative of the formation of leghaemoglobin needed for nitrogen fixation (Moller et al., 1992). Plants inoculated with representatives from the three-strain group (Arg237L^{Ts}, 31Arga, and 31Argb) did not produce root nodules. The control plants inoculated with sterile water also did not have nodules (Fig. 1D, Suppl. Fig. S1). In all cases, the plants bearing nodules were taller and greener than their non-nodulated counterparts. Also, sequence analysis using *rpoB* confirmed that the bacteria isolated from the nodules obtained represented those initially used as inoculum (i.e., sequences of the "reisolated" strains were identical to those of the original strains).

3.4. Description of two new Bradyrhizobium species

Novelty of the two putative species identified using genealogical concordance analysis of five widely employed taxonomic marker genes (Suppl. Figs. S2-S6) was supported by a range of genomic and phenotypic data. Phylogenetic patterns and ANI values inferred from whole genome data showed that both were clearly distinct from known *Bradyrhizobium* species, and those designated as unique and yet to be described in the GTDB (except for *Bradyrhizobium* sp004296405 that is conspecific to one of our new species). Coherence within the two delineated groups was reflected by a range of biochemical and physiological traits (Suppl. Tables S8-11), with one group also being capable of nodulating both cowpea and siratro. Based on these polyphasic data, we therefore regard the two groups delineated as distinct and novel species, for which we propose the names *B. spitzkopense* sp. nov. and *B. mpumalangense* sp. nov. with the genomes of strains Arg816^{Ts} and Arg237L^{Ts} respectively designated as nomenclatural types. Tables 2 and 3 provide protologues that summarize information regarding the etymology of these names, properties of the species, and information about their nomenclatural types.

3.5. Phylogenetic analysis of nifH and nodA

To determine how the 18 strains examined in this study relate to other *Bradyrhizobium* strains in terms of their symbiotic genes, phylogenetic analyses based on the *nodA* and *nifH* genes were performed. The aligned *nodA* dataset comprised 284 haplotypes, corresponding to 214 amino acid residues, while the *nifH* dataset had 594 haplotypes spanning 242 amino acid residues. See Suppl. Tables S5 and S6, as well as the study of Beukes et al. (2016) for the NCBI accession number, strain, host, source, country, and reference information of the *nodA* and *nifH* sequences used. The taxa included in the two datasets were not the same as not all strains had sequences for both genes, with the *nifH* dataset being larger and including numerous free-living nitrogen-fixers. Note also that, although we could neither amplify nor extract these loci from the genome of *B. mpumalangense* strain Arg237L^{Ts}, the genes could be amplified and sequenced in two other strains (31Arga and 31Argb) of the species.

Phylogenetic analysis of the *nodA* dataset showed the predominance of Clade III sequences in the Sub-Escarpment and Mesic Highveld Grassland bioregions, which was represented by all isolates of *B. spitzkopense* sp. nov. (Arg816^{Ts}, Arg826, Arg836, Arg846, Arg856, and Arg866), type strain WSM 4400^{T} of *B. australafricanum* and strain Arg105a19 of *B. arachidis* (Fig. 2 and Suppl. Fig. 8). However, within the Mesic Highveld Grassland bioregion, two strains of *B. mpumalangense* sp. nov. (31Arga, 31Argb), as well as *B. ivorense* isolate Arg68a19, and *B. brasilense/B. australafricanum* strain 62Arga22 formed part of Clade XV (Fig. 2 and Suppl. Fig. S8). In addition to these strains, Clade XV included Arg68, Arg62, Leo20, Leo78, Leo79, Leo84, Leo114, Leo121, Leo117, Leo132, Leo142, Leo166, Leo170 previously isolated from *Argyrolobium* and *Leobordea* species in the region (Beukes et al., 2016).

In the less well-resolved *nifH* phylogeny (Suppl. Fig. S9), *nodA* Clade XV strains (*B. mpumalangense* sp. nov. strains 31Arga and 31Argb) grouped with other strains from southern Africa, North Africa, Asia, and Canada from diverse hosts including cowpea, *Retama sphaerocarpa*, *Vigna riukiuensis*, and *Glycine* max. The *nodA* Clade III strains of *B. spitzkopense* sp. nov. formed part of a large clade containing strains from various legumes elsewhere in Africa and the world, and they originated from hosts such as cowpea, *C. cajan, Chamaecrista pumila*, and *Arachis hypogaea*, many of which have significant agricultural importance.

3.6. Genome-based analysis of host-specific nodulation genes

Analysis of the presence/absence for genes involved in host specificity showed that the genome of *B. mpumalangense* Arg237L^{Ts} did not contain any of the common nodulation or nitrogen fixation genes. We detected homologs for only *nolBUW* and *nfeD* (Suppl. Fig. S10). By contrast, *B. mpumalangense* strain 31Argb encoded nodulation and nitrogen fixation genes: *nodABCD1D2D3IJSM, nopP1P2, nolAUB*, the 4-sulfomuconolactone hydrolase gene, and *nifABD1D2HK1K2S1S2W2*. The genome of *B. spitzkopense* strain Arg816^{Ts} had *nodABCD1D2D3D4IUSZ, noeI, noeE, nopAX, nolAUWY, nfeD,* and *nifABDHK1K2S1S2W* (Suppl. Fig. S10).

When *hsn* gene composition was compared across other southern

Table 2
Protologue description of *Bradyrhizobium spitzkopense* sp. nov.

Species name	***Bradyrhizobium spitzkopense***
Species status	sp. nov.
Specific epithet	*spitzkopense*
Species etymology	*spitz.ko.pen'se,* N.L. neut. adj *spitzkopense,* pertaining to Spitzkop, the location where the type strain was isolated
Designated Genome	Arg816^{Ts}
Genome accession number	GCF_029989365.1
SRA accession	SRR29196437
Diagnostic traits and description of novel taxon	Cells are gram-negative motile rods, 0.577–0.603 μm in length and 0.166–0.177 μm wide. Colonies are circular, opaque, pulvinate, smooth, slimy and 1–2 mm in diameter following incubation on TYA for 7 days at 28 °C. Colonies also produced an alkaline reaction in YMA with bromothymol blue as a pH indicator. Strains grow at pH 5–10, 25–30 °C (optimum 28 °C) and tolerate up to 0.5 % (w/v) NaCl. Based on the API 20NE tests, strains can reduce nitrates to nitrite and nitrogen, but cannot produce indole or ferment D-glucose, and cannot assimilate gelatin. They have urease activity and can assimilate esculin ferric citrate, 4-nitrophenyl-βD-galactoyranoside, D-glucose, L-arabinose, potassium gluconate, capric acid, adipic acid, but do not assimilate gelatin, D-mannose, D-mannitol, and N-acetyl-glucosamine. Based on the Biolog GENIII MicroPlate™ tests, strains can utilize α-D-glucose, D-mannose, D-fucose, L-fucose and D-serine, D-fructose6-PO_4, D-galacturonic acid, L-galactonic acid lactone, D-glucuronic acid, glucuronamide, mucic acid, D-saccharic acid, p-hydroxy-phenylacetic acid, L-lactic acid, α-keto-glutaric acid, D-malic acid, L-malic acid, Acetoacetic Acid, acetic acid, formic acid, and sodium butyrate. Strains can grow in the presence of lithium chloride and potassium tellurite but cannot utilize β-methyl-D-glucoside, D-salicin, N-acetyl-β-D-mannosamine, N-acetyl-D-galactosamine, N-acetyl neuraminic acid, myo-inositol, glycyl-L-proline or pectin. The GC content for the genome of strain Arg816^{T} is 63.2 %, and it encodes nodulation genes *nodABCD1, D2, D3, D4IUSZ, noeI, nopAX, nolAUWY*, and *nfeD*.
Country of origin	South Africa
Region of origin	The collection site was in the Spitzkop area in the KwaZulu-Natal province of South Africa. The collection site was in the Sub-Escarpment Grassland bioregion at an altitude of 1171 m with stony-loam, well-drained soil on the side of a gravel road.
Isolation source	Nodule on a root of *Argyrolobium harveyanum* Oliv, which is native to parts of South Africa, Zimbabwe, Eswatini, Lesotho.
Date of isolation	30-Oct-19
16S rRNA gene accession	OP375594
Genome status	Draft genome
Genome size (bp)	9704017
Size of largest contig (bp)	1,094,147
GC mol %	63.16
N50	344780
Number of contigs	277
Read coverage	82.48x
Number of strains in species cluster	Six
Culture collection number	SARCC-3537
SeqCode registry URL	https://seqco.de/i:46768

Table 3
Protologue description of *Bradyrhizobium mpumalangense* sp. nov.

Species name	***Bradyrhizobium mpumalangense***
Species status	sp. nov.
Specific epithet	*mpumalangense*
Species etymology	*mpu.ma.lan.gen'se,* N.L. neut. adj *mpumalangense,* pertaining to Mpumalanga, the region where the type strain was isolated
Designated Genome	Arg237L^{Ts}
Genome accession number	GCF_029990105.1
SRA accession	SRR29197639
Diagnostic traits and description of novel taxon	Cells are gram-negative motile rods, 0.577–0.603 μm long and 0.166–0.177 μm wide. Colonies are circular, opaque, pulvinate, smooth, slimy and 1–2 mm in diameter following incubation on TYA for 7 days at 28 °C. Colonies also produce an alkaline reaction in YMA containing bromothymol blue as pH indicator. Strains grow at 25 to 30 °C (optimum 28 °C), at pH 5–10 and tolerates up to 0.5 % (w/v) NaCl. Based on API 20NE tests, strains have urease activity, and can assimilate esculin ferric citrate, 4-nitrophenyl-βD-galactoyranoside, D-glucose, L-arabinose, D-mannitol, N-acetyl-glucosamine, adipic acid, but cannot produce indole or ferment D-glucose, and cannot assimilate gelatin, capric acid or phenylacetic acid. Based on the Biolog GENIII MicroPlate™ tests, strains can utilize D-fucose, L-fucose, D-galacturonic acid, L-galactonic acid lactone, glucuronamide, methyl pyruvate, L-lactic acid, α-keto-glutaric acid, β-hydroxy-D, L-butyric acid. They could also grow in the presence of 1 % sodium lactate but could not grow in the presence of guanidine HCl. The GC content for the genome of strain Arg237L^{Ts} is 62.5 %, and it encodes nodulation genes *nolB, nolW, nolU, nfeD* genes.
Country of origin	South Africa
Region of origin	A site owned by Mondi plc near Mkhondo in the Mpumalanga province of South Africa. The collection site was in the Mesic Highveld Grassland bioregion at an altitude of 1212 m on the side of a gravel road.
Isolation source	Nodule on a root of *A. robustum*, which is native to parts of South Africa, Botswana, and Eswatini.
Date of isolation	28-Oct-2019
16S rRNA gene accession	OP375593
Genome status	Draft genome
Genome size (bp)	9335704
Size of largest contig (bp)	240991
GC mol %	62.5
N50	91003
Number of contigs	289
Read coverage	165.04x
Number of strains in species cluster	Three
Culture collection number	SARCC-3538
SeqCode registry URL	https://seqco.de/i:46769

African *Bradyrhizobium* strains (Table 4 and Suppl. Table S12), a clear distinction became evident when comparing strains in SA-specific clades to those occurring elsewhere. Strains belonging to *nodA* Clades XIV, XIII, and XV found only in SA, mostly encoded *nodH*, while strains belonging to the pantropical *nodA* Clade III and Zimbabwean Clade IX all encoded *nodZ* (fucosyl transferase), *noeE,* and *noeI* (Table 4). The *noeE* gene was also found in Clade XV strains Arg62 and Arg68, while Clade XIV strains (Pear76^{T} and Pear77^{T}) harboured pseudogenes of *noeE* that are truncated at the 5′ end with large internal deletions.

To get a better understanding of the evolutionary origins of the detected *hsn* genes, two additional rounds of BLAST searches were performed. See Suppl. Tables S13-S16 for TBLAST search results using a local database of 870 *Bradyrhizobium* genomes from NCBI and GTDB and see Suppl. Tables S17-S20 for the BLASTN results against NCBI's nucleotide database. Overall, the *nodH* gene was identified in only fourteen *Bradyrhizobium* strains. These included strains from *nodA* SA Clades XV, XIV, and XIII. This gene was also detected in strains with *nodA* Clade IV sequences from Australia, as well as strains with *nodA* alleles forming part of the European Clades I and VII, and in other rhizobial genera (*Rhizobium, Mesorhizobium, Sinorhizobium, Methylobacterium,* and *Microvirga*). None of the strains with *nodA* Clade III alleles contained a *nodH* sequence and interestingly, this gene sequence was not detected in the genome of *B. mpumalangense* strain 31Argb. In the phylogeny inferred from the *nodH* data (Fig. 3 and Suppl. Fig. S11), SA isolates of *nodA* Clades XV, XIV, and XIII formed a separate cluster that is most closely related to *Microvirga lotononis* HAMBI 3237 isolated in Zambia. On the other hand, *Bradyrhizobium* strains with the *nodA* alleles specific for Clades I and VII grouped with *Methylobacterium nodulans* ORS 2060, while *B. cenepequi* strain CNPSo 4026^{T} (*nodA* Clade IV) grouped separately.

Phylogenetic analysis of the *noeE* dataset showed that the two SA strains of *Bradyrhizobium brasilense* (Arg62) and *B. ivorense* (Arg68) from *nodA* Clade XV are nested among strains with *nodA* Clade III (Fig. 3 and Suppl. Fig. S13), albeit on a separate lineage from other Clade III strains. The *nodA* Clade III of *B. spitzkopense* strain Arg816^{Ts} and strain SEMIA 6144 from *Arachis hypogaea* grown in Zimbabwe were intermingled with other *nodA* Clade III strains. Interestingly, *Microvirga vignae* BR3299 from cowpea in Brazil (Radl et al., 2014) appeared to have acquired its *noeE* gene from *Bradyrhizobium*.

The *nodZ* phylogeny (Fig. 3 and Suppl. Fig. S12A) displayed similar patterns to the *noeE* tree, where the strains with *nodA* Clade III alleles and the Zimbabwean Clade IX strain were intermingled with other Clade III strains. The latter included the two SA Clade XIV strains (Pear77 ^{T} and Pear76^{T}) with a truncated *nodZ*, although they appeared to cluster in a group containing SA isolates from *Chamaecrista* and *Acacia* root nodules and a Brazilian strain isolated from cowpea (Suppl. Fig. S12A). The *nodZ* of *M. vignae* BR3299 appeared to be more closely related to *Mesorhizobium* and *Rhizobium* than to strains of *Bradyrhizobium* (Suppl. Figs. S12A and B). Genome analysis in SA *Argyrolobium* strains also revealed the presence of the *noeI* gene, but only in the case of the *B. spitzkopense* Arg816^{Ts} strain, reflecting the same grouping as other phylogenies with *nodA* Clade III strains (Suppl. Fig. S14).

4. Discussion

Our findings showed that the root nodule bacteria of *Argyrolobium* species native to the Great Escarpment and the associated Grassland biome represent diverse *Bradyrhizobium* species with symbiosis genes distinct from those nodulating Genisteae in the Mediterranean and Americas. The relatively high diversity of species we identified aligns well with previous reports concerning the rhizobial symbionts of native Genisteae in other parts of the world (Stepkowski et al., 2005, 2007, 2012, 2018). Collectively, these studies show that different populations of *Bradyrhizobium* nodulate Genisteae and its sister tribe, Crotalarieae, in their respective centres of diversity (Ardley et al., 2012; Stepkowski et al., 2018; Avontuur et al., 2021). In SA, which is the primary centre of diversification for Genisteae and Crotalarieae, native members of these tribes associate with species such as *B. altum, B. oropedii, B. ivorense*, *B. brasilense* and *B. arachidis*, *B. spitzkopense* and *B. mpumalangense* (this study; Avontuur et al., 2021). In the Mediterranean, native Genisteae are nodulated by species such as *B. retamae*, *B. rifense* and *B. hipponense* (Guerrouj et al., 2013; Chahboune et al., 2012; Rejili et al., 2020), while those from the Americas associate with species like *B. japonicum* and *B. lupini* (Costa et al., 2023; Peix et al., 2015). Although limited data is available for most Crotalarieae, these legumes also seem to have unique rhizobia in other parts of the world (e.g., Stepkowski et al., 2012; Sankhla et al., 2015). In the Core Cape Subregion of SA, where the climate is more Mediterranean, many Crotalarieae genera (such as *Aspalathus, Lebeckia,* and *Rafnia*) are predominantly nodulated by *Paraburkholderia* and *Mesorhizobium* strains (Dludlu et al., 2017, 2018a; Lemaire et al., 2015a,b).

Bradyrhizobium species associated with *Argyrolobium* species native to the Great Escarpment and surrounding Grassland biome encode

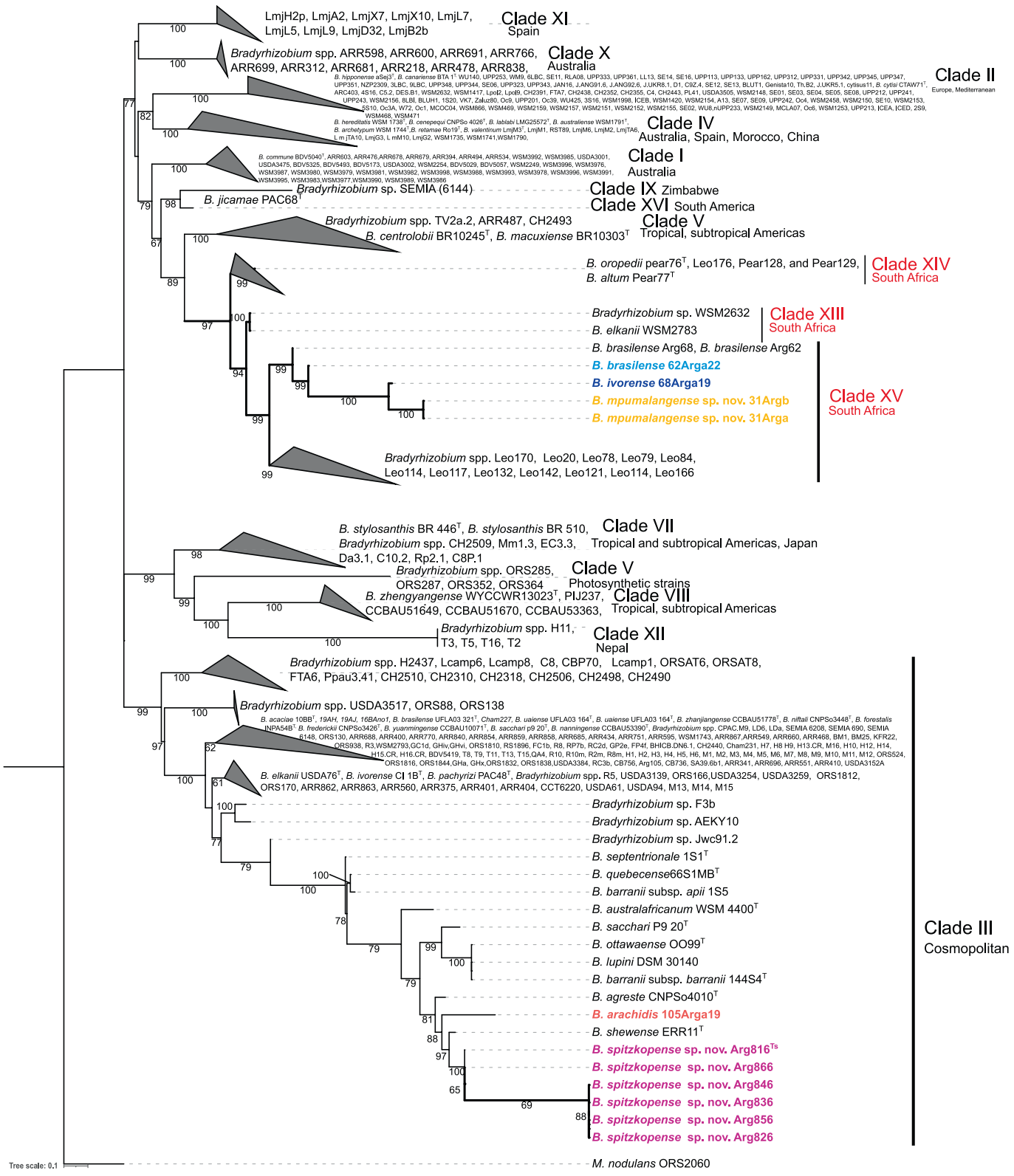

Fig. 2. A maximum likelihood phylogenetic tree inferred from the aligned amino acids for *nodA* from *Bradyrhizobium* using *Methylobacterium nodulans* ORS 2060 as outgroup. Major clades were collapsed into triangles, with strains from the current study indicated in coloured bold font as in Fig. 1B. Clade designations were based on previous studies (Beukes et al., 2016, Stepkowski et al., 2018). The dataset was expanded from those generated by Beukes et al. (2016) and Avontuur et al. (2021). For sequence accession numbers, Supplementary Table S5, as well as Table S2 of the Beukes et al. (2016) study. Bootstrap values of ≥ 60 % are indicated at the nodes, and the scale bar represents the number of nucleotide changes per site.

Table 4
Presence (+) or absence (−) of the *nodH, nodZ, noeE, noeI* in *Bradyrhizobium* strains from South Africa (SA), Botswana (BOT) and Zimbabwe with publicly available genomes.

Isolate	Species	*hsn* Genes[a]				*nodA* Clade[b]	Legume host	Geographic origin[c]	Bioregion[d]	Reference
		nodH	*nodZ*	*noeE*	*noeI*					
31Argb	*B. mpumalangense*	+	−	−	−	XV	*Argyrolobium rupestre*	Barberton, SA	Mesic Highveld Grassland	This study
Leo170	*B. mpumalangense*	+	−	−	−	XV	*Leobordea lanceolata*	Dullstroom, SA	Mesic Highveld Grassland	Beukes et al. 2016
Leo121	unknown	+	−	−	−	XV	*L. divaricata*	Dullstroom, SA	Mesic Highveld Grassland	Beukes et al. 2016
WSM2783	unknown	+	−	−	−	XIII	*L. carinata*	Mkhondo, SA	Mesic Highveld Grassland	Ardley et al. 2013
Pear77^T	*B. altum*	+	t	−	−	XIV	*Pearsonia obovata*	Dullstroom, SA	Mesic Highveld Grassland	Beukes et al. 2016; Avontuur et al. 2021
Pear76^T	*B. oropedii*	+	t	−	−	XIV	*P. obovata*	Dullstroom, SA	Mesic Highveld Grassland	Beukes et al. 2016; Avontuur et al. 2021
Arg62	*B. brasilense*	+	−	+	−	XV	*Argyrolobium* sp.	Emanzana, SA	Mesic Highveld Grassland	Beukes et al. 2016; Avontuur et al. 2021
Arg68	*B. ivorense*	+	−	+	−	XV	*Argyrolobium* sp.	Emanzana, SA	Mesic Highveld Grassland	Beukes et al. 2016; Avontuur et al. 2021
SEMIA 6144	unknown	−	+	+	+	IX	*Arachis hypogaea*	Zimbabwe	Kalahari Acacia-Baikiaea Woodlands	Delamuta et al. 2012
10BB^T	*B. acaciae*	−	+	+	+	III	*Acacia mearnsii*	Wakkerstroom, SA	Mesic Highveld Grassland	Avontuur et al. 2021
Arg816^{Ts}	*B. spitzkopense*	−	+	+	+	III	*A. harveyanum*	Spitzkop, SA	Sub-Escarpment Grassland	This study
Ghvi	unknown	−	+	+	+	III	*Vigna unguiculata*	Good Hope, BOT	Kalahari Acacia-Baikiaea Woodlands	Steenkamp et al. 2008
Gha	unknown	−	+	+	+	III	*V. unguiculata*	Good Hope, BOT	Kalahari Acacia-Baikiaea Woodlands	Steenkamp et al. 2008
Rp7b	unknown	−	+	+	+	III	*V. unguiculata*	Roodeplaat, SA	Central Bushveld	Steenkamp et al. 2008
Rc3b	unknown	−	+	+	+	III	*V. unguiculata*	Roodeplaat, SA	Central Bushveld	Steenkamp et al. 2008
Rc2d	unknown	−	+	+	+	III	*V. unguiculata*	Roodeplaat, SA	Central Bushveld	Steenkamp et al. 2008
CB756	*B. arachidis*	−	+	+	+	III	*Macrotyloma africanum*	Zimbabwe	Kalahari Acacia-Baikiaea Woodlands	Steenkamp et al. 2008
R5	*B. brasilense*	−	+	+	+	III	*V. unguiculata*	Rasesa, BOT	Kalahari Acacia-Baikiaea Woodlands	Steenkamp et al. 2008
Cham227	*B. brasilense*	−	+	+	+	III	*Chamaecrista* sp.	Pretoria, SA	Central Bushveld	Beukes et al. 2016
WSM 4400^T	*B. australafricanum*	−	+	+	+	III	*Glycine* sp.	Stutterheim, SA	Sub-Escarpment Grassland	Helene et al., 2020; Klepa et al., 2022

[a] Gene presence and absence recorded as positive or negative as observed from TBLASTN results (Table S12). The truncated *nodZ* genes of Pear77^T and Pear76^T are indicated with "t". *hsn* genes = host specificity genes.
[b] Clades inferred from *nodA* phylogeny (Fig. 2).
[c] Badplaas, Dullstroom, Barberton, Mkondo, Wakkerstroom are in the Mpumalanga (MP) province of South Africa, Spitzkop in the KwaZulu-Natal (KZN) province of South Africa, Roodeplaat and Pretoria in the Gauteng province of South Africa, Good Hope and Rasesa are in Botswana.
[d] Bioregions as defined by Trytsman et al. (2016) and Rutherford and Mucina (2006).

unique *nodA* alleles. Most of the strains examined belonged to *nodA* Clade XV, which also included strains from Crotalarieae species native to the area (Beukes et al., 2016; Avontuur et al., 2021). This *nodA* clade, together with *nodA* Clades XIV and XIII form a monophyletic assemblage in the *nodA* tree, and they have so far only been detected in strains recovered from species of Genisteae and Crotalarieae native to this geographic region (Avontuur et al., 2021; Beukes et al., 2016; this study). The fact that some of the strains investigated here formed part of the cosmopolitan *nodA* Clade III was not unexpected, as it is the most diverse *nodA* lineage of *Bradyrhizobium* (Steenkamp et al., 2008), comprising strains originating from the root nodules of various Leguminosae, spanning both native and cultivated legumes around the world (Beukes et al., 2016; Stepkowski et al., 2007; Stepkowski et al., 2012), including Sub-Saharan Africa (Degefu et al., 2017; Steenkamp et al., 2008). Nevertheless, the restriction of *nodA* Clades XV, XIV and XIII to the Great Escarpment and Grassland surroundings suggests that the symbiotic loci of these *Bradyrhizobium* species produce Nod factors specific to the unique legumes occurring in this region.

An interesting finding from the current study is that none of the *Argyrolobium* strains belonged to *nodA* Clade II, despite it being dominant among the strains associating with Genisteae in Europe, the Mediterranean, and the western parts of the Americas (Stepkowski et al., 2018). This distribution pattern suggested that Clade II coevolved with its Genisteae hosts following the tribe's initial emergence in the Mediterranean, the main centre of diversification of this tribe (Ainouche et al., 2003; Stepkowski et al., 2018). However, by having access to a larger collection of *Bradyrhizobium* strains from Genisteae native to SA, we could demonstrate that the emergence of *nodA* Clade II was not associated with diversification of Genisteae in its South African centre of origin. Instead, our data suggest that the initial origin and divergence of Genisteae and Crotalarieae in SA was linked to the emergence and diversification of a distinct version of *nodA* (i.e., the ancestor of Clades XV, XIV and XIII). It would be interesting to see whether the observed diversity patterns hold as the microsymbionts of more Genisteae and Crotalarieae are characterized in their native range.

Our results showed that both horizontal and vertical gene transfer of symbiosis loci drive the evolution and maintenance of nodulation and symbiotic nitrogen-fixation in *Bradyrhizobium* (Beukes et al., 2016; Stepkowski et al., 2018; Menna and Hungria, 2011). This was highlighted by the incongruence between the phylogenies of symbiotic genes and the core gene phylogeny, suggesting that horizontal gene transfer has played a significant role in the diversification of *Bradyrhizobium* symbionts. All strains carrying *nodA* Clade III alleles also clustered together based on *nifH* data, and such scenarios of phylogenetic congruence among different symbiotic genes are often reported for *Bradyrhizobium* (Beukes et al., 2016; Parker et al., 2002; Stepkowski

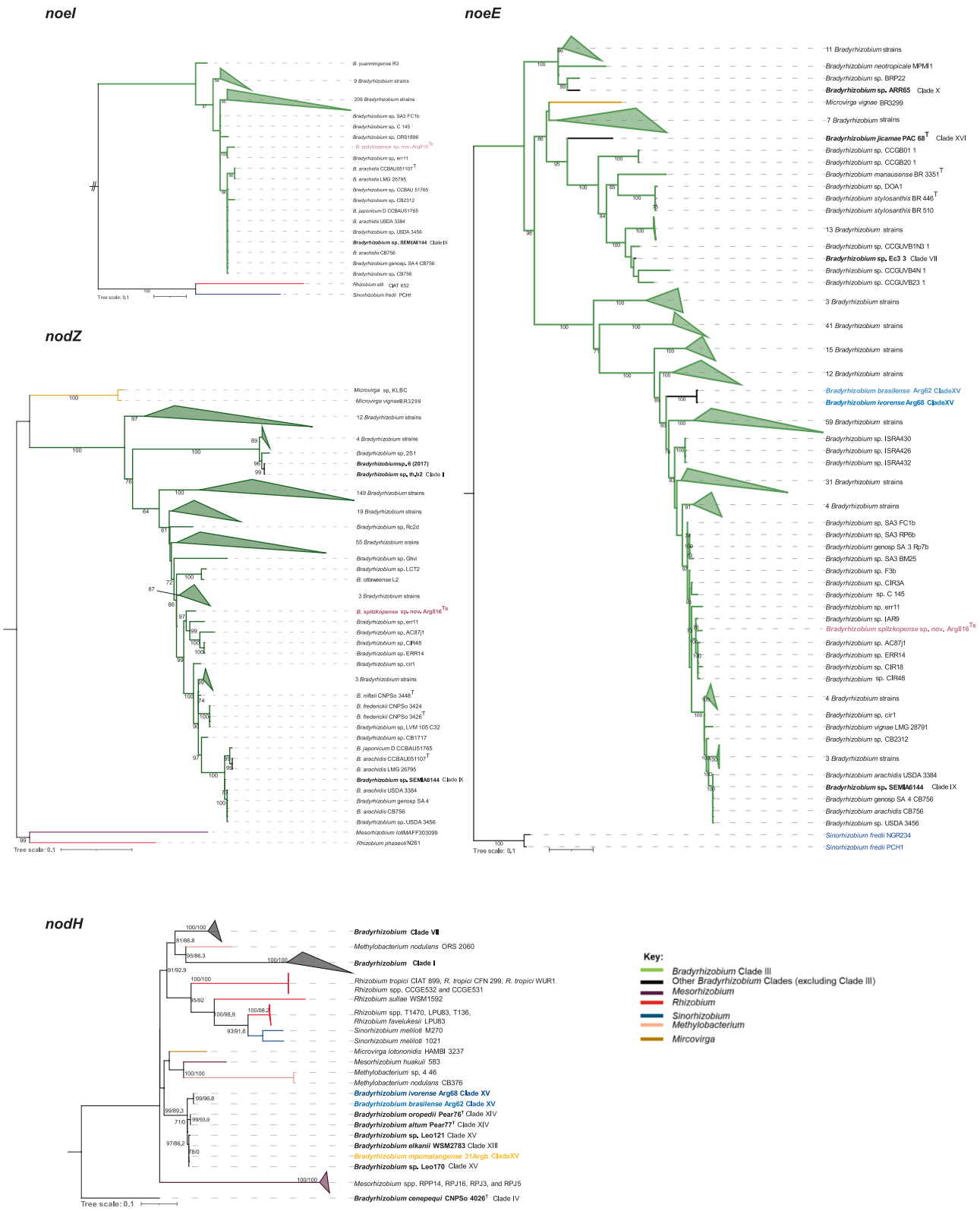

Fig. 3. Maximum likelihood phylogenies inferred from the aligned amino acid sequences for the *noeI*, *noeE*, *nodZ* and *nodH* genes using IQ-TREE. Strains from the current study are indicated in coloured bold font as in Fig. 1B, and clade designations were based on *nodA* phylogenies from this and previous studies (Beukes et al., 2016, Stepkowski et al., 2018) and indicated as in the key provided. Also, taxa from *Mesorhizobium* are shown in purple, *Rhizobium* in red, *Sinorhizobium* in navy blue, *Methylobacterium* in pink, and *Microvirga* in brown. Only UFBoot support values of $\geq$ 70 % are indicated. The scale bars indicate the number of nucleotide changes per site for each of the phylogenies. Strains from this study are coloured as in Fig. 1A. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

et al., 2005, 2007). By contrast, strains from *nodA* Clades XV, XIV and XIII grouped differently in the *nifH* phylogeny, accentuating their unique evolutionary origins involving HGT. Although such patterns of phylogenetic incongruence are less common, they have been reported for certain symbiotic genes in *Bradyrhizobium* (Okazaki et al., 2015; Okubo et al., 2016; Stepkowski et al., 2018). The involvement of HGT in evolution is also evident from the fact that our strains encoded different sets of *hsn* genes. Therefore, adaptations underpinning the coevolution of *Bradyrhizobium* with its legume host seem to be dynamic, involving nucleotide substitutions in particular genes, as well as the direct acquisition of novel genes (or their loss) to ultimately allow optimal nitrogen-fixing symbiotic interactions.

The fact that most of the *Bradyrhizobium* strains in *nodA* Clades XIII, XIV and XV also possess the *nodH* gene, suggests that these bacteria produce Nod factors that are sulfated at the terminal reducing N-acetylglucosamine end. This Nod factor modification is not commonly found in *Bradyrhizobium* but is well known in *Rhizobium* and *Sinorhizobium* where it is a host specificity signal for plants such as *Medicago sativa* and *Phaseolus vulgaris* (Laeremans et al., 1999; Wais et al., 2002; Faucher et al., 1988). The BLASTN search with the South African *nodH* sequence revealed that the closest relative of this gene was in *Microvirga* strains and other rhizobia such as *Mesorhizobium* and *Methylobacterium*, rather than other *Bradyrhizobium* strains. This, together with phylogenetic analysis suggests that these SA strains acquired their homologue of *nodH* horizontally from a source other than *Bradyrhizobium*. This observation was surprising as other *Bradyrhizobium* strains bearing *nodH* (e.g., those in *nodA* Clades I and VII) acquired it from rhizobia related to *Methylobacterium* that is known to produce sulfated Nod factors (Renier et al., 2008), while sharing common ancestry with *Rhizobium/Sinorhizobium* strains. The mosaic distribution of *nodH* in *Bradyrhizobium* may suggest that this gene was separately acquired, which, probably, was preceded by the loss of genes involved in Nod factor fucosylation and fucose modifications.

None of the *Bradyrhizobium* strains from *nodA* Clades XIII, XIV and XV encode *nodZ* and *noeI*, indicating that their Nod factors are not methylfucosylated. Also, only a few of these strains encode *noeE*, suggesting that sulfation of a fucose moiety would also not be possible. This is even though *nodZ*, *noeE* and *noeI* are commonly detected in *Bradyrhizobium* (Ormeno-Orrillo et al., 2013; Stepkowski et al., 2005, 2007, 2012), especially among broad-host range rhizobia (Ormeno-Orrillo et al., 2013; Stepkowski et al., 2007). However, the *noeI* gene in Genisteae-nodulating *Bradyrhizobium* strains is often mutated or absent, suggesting that C-2 methylated fucose is not a favoured modification recognised by their hosts (Stepkowski et al., 2018). These genes are thought to be ancestral to *Bradyrhizobium*, and they were only later acquired by other rhizobia via HGT (Ormeno-Orrillo et al., 2013; Stepkowski et al., 2007). However, the presence of a *nodZ* pseudogene in two of the *nodA* Clade XIV strains, together with a seemingly intact *noeE* gene in two *nodA* Clade XV strains, suggests that these evolutionary processes are dynamic, potentially involving many *hsn* gene losses and gains at both ancestral and contemporary time scales.

In-depth studies of the *hsn* gene repertoires of *Bradyrhizobium* species, combined with laboratory-based nodulation tests would go a long way towards understanding the host range of the unique strains identified here. For example, despite examining only a limited number of legume hosts and *hsn* genes, the absence of *nodZ* and presence of *nodH* might have impacted our nodulation results. The *nodZ* gene is known to be required for the nodulation of siratro (Stacey et al., 1994; Sanjuan et al., 1992), as its transfer to *Rhizobium leguminosarum* biovar *viciae* (a common *Vicia* symbiont) extended the bacterium's host range to other legumes such as cowpea, siratro and *Lotus* (Bras et al., 2000; López-Lara et al., 1996). Also, a mutation in *nodH* reportedly rendered some *Rhizobium* strains unable to nodulate their *M. sativa* host but conferred the ability to nodulate other legumes such as *Vigna sativa* (Debellé and Sharma, 1986; Roche et al., 1991). We hope to include testing of these strains on the original host or closely related Genisteae hosts, especially those native to the SA Grassland biome, in upcoming studies. Future work should also seek to further uncover and understand the *hsn* gene repertoires they have acquired during their coevolution with respective hosts. Such information will aid in understanding the evolutionary trends of rhizobia in this region in relation to the adaptation to their host requirements for symbiosis.

In contrast to the strains of *B. spitzkopense* (of which the strain with a genome sequence contains *nodZ*), none of the *B. mpumalangense* strains examined were able to nodulate either cowpea or siratro (one of the genome sequences for this species, 31Argb, lacked a *nodZ* sequence). This is despite the fact that we were able to amplify the *nodA* and *nifH* genes for *B. mpumalangense* strains 31Arga and 31Argb, and identify common nodulation (e.g. *nodABCD1D2D3HIJSM, nopP1P, nolAUB*) and nitrogen fixation genes, (e.g. *nifABD1D2HK1K2S1S2W2*) in the genome of *B. mpumalangense* strain 31Argb. In the genome of strain Arg237L^{Ts} we only detected the gene for nodulation efficiency and competitiveness (*nfeD*) (García-Rodríguez and Toro, 2000) and *hsn* genes *nolBUW*, suggesting the loss of the remainder of the symbiotic genes. Why *B. mpumalangense* strains 31Arga and 31Argb did not nodulate the two legumes remain unclear, but perhaps a functional *nodZ* is vital for them to nodulate cowpea and siratro.

This is the first study to describe novel *Bradyrhizobium* species from Genisteae native to SA. We propose the names *B. spitzkopense* sp. nov. (Arg816^{Ts}) and *B. mpumalangense* sp. nov. (Arg237L^{Ts}) for these two new species. Their descriptions follow the rules set out in the SeqCode, thereby enabling the description and validation of species from SA while taking into consideration legislative complexities related to the Nagoya protocol and the regulations linked to the National Environmental Management: Biodiversity Act (Hamer et al., 2021; Knight et al., 2023). Without having to deposit axenic cultures for the two organisms in international culture collections, the SeqCode allows for the validation of the names of the species, which would otherwise not have been possible under the International Code of Nomenclature of Prokaryotes (Oren et al., 2023). However, axenic cultures for the two species have been deposited in two local culture collections recognized by the World Federation of Culture Collections, making them available for use to other scientists around the world. Indeed, our results provide a sound foundation for future studies aiming to document and formally describe the vast untapped diversity of rhizobia in the Grassland biome of SA. Combined with genomics and nodulation studies, such studies would also prove invaluable for unravelling the complexities of their *hsn* gene origin.

CRediT authorship contribution statement

Mabodiba M. Maake: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Chrizelle W. Beukes:** Writing – review & editing, Validation, Supervision. **Magriet A. van der Nest:** Writing – review & editing, Supervision. **Juanita R. Avontuur:** Writing – review & editing. **Esther K. Muema:** Writing – review & editing, Methodology. **Tomasz Stępkowski:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Stephanus N. Venter:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition. **Emma T. Steenkamp:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported via the joint science and technology research collaboration programme of the National Research Foundation (NRF) of South Africa and the National Centre for Research and Development (NCBR) of Poland (Grants: POL180626349029 and PL-RPA2/06/BRADYAF/2019), as well as the by the South African Department of Science and Innovation (DSI) via its DSI-NRF Centre of Excellence in Plant Health Biotechnology (CPHB) at the Forestry and Agricultural Biotechnology Institute (FABI, University of Pretoria). We thank the NRF and Deutscher Akademischer Austauschdienst German Academic Exchange Service (DAAD) for providing Mabodiba Maake with PhD Scholarship (Grant: MND210506599713-PR-2023).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2025.108471.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics. 21, 2104–2105.
Ainouche, A., Bayer, R.J., Cubas, P., Misset, M-T., 2003. Phylogenetic relationships within tribe Genisteae (Papilionoideae) with special reference to genus *Ulex*. Adv. Legume Syst. 10, 239–252.
Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Aoki, S., Ito, M., Iwasaki, W., 2013. From β-to α-proteobacteria: the origin and evolution of rhizobial nodulation genes *nodIJ*. Mol. Biol. Evol. 30, 2494–2508.
Ardley, J.K., Parker, M.A., de Meyer, S.E., Trengove, R.D., O'Hara, G.W., Reeve, W.G., Yates, R.J., Dilworth, M.J., Willems, A., Howieson, J.G., 2012. *Microvirga lupini* sp. nov., *Microvirga lotononidis* sp. nov. and *Microvirga zambiensis* sp. nov. are alphaproteobacterial root-nodule bacteria that specifically nodulate and fix nitrogen with geographically and taxonomically separate legume hosts. Int. J. Syst. Evol. Microbiol. 62, 2579–2588.
Ardley, J.K., Reeve, W.G., O'Hara, G.W., Yates, R.J., Dilworth, M.J., Howieson, J.G., 2013. Nodule morphology, symbiontic specificity and association with unusual rhizobia are distinguisihing features of the genus *Listia* within the southern African crotalarioid clade *Lotononis* sl. Ann. Bot. 122, 1–5.
Atkinson, E.M., Palcic, M.M., Hindsgaul, O., Long, S.R., 1994. Biosynthesis of *Rhizobium meliloti* lipooligosaccharide Nod factors: NodA is required for an N-acyltransferase activity. Proc. Natl. Acad. Sci. 91, 8418–8422.
Avontuur, J.R., Palmer, M., Beukes, C.W., Chan, W.Y., Coetzee, M.P.A., Blom, J., Stepkowski, T., Kyrpides, N.C., Woyke, T., Shapiro, N., Whitman, W.B., Venter, S.N., Steenkamp, E.T., 2019. Genome-informed *Bradyrhizobium* taxonomy: where to from here? Syst. Appl. Microbiol. 42, 427–439.
Avontuur, J.R., Palmer, M., Beukes, C.W., Chan, W.Y., Tasiya, T., van Zyl, E., Coetzee, M. P., Stepkowski, T., Venter, S.N., Steenkamp, E.T., 2021. *Bradyrhizobium altum* sp. nov., *Bradyrhizobium oropedii* sp. nov. and *Bradyrhizobium acaciae* sp. nov. from South Africa show locally restricted and pantropical *nodA* phylogeographic patterns. Mol. Phylogenet. Evol. 167, 107338.
Banasiewicz, J., Granada, C.E., Lisboa, B.B., Grzesiuk, M., Matuśkiewicz, W., Bałka, M., Schlindwein, G., Vargas, L.K., Passaglia, L.M., Stępkowski, T., 2021. Diversity and phylogenetic affinities of *Bradyrhizobium* isolates from Pampa and Atlantic Forest Biomes. Syst. Appl. Microbiol. 44, 126203.
Barcellos, F.G., Menna, P., Da Silva Batista, J.S., Hungria, M., 2007. Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium* (*Ensifer*) fredii and *Bradyrhizobium elkanii* in a Brazilian Savannah soil. Appl. Environ. Microbiol. 73, 2635–2643.
Beukes, C.W., Stepkowski, T., Venter, S.N., Clapa, T., Phalane, F.L., le Roux, M.M., Steenkamp, E.T., 2016. Crotalarieae and Genisteae of the South African Great Escarpment are nodulated by novel *Bradyrhizobium* species with unique and diverse symbiotic loci. Mol. Phylogenet. Evol. 100, 206–218.
Bouznif, B., Guefrachi, I., Rodríguez De La Vega, R.C., Hungria, M., Mars, M., Alunni, B., Shykoff, J.A., 2019. Phylogeography of the *Bradyrhizobium* spp. associated with peanut, *Arachis hypogaea*: fellow travelers or new associations? Front. Microbiol. 10, 453589.
Bras, C.P., Jordá, M.A., Wijfjes, A.H., Harteveld, M., Stuurman, N., Thomas-Oates, J.E., Spaink, H.P., 2000. A *Lotus japonicus* nodulation system based on heterologous expression of the fucosyl transferase NodZ and the acetyl transferase NolL in *Rhizobium leguminosarum*. Mol. Plant Microbe Interact. 13, 475–479.
Cardoso, D., Pennington, R.T., de Queiroz, L.P., Boatwright, J.S., van Wyk, B.E., Wojciechowski, M.F., Lavin, M., 2013. Reconstructing the deep-branching relationships of the papilionoid legumes. S. Afr. J. Bot. 89, 58–75.
Chahboune, R., Carro, L., Peix, A., Ramirez-Bahena, M.H., Barrijal, S., Velazquez, E., Bedmar, E.J., 2012. *Bradyrhizobium rifense* sp. nov. isolated from effective nodules of *Cytisus villosus* grown in the Moroccan Rif. Syst. Appl. Microbiol. 35, 302–305.
Clark, V.R., Barker, N.P., Mucina, L., 2011. The Great Escarpment of southern Africa: a new frontier for biodiversity exploration. Biodivers. Conserv. 20, 2543–2561.
Cleenwerck, I., 2002. Re-examination of the genus *Acetobacter*, with descriptions of *Acetobacter cerevisiae* sp. nov. and *Acetobacter malorum* sp. nov. Int. J. Syst. Evol. Microbiol. 52, 1551–1558.
Costa, A., Panzera, M., Roldán, D.M., Viera, F., Fernández, S., Zabaleta, M., Amarelle, V., Fabiano, E., 2023. Diversity of *Bradyrhizobium* strains that nodulate *Lupinus* species native to Uruguay. Environ. Sustain. 6, 195–211.
da Silva, M., Desmeth, P., Venter, S.N., Shouche, Y., Yurkov, A., 2023. How legislations affect new taxonomic descriptions. Trends Microbiol. 31, 111–114.
Debellé, F., Sharma, S.B., 1986. Nucleotide sequence of *Rhizobium meliloti* RCR2011 genes involved in host specificity of nodulation. Nucleic Acids Res. 14, 7453–7472.
Degefu, T., Wolde-Meskel, E., Woliy, K., Frostegard, A., 2017. Phylogenetically diverse groups of *Bradyrhizobium* isolated from nodules of tree and annual legume species growing in Ethiopia. Syst. Appl. Microbiol. 40, 205–214.
Delamuta, J.R., Ribeiro, R.A., Menna, P., Bangel, E.V., Hungria, M., 2012. Multilocus sequence analysis (MLSA) of *Bradyrhizobium* strains: revealing high diversity of tropical diazotrophic symbiotic bactria. Braz. J. Microbiol. 43, 698–710.
Dludlu, M.N., Chimphango, S.B., Stirton, C.H., Muasya, A.M., 2017. Differntial preference of *Burkholderia* and *Mesorhiobium* to pH and soil types in the Core Cape Subregion, South Africa. Genes. 9, 2.
Dludlu, M., Chimphango, S., Walker, G., Stirton, C., Muasya, A., 2018. Horizontal gene transfer among rhizobia of the Core Cape Subregion of southern Africa. S. Afr. J. Bot. 118, 342–352.
Edwards, T.J., van Wyk, B.E., 2005. A synopsis of *Argyrolobium* (tribe Genisteae, Papilionoideae) in South Africa. S. Afr. J. Bot. 71, 380–417.
Faucher, C., Maillet, F., Vasse, J., Rosenberg, C., van Brussel, A., Truchet, G., Denarié, J., 1988. *Rhizobium meliloti* host range *nodH* gene determines production of an alfalfa-specific extracellular signal. J. Bacteriol. 170, 5489–5499.
Garcia, M., Dunlap, J., Loh, J., Stacey, G., 1996. Phenotypic characterization and regulation of the *nolA* gene of *Bradyrhizobium japonicum*. Mol. Plant Microbe. Interact. 9, 625–636.
García-Rodríguez, F.M., Toro, N., 2000. *Sinorhizobium meliloti nfe* (nodulation formation efficiency) genes exhibit temporal and spatial expression patterns similar to those of genes involved in symbiotic nitrogen fixation. Mol. Plant Microbe. Interact. 13, 583–591.
Gerhardt, P., Murray, R.G.E., Wood, A.W., Krieg, N.R., 1994. Methods for General and Molecular Bacteriology. American Society for Microbiology, Washington. DC.
Gronemeyer, J.L., Bunger, W., Reinhold-Hurek, B., 2017. *Bradyrhizobium namibiense* sp. nov., a symbiotic nitrogen-fixing bacterium from root nodules of *Lablab purpureus*, hyacinth bean, in Namibia. Int. J. Syst. Evol. Microbiol. 67, 4884–4891.
Guerrouj, K., Ruiz-Diez, B., Chahboune, R., Ramirez-Bahena, M.H., Abdelmoumen, H., Quinones, M.A., El Idrissi, M.M., Velazquez, E., Fernandez-Pascual, M., Bedmar, E.J., Peix, A., 2013. Definition of a novel symbiovar (sv. retamae) within *Bradyrhizobium retamae* sp. nov., *nodulating Retama sphaerocarpa* and *Retama monosperma*. Syst. Appl. Microbiol. 36, 218–223.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59 (3), 307–321.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic acids symposium series, Vol. 41, No. 41, pp. 95–98.
Hamer, M., Behr, K., Engelbrecht, I., Richards, L., 2021. Permit requirements, associated challenges and recommendations for biodiversity collections and research in South Africa. S. Afr. J. Sci. 117, 1–4.
Han, L.L., Wang, E.T., Han, T.X., Liu, J., Sui, X.H., Chen, W.F., Chen, W.X, 2009. Unique community structure and biogeography of soybean rhizobia in the saline-alkaline soils of Xinjiang, China. Plant Soil. 324, 291–305.
Hedlund, B.P., Chuvochina, M., Hugenholtz, P., Konstantinidis, K.T., Murray, A.E., Palmer, M., Parks, D.H., Probst, A.J., Reysenbach, A-L., Rodriguez-R, L.M., et al., 2022. SeqCode: a nomenclatural code for prokaryotes described from sequence data. Nat. Microbiol. 7, 1702–1708.
Held, M., Hossain, M.S., Yokota, K., Bonfante, P., Stougaard, J., Szczyglowski, K., 2010. Common and not so common symbiotic entry. Trends Plant Sci. 15, 540–545.
Helene, L.C.F., O'Hara, G., Hungria, M., 2020. Characterization of *Bradyrhizobium* strains indigenous to Western Australia and South Africa indicates remarkable genetic diversity and reveals putative new species. Syst. Appl. Microbiol. 43, 126053.
Horn, K., Parker, I.M., Malek, W., Rodríguez-Echeverría, S., Parker, M.A., 2014. Disparate origins of *Bradyrhizobium* symbionts for invasive populations of *Cytisus scoparius* (Leguminosae) in North America. FEMS Microbiol. Ecol. 89 (1), 89–98.
Howieson, J.G., Dilworth, M.J., 2016. Working with rhizobia. Centre for Rhizobium Studies Murdoch University, ACIAR.
Jorrin, B., Maluk, M., Atoliya, N., Kumar, S.C., Chalasani, D., Tkacz, A., Singh, P., Basu, A., Pullabhotla, S.V., Kumar, M., 2021. Genomic diversity of pigeon pea (*Cajanus cajan* L. Millsp.) endosymbionts in India and selection of potential strains for use as agricultural inoculants. Front. Plant Sci. 12, 680981.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods. 14, 587–589.
Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K., 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. DNA Res. 9, 189–197.
Kaneko, T., Maita, H., Hirakawa, H., Uchiike, N., Minamisawa, K., Watanabe, A., Sato, S., 2011. Complete genome sequence of the soybean symbiont *Bradyrhizobium japonicum* strain USDA6T. Genes. 2, 763–787.
Katoh, K., Rozewicki, J., Yamada, K.D., 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinformatics.

Kereszt, A., Mergaert, P., Kondorosi, E., 2011. Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims? Mol. Plant Microbe Interact. 24, 1300–1309.
Klepa, M.S., Helene, L.C., O'Hara, G, Hungria, M., 2022. *Bradyrhizobium cenepequi* sp. nov., *Bradyrhizobium semiaridum* sp. nov., *Bradyrhizobium hereditatis* sp. nov., and *Bradyrhizobium australafricanum* sp. nov. symbionts of different leguminous plants of Western Australia and South Africa and defiinition of three novel symbiovars. Int. J. Syst. Evol. Microbiol. 72, 005446.
Knight, J., Flack-Davison, E., Engelbrecht, S., Visagie, R.G., Beukes, W., Mwale, M., 2023. Compliance with the Nagoya protocol in the southern african development community. S. Afr. J. Bot. 159, 302–318.
Krueger, F., 2015. Trim Galore!: A wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data. Babraham Institute.
Kück, P., Longo, G.C., 2014. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front. Zool. 11, 1–8.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.
Laeremans, T., Snoeck, C., Mariën, J., Verreth, C., Martínez-Romero, E., Promé, J.-C., Vanderleyden, J., 1999. *Phaseolus vulgaris* recognizes *Azorhizobium caulinodans* Nod factors with a variety of chemical substituents. Mol. Plant Microbe Interact. 12, 820–824.
Lemaire, B., Dlodlo, O., Chimphango, S., Stirton, C., Schrire, B., Boatwright, S., Honnay, O., Smets, E., Sprent, J., James, E., 2015a. Symbiotic diversity, specificity and distribution of rhizobia in native legumes of the Core Cape Subregion (South Africa). FEMS Microbiol. Ecol. 91, 2–17.
Lemaire, B., Van Cauwenberghe, J., Chimphango, S., Stirton, C., Honnay, O., Smets, E., Muasya, A.M., 2015b. Recombination and horizontal transfer of nodulation and ACC deaminase (*asdS*) genes within Alpha- and Betaproteobacteria nodulating legumes of the Cape Fynbos biome. FEMS Microbiol. Ecol. 91, fiv118.
Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J.C., Dénarié, J., 1990. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature. 344, 781–784.
Letunic, I., Bork, P., 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 52, W78–W82.
Lewin, A., Rosenberg, C.H., Meyer, Z.A., Wong, C.H., Nelson, L., Manen, J.F., Stanley, J., Dowling, D.N., Denarie, J., Broughton, W.J., 1987. Multiple host-specificity loci of the broad host-range *Rhizobium* sp. NGR234 selected using the widely compatible legume *Vigna unguiculata*. Plant Mol. Biol. 8, 447–459.
Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25, 1451–1452.
López-Lara, I.M., Blok-Tip, L., Quinto, C., Garcia, M.L., Stacey, G., Bloemberg, G.V., Lamers, G.E., Lugtenberg, B.J., Thomas-Oates, J.E., Spaink, H.P., 1996. NodZ of *Bradyrhizobium* extends the nodulation host range of Rhizobium by adding a fucosyl residue to nodulation signals. Mol. Microbiol. 21, 397–408.
Martínez-Romero, E., 2009. Coevolution in *Rhizobium*-legume symbiosis? DNA Cell Biol. 28, 361–370.
Menna, P., Hungria, M., 2011. Phylogeny of nodulation and nitrogen-fixation genes in *Bradyrhizobium*: supporting evidence for the theory of monophyletic origin, and spread and maintenance by both horizontal and vertical transfer. Int. J. Syst. Evol. Microbiol. 61, 3052–3067.
Mergaert, P., van Montagu, M., Holsters, M., 1997. Molecular mechanisms of Nod factor diversity. Mol. Microbiol. 25, 811–817.
Minh, B.Q., Nguyen, M.A., Von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 39, 1188–1195.
Moller, E.M., Bahnweg, G., Sandermann, H., Geiger, H.H., 1992. A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Res. 20, 6115–6116.
Moulin, L., Béna, G., Boivin-Masson, C., Stępkowski, T., 2004. Phylogenetic analyses of symbiotic nodulation genes support vertical and lateral gene co-transfer within the *Bradyrhizobium* genus. Mol. Phylogenet. Evol. 30, 720–732.
Na, S.-I., Kim, Y.O., Yoon, S.-H., Ha, S.-M., Baek, I., Chun, J., 2018. UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. J. Microbiol. 56, 280–285.
Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.
Nurk, S., Bankevich, A., Antipov, D., Gurevich, A.A., Korobeynikov, A., Lapidus, A., Prjibelski, A.D., Pyshkin, A., Sirotkin, A., Sirotkin, Y., 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J. Comput. Biol. 20, 714–737.
Okazaki, S., Noisangiam, R., Okubo, T., Kaneko, T., Oshima, K., Hattori, M., Teamtisong, K., Songwattana, P., Tittabutr, P., Boonkerd, N., Saeki, K., 2015. Genome analysis of a novel *Bradyrhizobium* sp. DOA9 carrying a symbiotic plasmid. PloS one 10 (2), e0117392.
Okubo, T., Piromyou, P., Tittabutr, P., Teaumroong, N., Minamisawa, K., 2016. Origin and Evolution of nitrogen fixation genes on symbiosis islands and plasmid in *Bradyrhizobium*. Microbes Environ. 31, 260–267.
Oren, A., Arahal, D.R., Göker, M., Moore, E.R., Rossello-Mora, R., Sutcliffe, I.C., 2023. International code of nomenclature of prokaryotes. Prokaryotic code (2022 revision). Int. J. Syst. Evol. Microbiol. 73, 005585.
Ormeno-Orrillo, E., Martinez-Romero, E., 2019. A Genomotaxonomy view of the *Bradyrhizobium* genus. Front. Microbiol. 10, 1334.
Ormeno-Orrillo, E., Servin-Garciduenas, L.E., Imperial, J., Rey, L., Ruiz-Argueso, T., Martinez-Romero, E., 2013. Phylogenetic evidence of the transfer of *nodZ* and *nolL* genes from *Bradyrhizobium* to other rhizobia. Mol. Phylogenet. Evol. 67, 626–630.
Parker, M.A., 2004. rRNA and *dnaK* relationships of *Bradyrhizobium* sp. nodule bacteria from four papilionoid legume trees in Costa Rica. Syst. Appl. Microbiol. 27, 334–342.
Parker, M.A., 2015. The spread of *Bradyrhizobium* lineages across host legume clades: from *Abarema* to *Zygia*. Microb. Ecol. 69, 630–640.
Parker, M.A., Lafay, B., Burdon, J.J., van Berkum, P., 2002. Conflicting phylogeographic patterns in rRNA and *nifD* indicate regionally restricted gene transfer in *Bradyrhizobium*. Microbiology. 148, 2557–2565.
Parks, D.H., Chuvochina, M., Rinke, C., Mussig, A.J., Chaumeil, P.-A., Hugenholtz, P., 2022. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Res. 50, D785–D794.
Parte, A.C., Carbasse, J.S., Meier-Kolthoff, J.P., Reimer, L.C., Göker, M., 2020. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. Int. J. Syst. Evol. Microbiol. 70, 5607.
Peix, A., Ramirez-Bahena, M.H., Flores-Felix, J.D., Alonso de la Vega, P., Rivas, R., Mateos, P.F., Igual, J.M., Martinez-Molina, E., Trujillo, M.E., Velazquez, E., 2015. Revision of the taxonomic status of the species *Rhizobium lupini* and reclassification as *Bradyrhizobium lupini* comb. nov. Int. J. Syst. Evol. Microbiol. 65, 1213–1215..
Perret, X., Staehelin, C., Broughton, W.J., 2000. Molecular basis of symbiotic promiscuity. Microbiol. Mol. Biol. Rev. 64, 180–201.
Polhill, R., 1968. *Argyrolobium* Eckl. & Zeyh. (Leguminosae) in tropical Africa. Kew Bull. 22, 145–168.
Polhill, R., van Wyk, B., Lewis, G., Schrire, B., Mackinder, B., Lock, M., 2005. Genisteae. Legumes of the World. 283–297.
POWO, 2019. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet. Kew Science. http://www.plantsoftheworldonline.org/.
Radl, V., Simoes-Araujo, J.L., Leite, J., Passos, S.R., Martins, L.M.V., Xavier, G.R., Rumjanek, N.G.A., Baldani, J.I., Zilli, J.E., 2014. *Microvirga vignae* sp. nov., a root nodule symbiotic bacterium isolated from cowpea grown in semi-arid Brazil. Int. J. Syst. Evol. Microbiol. 64, 725–730.
Rahi, P., 2021. Regulating access can restrict participation in reporting new species and taxa. Nat. Microbiol. 6, 1469–1470.
Rejili, M., Off, K., Brachmann, A., Marín, M., 2020. *Bradyrhizobium hipponense* sp. nov., isolated from *Lupinus angustifolius* growing in the northern region of Tunisia. Int. J. Syst. Evol. Microbiol. 70, 5539–5550.
Renier, A., Jourand, P., Rapior, S., Poinsot, V., Sy, A., Dreyfus, B., Moulin, L., 2008. Symbiotic properties of *Methylobacterium nodulans* ORS 2060T: a classic process for an atypical symbiont. Soil Biol. Biochem. 40, 1404–1412.
Richter, M., Rosselló-Móra, R., 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. 106, 19126–19131.
Richter, M., Rosselló-Móra, R., Glöckner, F.O., Peplies, J., 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics. 32, 929–931.
Roche, P., Debellé, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Dénarié, J., Promé, J.-C., 1991. Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. Cell 67, 1131–1143.
Rodríguez-Echeverría, S., 2010. Rhizobial hitchhikers from Down Under: invasional meltdown in a plant–bacteria mutualism? J. Biogeogr. 37, 1611–1622.
Röhrig, H., Schmidt, J., Wieneke, U., Kondorosi, E., Barlier, I., Schell, J., John, M., 1994. Biosynthesis of lipooligosaccharide nodulation factors: *Rhizobium* NodA protein is involved in N-acylation of the chitooligosaccharide backbone. Proc. Natl. Acad. Sci. 91, 3122–3126.
Rutherford, M., Mucina, L., 2006. The vegetation of South Africa, Lesotho and Swaziland. South African National Biodiversity Institute, Pretoria.
Sadowsky, M.J., Cregan, P.B., Gottfert, M., Sharma, A., Gerhold, D., Rodriguez-Quinones, F., Keyser, H.H., Hennecke, H., Stacey, G., 1991. The *Bradyrhizobium japonicum nolA* gene and its involvement in the genotype-specific nodulation of soybeans. Proc. Natl. Acad. Sci. 88, 637–641.
Sánchez-Cañizares, C., Rey, L., Duran, D., Temprano, F., Sánchez-Jiménez, P., Navarro, A., Polajnar, M., Imperial, J., Ruiz-Argüeso, T., 2011. Endosymbiotic bacteria nodulating a new endemic lupine *Lupinus mariae-josephi* from alkaline soils in Eastern Spain represent a new lineage within the *Bradyrhizobium* genus. Syst. Appl. Microbiol. 34, 207–215.
Sanjuan, J., Carlson, R.W., Spaink, H.P., Bhat, U.R., Barbour, W.M., Glushka, J., Stacey, G., 1992. A 2-O-methylfucose moiety is present in the lipo-oligosaccharide nodulation signal of *Bradyrhizobium japonicum*. Proc. Natl. Acad. Sci. 89, 8789–8793.
Sankhla, I.S., Meghwal, R.R., Tak, N., Tak, A., Gehlot, H.S., 2015. Phenotypic and molecular characterization of microsymbionts associated with *Crotalaria medicagenia*: a native legume of the Indian Thar Desert. Plant Arch 15, 1003–1010.
Santi, C., Bogusz, D., Franche, C., 2013. Biological nitrogen fixation in non-legume plants. Ann. Bot. 111, 743–767.
Shang, J.Y., Zhang, P., Jia, Y.W., Lu, Y.N., Wu, Y., Ji, S., Chen, L., Wang, E.T., Chen, W. X., Sui, X.H., 2022. Coordinated regulation of symbiotic adaptation by NodD proteins and NolA in the type I peanut bradyrhizobial strain *Bradyrhizobium zhanjiangense* CCBAU51778. Microbiol. Res. 265, 127188.
Somasegaran, P., Hoben, H.J., 2012. Handbook for rhizobia: methods in legume-Rhizobium technology. Springer Science & Business Media.
Sprent, J.I., 2007. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurence of nodulation. New Phytol. 174, 11–25.
Stacey, G., Luka, S., Sanjuan, J., Banfalvi, Z., Nieuwkoop, A.J., Chun, J.Y., Forsberg, L.S., Carlson, R., 1994. *nodZ*, a unique host-specific nodulation gene, is involved in the fucosylation of the lipooligosaccharide nodulation signal of *Bradyrhizobium japonicum*. J. Bacteriol. 176, 620–633.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30, 1312–1313.

Steenkamp, E.T., Stępkowski, T., Przymusiak, A., Botha, W.J., Law, I.J., 2008. Cowpea and peanut in southern Africa are nodulated by diverse *Bradyrhizobium* strains harboring nodulation genes that belong to the large pantropical clade common in Africa. Mol. Phylogenet. Evol. 48, 1131–1144.

Stepkowski, T., Moulin, L., Krzyzanska, A., McInnes, A., Law, I.J., Howieson, J., 2005. European origin of *Bradyrhizobium* populations infecting lupins and serradella in soils of Western Australia and South Africa. Appl. Environ. Microbiol. 71, 7041–7052.

Stepkowski, T., Hughes, C.E., Law, I.J., Markiewicz, L., Gurda, D., Chlebicka, A., Moulin, L., 2007. Diversification of lupine *Bradyrhizobium* strains: evidence from nodulation gene trees. Appl. Environ. Microbiol. 73, 3254–3264.

Stepkowski, T., Watkin, E., McInnes, A., Gurda, D., Gracz, J., Steenkamp, E.T., 2012. Distinct *Bradyrhizobium* [corrected] communities nodulate legumes native to temperate and tropical monsoon Australia. Mol. Phylogenet. Evol. 63, 265–277.

Stepkowski, T., Banasiewicz, J., Granada, C.E., Andrews, M., Passaglia, L.M.P., 2018. Phylogeny and phylogeography of rhizobial symbionts nodulating legumes of the tribe Genisteae. Genes 9.

Tavaré, S., 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lect. MAth. Life Sci 17, 57–86.

Thuiller, W., Midgley, G.F., Rougeti, M., Cowling, R.M., 2006. Predicting patterns of plant species richness in megadiverse South Africa. Ecography. 29, 733–744.

Tian, C.F., Young, J.P.W., Wang, E.T., Tamimi, S.M., Chen, W.X., 2010. Population mixing of *Rhizobium leguminosarum* bv. viciae nodulating *Vicia faba*: the role of recombination and lateral gene transfer. FEMS Microbiol. Ecol. 73, 563–576.

Trytsman, M., Westfall, R.H., Breytenbach, P.J., Calitz, F.J., van Wyk, A.E., 2016. Diversity and biogeographical patterns of legumes (Leguminosae) indigenous to southern Africa. PhytoKeys. 53–96.

Urquiaga, M.C.D.O., Klepa, M.S., Somasegaran, P., Ribeiro, R.A., Delamuta, J.R.M., Hungria, M., 2019. *Bradyrhizobium frederickii* sp. nov., a nitrogen-fixing lineage isolated from nodules of the caesalpinioid species *Chamaecrista fasciculata* and characterized by tolerance to high temperature *in vitro*. Int. J. Syst. Evol. Microbiol. 69 (12), 3863–3877.

van Lill, M., Venter, S.N., Muema, E.K., Palmer, M., Chan, W.Y., Beukes, C.W., Steenkamp, E.T., 2024. SeqCode facilitates naming of South African rhizobia left in limbo. Syst. Appl. Microbiol. 47, 126504.

Venter, S.N., Palmer, M., Beukes, C.W., Chan, W.Y., Shin, G., van Zyl, E., Seale, T., Coutinho, T.A., Steenkamp, E.T., 2017. Practically delineating bacterial species with genealogical concordance. Antonie Van Leeuwenhoek. 110, 1311–1325.

Vinuesa, P., Silva, C., Lorite, M.J., Izaguirre-Mayoral, M.L., Bedmar, E.J., Martínez-Romero, E., 2005. Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from rrs, *atpD*, *recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetland. Syst. Appl. Microbiol. 28, 702–716.

Vinuesa, P., Silva, C., Werner, D., Martínez-Romero, E., 2005. Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in *Bradyrhizobium* species cohesion and delineation. Mol. Phylogenet. Evol. 34, 29–54.

Vinuesa, P., Rojas-Jiménez, K., Contreras-Moreira, B., Mahna, S.K., Prasad, B.N., Moe, H., Selvaraju, S.B., Thierfelder, H., Werner, D., 2008. Multilocus sequence analysis for assessment of the biogeography and evolutionary genetics of four *Bradyrhizobium* species that nodulate soybeans on the Asiatic continent. Appl. Environ. Microbiol. 74, 6987–6996.

Wais, R.J., Keating, D.H., Long, S.R., 2002. Structure-function analysis of nod factor-induced root hair calcium spiking in Rhizobium-legume symbiosis. Plant Physiol. 129, 211–224.

Zilli, J.E., Barauna, A.C., da Silva, K., de Meyer, S.E., Farias, E.N., Kaminski, P.E., da Costa, I.B., Ardley, J.K., Willems, A., Camacho, N.N., Dourado Fdos, S., O'Hara, G., 2014. *Bradyrhizobium neotropicale* sp. nov., isolated from effective nodules of *Centrolobium paraense*. Int. J. Syst. Evol. Microbiol. 64, 3950–3957.