Systematic and Applied Microbiology 47 (2024) 126504

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

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

SeqCode facilitates naming of South African rhizobia left in limbo

Melandré van Lill [a], Stephanus N. Venter [a], Esther K. Muema [b], Marike Palmer [c], Wai Y. Chan [a], Chrizelle W. Beukes [d], Emma T. Steenkamp [a,*]

[a] Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa
[b] Department of Soil Science, Faculty of AgriSciences, Stellenbosch University, South Africa
[c] Department of Microbiology, University of Manitoba, Winnipeg, MB, Canada
[d] The James Hutton Institute, Invergowrie, Dundee DD2 5DA, United Kingdom

ARTICLE INFO

Keywords:
Biodiversity
ICNP
Phylogenomics
Mesorhizobium
Vachellia karroo

ABSTRACT

South Africa is well-known for the diversity of its legumes and their nitrogen-fixing bacterial symbionts. However, in contrast to their plant partners, remarkably few of these microbes (collectively referred to as rhizobia) from South Africa have been characterised and formally described. This is because the rules of the International Code of Nomenclature of Prokaryotes (ICNP) are at odds with South Africa's National Environmental Management: Biodiversity Act and its associated regulations. The ICNP requires that a culture of the proposed type strain for a novel bacterial species be deposited in two international culture collections and be made available upon request without restrictions, which is not possible under South Africa's current national regulations. Here, we describe seven new *Mesorhizobium* species obtained from root nodules of *Vachellia karroo*, an iconic tree legume distributed across various biomes in southern Africa. For this purpose, 18 rhizobial isolates were delineated into putative species using genealogical concordance, after which their plausibility was explored with phenotypic characters and average genome relatedness. For naming these new species, we employed the rules of the recently published Code of Nomenclature of Prokaryotes described from Sequence Data (SeqCode), which utilizes genome sequences as nomenclatural types. The work presented in this study thus provides an illustrative example of how the SeqCode allows for a standardised approach for naming cultivated organisms for which the deposition of a type strain in international culture collections is currently problematic.

Introduction

South Africa is known for the astounding diversity of microbes inhabiting its terrestrial ecosystems (Cowan et al., 2013; Guo and Arnolds, 2018). This is particularly notable in soil environments where myriads of microorganisms mediate soil fertility through nutrient cycling (Cowan et al., 2022). For example, research during the last decade demonstrated that South African soils are home to a large variety of novel rhizobial bacteria (e.g., Beukes et al., 2013, 2019a; Dludlu et al., 2018; Hassen et al., 2020; Lemaire et al., 2015; Ndlovu et al., 2013). These oligotrophic soil inhabitants convert chemically inert molecular nitrogen to bioavailable ammonia when they engage in symbioses with plants from the family Leguminosae (Poole et al., 2018; Young et al., 2001). The reaction is catalysed by rhizobial nitrogenase within specialized organs known as nodules occurring on the roots or stem of the legume (Poole et al., 2018). Despite the ecological and agricultural importance of this symbiosis (Davies-Barnard and Friedlingstein, 2020; Soumare et al., 2020), only a few rhizobial species from South Africa have been described and named (Avontuur et al., 2022; Beukes et al., 2019b; Claassens et al., 2023; Mavima et al., 2021; Steenkamp et al., 2015).

There are many reasons for the slow pace at which new taxa are described and documented (Godfray et al., 2004; Pallen et al., 2021), but for signatories of the Convention on Biological Diversity (CBD) and the Nagoya Protocol (NP), national biodiversity laws and associated regulations present additional hurdles (Rahi, 2021). The CBD promotes protection of biological diversity through fair and equitable sharing of biological resources (David Cooper and Noonan-Mooney, 2013), while the NP provides the legal framework for benefit-sharing (Ramos, 2021). Although South Africa supports both these agreements through its National Environmental Management: Biodiversity Act (NEMBA), implementation of relevant legislation, policies and procedures are in some cases lagging (Da Silva et al., 2023; Knight et al., 2023). The latter is particularly evident in the exceedingly complex nature of current

* Corresponding author.
E-mail addresses: melandre.vanlill@up.ac.za (M. van Lill), emma.steenkamp@fabi.up.ac.za (E.T. Steenkamp).

https://doi.org/10.1016/j.syapm.2024.126504
Received 29 January 2024; Received in revised form 13 March 2024; Accepted 18 March 2024
Available online 25 March 2024

permitting requirements for distributing indigenous biological resources inside and outside the country (Hamer et al., 2021). Taxonomic enterprises are thus at odds with the International Code of Nomenclature of Prokaryotes (ICNP) which stipulates in Rule 30 that only axenic cultures deposited in two international culture collections and made available without any restrictions are accepted as appropriate nomenclatural types (Oren et al., 2023). Under the ICNP, new bacterial names are thus essentially disqualified from validation as long as the distribution of type strains are restricted (Da Silva et al., 2023; Lewis, 2012; Oren et al., 2023; Ramos, 2021).

In 2022, an alternative system for naming prokaryotes was introduced to address the fact that most prokaryotes are precluded from being validly named under the ICNP due to lack of axenic cultures (Hedlund et al., 2022; Murray et al., 2020). It was formally published as the Code of Nomenclature of Prokaryotes Described from Sequence Data (SeqCode) and broadly uses the same naming rules as the ICNP but designates genome sequences as nomenclatural types (Hedlund et al., 2022). In other words, genome sequences of uncultivated and slow-growing species, as well as species from countries that restrict the export of biological material, are all legitimate sources of nomenclatural types under the SeqCode (Hedlund et al., 2022; Palmer et al., 2022). Therefore, given that genome data form an essential part of rhizobial taxonomy (de Lajudie et al., 2019; Ormeño-Orrillo et al., 2015), the SeqCode provides a fitting avenue for documenting and naming new taxa from countries such as South Africa.

The overall goal of this study was to describe and name new rhizobial species associated with *Vachellia karroo* using the SeqCode. This legume is one of the most widely distributed tree species in southern Africa, occurring throughout South Africa (Dingaan and du Preez, 2017) where it thrives across a range of climatic and edaphic conditions (Maroyi, 2017). Because of its high adaptability to different environments, *V. karroo* is a pioneer species in most biomes, as well as a woody invader in savannas (Taylor and Barker, 2012). In fact, the legume's ecological success has been linked to the large diversity of its rhizobial symbionts (Beukes et al., 2019a). Among these, *Mesorhizobium* seems to be a predominant genus, as it has been associated with *V. karroo* from across South Africa including sites in the Succulent Karoo, Desert, Grassland and Savanna biomes (Beukes et al., 2019a; Muema et al., 2022). Although *Mesorhizobium* strains are commonly associated with South African legumes (Beukes et al., 2019a; Gerding et al., 2012; Hassen et al., 2020), none so far have been formally described and named.

The specific aims of our study were to taxonomically delineate a set of *Mesorhizobium* strains previously isolated from the root nodules of *V. karroo* and to describe and name novel species according to the SeqCode's requirements (Hedlund et al., 2022), while also adhering to the minimal standards for defining new rhizobial species (de Lajudie et al., 2019). Apart from advancing the field of *Mesorhizobium* systematics, the results of this study would illustrate how the SeqCode could facilitate valid naming of species for which pure cultures are not available beyond the country of origin's borders. The SeqCode will undoubtedly be invaluable for documenting and naming new bacteria in countries striving to protect their biodiversity through highly restrictive access and benefit sharing legislation.

Materials and methods

Bacterial strains and growth conditions

A collection of 18 *Mesorhizobium* strains obtained from the root nodules of *V. karroo* was utilized in this study (Fig. 1). The strains were recovered with "trapping" experiments under controlled greenhouse conditions during which the legume was grown in soils collected from the rhizosphere of *V. karroo* occurring in diverse locations in the Limpopo, Free State, North West, Eastern Cape, Western Cape and Northern Cape Provinces of South Africa. Of these, 17 strains (VK1D, VK4B, VK25D, VK25A, VK3E, VK9D, VK2D, VK2B, VK24D, VK23B, VK23D, VK23A, VK4C, VK22E, VK22B, VK23E, and VK3C) were available from a previous study (Beukes et al., 2019a) and the remaining strain (MSK 1335) was recovered by Muema et al. (2022).

For routine culturing, Yeast Mannitol (YM; VWR Chemicals, Leuven, Belgium) agar (Oxoid, Hampshire, England) and incubation at 28 °C were used. In instances where strains excessively produced exopolysaccharides, YM was replaced by Tryptone Yeast (TY; Oxoid) agar supplemented with 44 g/l $CaCl_2{\cdot}2H_2O$ (Howieson and Dilworth, 2016). All

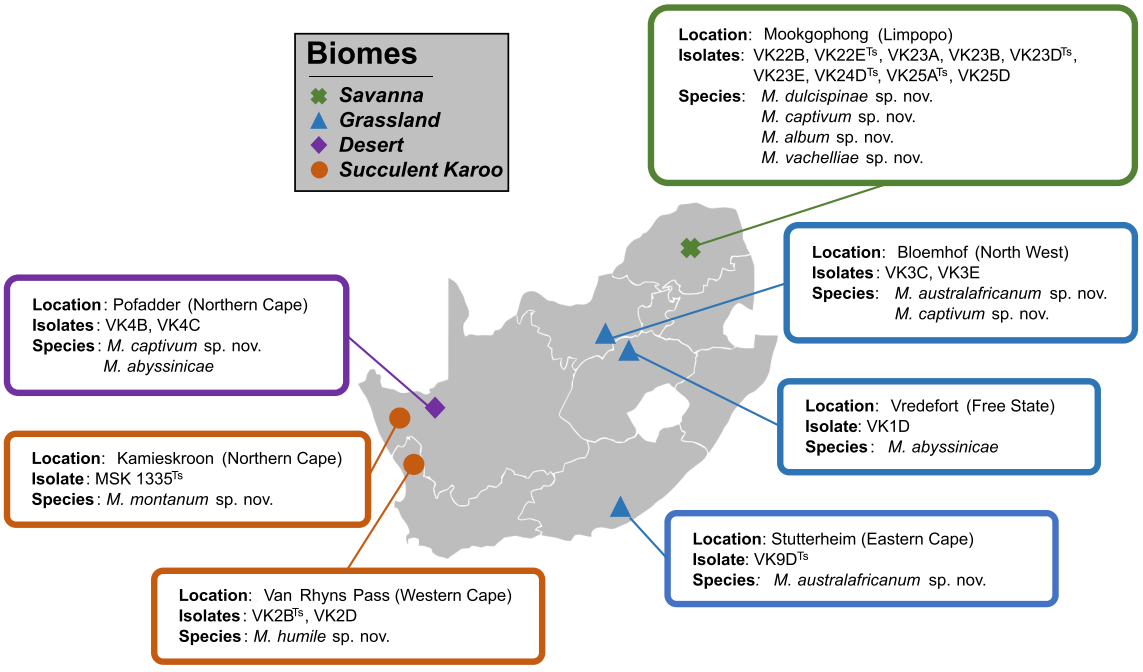

Fig. 1. The geographic sampling locations of the rhizosphere soil used as a trapping medium to obtain the 18 *Mesorhizobium* strains associated with *V. karroo* across South Africa. The biomes were defined according to Driver *et al.*, 2012 and are represented by different symbols. Additionally, this figure also depicts the isolates identified following the trapping experiments with the sampled soil and their respective species allocation.

strains were cryopreserved at −80 °C in 20 % (v/v) glycerol and Microbank™ storage beads (Pro-lab Diagnostics, Canada).

Confirmation of Mesorhizobium membership

To confirm that the bacteria represent *Mesorhizobium*, 16S ribosomal RNA (rRNA) sequence analysis were conducted. For this purpose, DNA extracted with the Quick-DNA™ MiniPrep kit (Zymo Research, United Sates of America) was used as template in PCR with primers 27F (5′ AGA GTT TGA TCC TGG CTC AG 3′) and 1492R (5′ GGT TAC CTT GTT ACG ACT T 3′) to amplify the 16S rRNA gene (Suzuki and Giovannoni, 1996).

The PCR products were purified enzymatically by incubation at 37 °C for 15 min in the presence of 20 U/µl Exonuclease I (Thermo Fisher Scientific, Waltham, United States of America) and 2 U/µl FastAP thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific). Following inactivation of the enzymes at 80 °C for 15 min, amplicons were sequenced using the BigDye™ Terminator v3.1 Cycle Sequencing kit (Thermo Fisher Scientific), an AB13100 Automated Capillary DNA sequencer (Thermo Fisher Scientific) and the respective PCR primers. The sequence files obtained were then viewed and edited using ChromasLite version 2.01 (Technelysium, Queenswood, Australia) and BioEdit version 7.05 (Hall, 1999).

The sequences for all 18 strains were subjected to the Basic Local Alignment Search Tool (BLAST) analysis (Altschul et al., 1990) against the nonredundant nucleotide database of the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov). The 16S rRNA sequences were additionally subjected to phylogenetic analysis. For this purpose, a 16S rRNA dataset was assembled using published sequences for the type strains of *Mesorhizobium* according to the List of Prokaryotic Names with Standing in Nomenclature (LPSN; Parte et al., 2020). The 16S rRNA sequences were aligned with MAFFT (Multiple Alignment using Fast Fourier Transform) version 7.310 using the Q-INS-I strategy that takes secondary structure into account (Katoh et al., 2019). The alignment was then used to construct a maximum likelihood (ML) phylogeny in IQ-TREE release 2.2.2.6 (Nguyen et al., 2015; Trifinopoulos et al., 2016). Branch support was estimated by performing Ultrafast Bootstrap approximation (UFBoot; Hoang et al., 2018) of 1000 pseudoreplicate alignments and single branch tests of 1000 replicates per branch [SH-aLRT (Guindon et al., 2010)]. The phylogenetic tree was viewed on the Interactive Tree of Life version 5 (iTOL; https://itol.embl.de/; Letunic and Bork, 2021) and edited using Inkscape version 1.3 (https://inkscape.org/).

Genome sequencing and assembly

High-quality DNA of the 18 *Mesorhizobium* strains used in this study was extracted from five-day old YM or TY cultures using either the Promega Wizard Genomic DNA purification kit (Anatech Limited, South Africa) or the hexadecyltrimethylammonium bromide (CTAB) method (Cleenwerck et al., 2002; William et al., 2012). Genomes were sequenced by the Biotechnology Platform of the Agricultural Research Council (ARC-BTP), South Africa, using the HiSeq 3000 sequencing platform (Illumina, San Diego, United States of America) and the NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina® (New England Biolabs, Ipswich, United States of America).

Quality checks and read trimming were performed using FastQC version 0.11.9 and Trim Galore version 0.6.7, respectively (Babraham Bioinformatics, Cambridge). Draft genome assemblies were constructed using SPAdes version 3.13.1 (Bankevich et al., 2012) and assembly statistics were estimated with QUAST version 5.0.2 (Center for Algorithmic Biotechnology, United States of America). Average sequencing coverage was calculated using BBMap version 38.18 (Department of Energy, Joint Genome Institute, United States of America). The quality of the draft genomes was further interrogated using the Microbial Genomes Atlas (MiGA) webserver (https://disc-genomics.uibk.ac.at/miga//; Rodriguez-R et al., 2018) and CheckM version 1.0.18 (Parks et al., 2015). To estimate the position of our 18 strains in the taxonomic framework implemented in the Genome Taxonomy Database (GTDB; https://gtdb.ecogenomic.org/; Parks et al., 2022), we utilised the taxonomic classification toolkit, GTDB-Tk version 2.3.2 (Chaumeil et al., 2022), as implemented in KBase (the workflow is available at https://kbase.us/n/159176/19/; Arkin et al., 2018). The draft genome assemblies and raw sequence data were deposited into the NCBI database under the BioProject accession number: PRJNA1009259.

For the strains designated as nomenclatural types (see below), Barrnap version 0.9 (https://github.com/tseemann/barrnap) was used to predict and extract the ribosomal rRNA sequences (i.e., 16S, 5S and 23S). Transfer RNAs and the standard amino acids that they decode were predicted using tRNAscan-SE version 2.0 (Chan et al., 2021; Chan and Lowe, 2019). The genome coordinates of the common nodulation gene *nodA*, N-acetylglucosaminyltransferase gene *nodC* and nitrogen-fixing gene *nifH* were predicted for the nomenclatural type sequences *in silico* using Prodigal version 2.6.3 (Hyatt et al., 2010) and searched againts other known sequences using BLAST + version 2.11.0 (Camacho et al., 2009).

Species delineation and phylogenomic analysis

For delineating the strains to the species level, species hypotheses were based on genealogical concordance. For inferring phylogenetic relationships among *Mesorhizobium* species, the Up-to-date Bacterial Core Gene set (UBCG version 3.0) was used (Na et al., 2018). The 92 UBCG sequences were predicted and extracted using Prodigal and the hmmsearch 3.1b2 function of HMMER3 (Sun and Buhler, 2007) as part of the UBCG pipeline (Na et al., 2018). The 92 gene dataset included information for all genomes sequenced in this study, as well as those available for type strains of *Mesorhizobium sensu stricto* as demarcated in Release 08-RS214 (28th April 2023) of the GTDB. The dataset also included gene sequences for the representatives of undescribed novel species catalogued by the GTDB (Parks et al., 2022). Information regarding the entire dataset can be found in Supplementary Table S1.

The extracted gene sequences were aligned with MAFFT using default settings (Katoh et al., 2019), and then partitioned and concatenated using FASconCAT-G version 1.02 (Kück and Longo, 2014). The returned dataset was subjected to ML phylogenetic analysis in IQ-TREE, using independent parameter estimation for each gene partition and the ModelFinder tool for determining optimal model parameters (Kalyaanamoorthy et al., 2017). Branch support was estimated with UFBoot and SH-aLRT of 1000 pseudoreplicates.

Using the UBCG pipeline, we evaluated the phylogenetic clustering for our 18 isolates in each of the 92 genes using FastTree (Price et al., 2010) and the Gene Support Index. The latter reflects the number of single-gene phylogenies supporting a particular bifurcation of the 92-gene UBCG phylogeny (Na et al., 2018). These values were used to identify concordant groups (i.e., clusters supported by all 92 single gene trees), representing putative species (Venter et al., 2017).

In cases where whole genome sequences were not available for relevant *Mesorhizobium* type strains, a smaller dataset was constructed, consisting of the five protein-coding genes typically used for multilocus sequences analysis (MLSA) in this genus (Laranjo et al., 2014). These genes are *recA* (encoding recombinase protein A), *atpD* (encoding ATP synthase beta subunit), *glnII* (encoding glutamine synthetase II), *rpoB* (encoding DNA-directed RNA polymerase subunit beta) and *gyrB* (encoding DNA gyrase subunit B), of which only *recA* and *rpoB* are included in the UBCG set (Na et al., 2018). The single gene data, as well as the partitioned concatenated data were subjected to ML phylogenetic analysis, with branch support also estimated using UFBoot, all as described above.

Overall-genome-relatedness index analysis

Average nucleotide identity (ANI) values were used as a measure of

the similarity among genomes and are routinely included in novel rhizobial species descriptions (Goris et al., 2007; Ormeño-Orrillo et al., 2015; Yoon et al., 2017). Pair-wise ANIb values among genomes were estimated using JSpecies version 1.2.1 (Richter and Rosselló-Móra, 2009) and calculated by using BLAST + 2.11.0 (Camacho et al., 2009). This software performs reciprocal BLAST searches between genome pairs in which 1020 nucleotide sections of the query genome are compared against the reference genome. From these data, JSpecies estimates normalised average identity values for homologous regions showing ≥ 30 % similarity over ≥ 70 % of their alignment length.

Phenotypic tests

All strains were grown on TY agar for five days and a single colony was used to perform Gram staining (Vincent, 1970). Gram-stained cells were viewed at 100x magnification using a Zeiss Axioskop 2 plus microscope with Zen Blue software and the general cell shape was determined. The motility of the bacteria was determined by inoculating the *Mesorhizobium* strains into a semisolid growth medium containing 0.4 % (w/v) YM agar (Gao et al., 2020).

Growth of the strains on TY medium was evaluated by incubation for five days at 4 °C, 10 °C, 15 °C, 20 °C, 25 °C, 28 °C, 30 °C, 35 °C, 37 °C and 40 °C. Other growth characteristics were determined using incubation for five days at 28 °C. These included the pH growth range, which was determined by growing a bacterial suspension in YM broth with adjusted pH 4 to pH 10 by using hydrogen chloride and sodium hydroxide and incubated in an orbital shaking incubator (MRC, Harlow, United Kingdom) at *ca.* 150 rotations per minute. Additionally, sodium chloride (NaCl) tolerance was tested on YM agar plates supplemented with 0.3 %, 0.5 %, 1.0 %, 1.5 %, 2.0 % and 2.5 % (w/v) of NaCl, respectively.

Biochemical attributes of the strains were determined using API® 20NE strips (bioMérieux, France), according to the manufacturer's instructions with modifications made to the incubation period and temperature (i.e., five days at 28 °C). The strains were also analysed using the Biolog GENIII MicroPlate™ which includes 71 different carbon source utilization assays and 23 chemical sensitivity assays (Biolog, France), with incubation time and temperature identical to that of the API® 20NE strips.

Nodulation confirmation and symbiovar identification

The nodulation capabilities of representative strains of each proposed novel species were tested on *V. karroo* as described previously (Beukes et al., 2019a). Briefly, seeds of *V. karroo* were pretreated with concentrated sulphuric acid, rinsed with sterile distilled water (sH_2O) and then soaked in sH_2O for three hours. The imbibed seeds were then placed on water agar (1.5 %; w/v) and incubated in the dark for two days at 28 °C to germinate. Three germinated seeds were then planted in a plastic nursery pot filled to 80 % of its capacity with sterile river sand, after which inoculum for a particular strain was applied directly onto the radicle of each seedling. For each isolate, the inoculum consisted of a YM culture following incubation at 28 °C for 3 days in a shaking incubator (MRC, for *ca.* 150 rotations per minute). For the negative control, uninoculated seedlings were used. Plants were left to grow under nitrogen-free conditions (Howieson and Dilworth, 2016) in a greenhouse set at a cycle of 28 °C for 14 h of daytime and 15 °C for 10 h of nighttime. Plants were watered three times a week with sH_2O until the sand was saturated, and treated with nitrogen-free Hoagland plant growth solution once a week (Howieson and Dilworth, 2016).

After seven weeks of growth, the plants were harvested and examined for nodule development. If present, nodules were excised from the roots and surface sterilised by soaking them for 2 min in 3.5 % (v/v) sodium hypochlorite (NaOCl) and then for 1 min in 80 % (v/v) ethanol. Following five rinses with sH_2O, the content of individual nodules was streaked onto YM agar and incubated at 28 °C (Howieson and Dilworth, 2016). Subculturing on YM agar was performed to obtain pure cultures.

The identities of the root nodule bacteria were verified by comparison to the original strains through sequencing of the *atpD* gene region (in the case of MSK 1335) and the *rpoB* gene region (all other strains). For this purpose, DNA was extracted using the Quick-DNA™ MiniPrep kit (Zymo Research, United States of America) and used as templates in PCR with primers atpD-273f and atpD-771r for *atpD* (Gaunt et al., 2001) or ropB82F and ropB1580R for *rpoB* (Capela et al., 2001). The constituents of the amplification reaction mixtures and the PCR cycling conditions used were the same as those used for the 16S rRNA region above. The only differences were that *atpD* employed an annealing temperature of 56 °C for 40 s, while the *rpoB* cycling conditions included an initial denaturation at 95 °C for 5 min, followed by 3 cycles of denaturation at 94 °C for 2 min, annealing at 58 °C for 2 min and elongation at 72 °C for 1 min, after which another 30 cycles at the same temperatures of denaturisation for 30 s, annealing for 1 min and extension for 1 min and a final elongation step was performed at 72 °C for 5 min. The *rpoB* and *atpD* amplicons were purified and sequenced as before. Their sequences were then compared to those of the original isolates used as inoculum, in order to prove Koch's postulates and confirm the rhizobial abilities/lifestyle of these strains.

Phylogenetic analysis to investigate the relatedness of our strains to known symbiovars (symbiotic variants) was conducted using the *nodC* genes extracted from whole genome sequences generated in the current study, along with additional *nodC* sequences with high sequence similarity based on a BLAST search against NCBI's nucleotide database (accessed 4 March 2024) and other *Mesorhizobium* type strains according to the LPSN (accessed 4 March 2024). The *nodC* sequences were aligned with MAFFT using default settings (Katoh et al., 2019). Maximum likelihood phylogenetic analysis was performed using IQ-TREE with branch support estimated with UFBoot and SH-aLRT of 1000 pseudoreplicates each.

Results

Mesorhizobium strains from different biomes in South Africa

Analysis of the 16S rRNA sequences confirmed that all 18 of the strains included in this study are members of *Mesorhizobium sensu stricto*, as defined in the GTDB's most recent release (Parks et al., 2022). The nucleotide BLAST results showed that they all shared ≥ 96 % sequence similarity with other *Mesorhizobium* strains (Camacho et al., 2009). The 16S rRNA phylogeny separated the 18 *Mesorhizobium* strains into different clusters, but with limited resolution and statistical support (Supplementary Figure S1). Nevertheless, these data confirmed that the 18 strains isolated from the root nodules of *V. karroo* represent diverse *Mesorhizobium* species from the Succulent Karoo (n = 3), Desert (n = 2), Grassland (n = 4) and Savanna (n = 9) biomes in South Africa (Fig. 1). We accordingly proceeded to sequence the genomes of these bacteria for further analyses.

High-quality genome assemblies for 18 Mesorhizobium strains

Draft genomes were sequenced and assembled for all of the *Mesorhizobium* strains isolated from the root nodules of *V. karroo* studied here (Supplementary Table S2). Overall, the genomes were between 6 374 021 and 7 496 434 base pairs (bp) in size, with %GC mole content ranging from 62.94 % to 64.03 %, which are comparable to other *Mesorhizobium* genomes (Geddes et al., 2020). In most instances, our assemblies also contained fewer than 100 contigs. The assemblies all had an average read coverage of ≥ 74.362x, which is well above the recommended average coverage for accurate whole genome comparisons (Field et al., 2008) and for use in taxonomic analyses (Chun et al., 2018). The high quality and integrity of our genome assemblies were also reflected by high N50 and N75 values. These parameters represent the contig lengths where 50 % and 75 %, respectively, of the genome is contained in the shortest contig of that length. All the recorded N50

values exceeded 105 172 bp, and the N75 values all exceeded 55 546 bp. Similarly, the L50 and L75 values reflect the number of contigs needed to cover 50 % and 75 % of the genome, respectively, and for our assemblies, these were respectively 5–19 contigs and 10–40 contigs. The sequence data and draft assemblies generated for all 18 *Mesorhizobium* strains have been deposited in the NCBI's sequence read archive (SRA) and genome database, respectively (see Supplementary Table S2 accession numbers).

Robust 92-gene phylogeny for Mesorhizobium sensu stricto

Analysis of the combined nucleotide data for 92 UBCGs allowed inference of a robustly resolved *Mesorhizobium* phylogeny (Fig. 2, Supplementary Figure S2). The taxa included in the analysis were representative of species included in *Mesorhizobium senso stricto* as defined in the latest release of the GTDB (Supplementary Table S1). Our dataset therefore excluded *M. oceanicum* (now named *Aquibium oceanicum*) (Kim et al., 2022), as well as *M. alexandrii*, *M. camelthorni*, *M. dentrificans*, *M. ephedrae*, *M. liriopis*, *M. microcysteis*, *M. soli*, *M. rhizophilum*, *M. sediminum* and *M. zhangyense*. The dataset also excluded *Mesorhizobium* species that were reclassified on GTDB as members of other species of *Mesorhizobium senso stricto*. These were *M. delmotii* (synonym of *M. temperatum*), *M. sanjuanii* (synonym of *M. helmanticense*), and *M. hominis* and *M. hungaricum* (both synonyms of *M. terrae*).

The 18 *Mesorhizobium* strains from *V. karroo* formed seven distinct lineages with 100 % bootstrap support and SH-aLRT $\geq$ 0.8 for all branches. Our strains mostly did not group closely to known *Mesorhizobium* species or representatives of undescribed species included in the GTDB. This was particularly evident for the four clusters containing strains VK25A + VK25D, strains VK2B + VK2D, strains VK23A + VK23B + VK23D, and strains VK9D + VK3E + MSK 1335. Strain VK24D was most closely related to GTDB representative, *Mesorhizobium* sp. ORS3324, while the clusters containing strains VK22E + VK22B + VK3C + VK23E + VK4C and strains VK1D + VK4B grouped closely to GTDB representatives *Mesorhizobium* sp. M1B.F.Ca.ET.045.04.1.1 and *Mesorhizobium* sp. M4B.F.Ca.ET.203.01.1.1, respectively. From the phylogeny alone, however, it was impossible to infer species boundaries, especially in the case of clusters that are closely related to other taxa.

Species delineation based on genealogical concordance

For species-level identifications, we utilized the fact that genealogical concordance among multiple gene regions typically demarcates species boundaries (Venter et al., 2017). Analysis with the UBCG pipeline accordingly indicated that all the seven clusters of strains mentioned above had a gene support index value of 92 (Fig. 2). In other words, the respective clusters were supported by phylogenies inferred from each of the 92 single UBCGs.

Despite the increased availability of *Mesorhizobium* genomes in public repositories, some type strains potentially related to our strains lacked whole genome sequences. We accordingly generated single gene trees using nucleotide sequences for *recA*, *atpD*, *glnII*, *rpoB*, and *gyrB* (Supplementary Figures S3-S7) and a corresponding MLSA phylogeny (Supplementary Figure S8). In the concatenated five-gene tree, strains VK2B and VK2D were closely related to *M. acaciae* RITF741^T that lacks a genome sequence (Supplementary Figure S8), but their clustering with this known species was only supported by the *glnII* genealogy. However, the *recA*, *gyrB* and *rpoB* genealogies grouped strains VK1D and VK4B with *M. abyssinicae* AC98c^T, with high bootstrap support. For this type strain, no *atpD* and *glnII* sequence data are publicly available and its genome has not yet been sequenced.

Based on the results of these two sets of analyses, seven species hypotheses were delineated. Of these, six were novel and represented by (i) strains VK2B and VK2D, (ii) strains VK25A and VK25D, (iii) strains VK23A, VK23B and VK23D, (iv) strains VK9D, VK3E, and MSK 1335, (v) strains VK22E, VK22B, VK3C, VK23E and VK4C and (vi) strain VK24D. Strains VK1D and VK4B are probably members of *M. abyssinicae*. We accordingly subjected these hypotheses to further genome and phenotype-based analyses to identify corroborating evidence.

Genome-based support for species hypotheses

The species hypotheses examined in this study, all yielded ANI values below 96 % when compared with their closest known relatives (Fig. 3, Supplementary Table S3). The only exception was the two putative *M. abyssinicae* strains (VK1D and VK4B) that shared ANI values of > 97.7 % with GTDB representative *Mesorhizobium* sp. M4B.F.Ca.ET.203.01.1.1. For all the multi-strain species hypotheses, ANI values well above 96 % were recorded among the strains of that species (Fig. 3, Supplementary Table S3). The only two exceptions were the cluster containing strains VK9D, VK3E and MSK 1335 and the cluster containing strains VK22E, VK22B, VK3C, VK23E and VK4C. In the case of MSK 1335, comparisons with strains VK9D and VK3E yielded ANI values of 91.7 % and 91.8 %, respectively, although the latter two strains shared an ANI of $\geq$ 96.5 %. These data thus indicate that strain MSK 1335 represents a species separate from strains VK9D and VK3E. In the case of the second cluster, strain VK4C shared ANI values of 94.2–94.5 % with strains VK22E, VK22B, VK3C and VK23E that shared ANI values of $\geq$ 95.7 % among themselves. These data suggest that strain VK4C may represent a distinct species, although further study of the five strains is required to unequivocally support strain VK4C's novelty.

In total, the genome data supported seven novel species hypotheses. We accordingly propose that a representative genome for each serve as the nomenclatural type (indicated with ^{Ts}) with the names *M. vachelliae* sp. nov. (VK25A^{Ts} and VK25D), *M. humile* sp. nov. (VK2B^{Ts} and VK2D), *M. australafricanum* sp. nov. (VK9D^{Ts} and VK3E), *M. montanum* sp. nov. (MSK 1335^{Ts}), *M. dulcispinae* sp. nov. (VK23D^{Ts}, VK23B, and VK23A), *M. captivum* sp. nov. (VK22E^{Ts}, VK22B, VK3C, VK23E, and VK4C), and *M. album* sp. nov. (VK24D^{Ts}). The eighth group, which included strains VK1D and VK4B, is conspecific to the unnamed GTDB representative *Mesorhizobium* sp. M4B.F.Ca.ET.203.01.1.1, and likely also *M. abyssinicae* AC98c^T. The formal protologues for the seven new species are given in Tables 2-8.

Phenotypic support for species hypotheses

All *Mesorhizobium* strains examined were Gram-negative, rod-shaped motile rods capable of growing across a broad temperature range from 4 °C to 40 °C, pH ranges from 5 to 10, and tolerate different NaCl concentrations from 0.3 % to 2.5 % (v/w) (Fig. 2, Table 9, Supplementary Table S4). Results for the API® 20NE tests (Table 9, Supplementary Table S5) and Biolog GenIII Microplate™ assays (Table 9, Supplementary Table S6) generally demonstrated high variability within species, although several traits reflected phenotypic coherence among members of a species. For example, only *M. album* was able to reduce NO_3 to N_2, while *M. montanum* and *M. vachelliae* were able to reduce NO_3 to NO_2. The strains of *M. dulcispinae* and *M. captivum* demonstrated variable reducing capacity and *M. humile* and *M. australafricanum* could not reduce NO_3 or NO_2. Additionally, all strains of *M. australafricanum* could utilise a variety of sugars, carboxylic acids, esters and fatty acids as sole carbon sources, while strains of *M. vachelliae* were all able to utilise dextrin, D-maltose, D-trehalose, D-cellobiose, gentiobiose, sucrose, D-turanose, stachyose, methyl pyruvate, D-lactic acid methyl ester, L-lactic acid, D-malic acid, L-malic acid, bromo-succinic acid, D-sorbitol, D-mannitol, myo-inositol and glycerol. In the case of *M. captivum* and *M. dulcispinae*, all strains were able to utilise D-glucose-6-PO_4, D-Fructose-6-PO_4, D-sorbitol, D-mannitol, myo-inositol and glycerol as carbon sources.

The putative *M. abyssinicae* strains VK1D and VK4B shared characteristics with the type strain AC98c^T of this species. This included the ability to utilise dextrin and glycyl-L-proline, as well as growth on L-Serine as sole carbon source (Degefu et al., 2013). Strains VK1D and

(caption on next page)

Fig. 2. Condensed UBCG phylogeny of *Mesorhizobium* genomes, together with the growth characteristics of the strains identified from root nodules of *Vachellia karroo* obtained from trapping experiments using soil sampled across South Africa. Concatenation of the 92 core-gene alignments resulted in a UBCG phylogeny that is rooted with *B. japonicum* USDA 6 ^{T}, as well as a condensed branch shown as an isosceles triangle (at the top of the phylogeny) that includes *Mesorhizobium* taxa more distantly related to the strains from this study. The full phylogeny can be viewed in the Supplementary Figure S2. Only representative genomes and type strains that are classified as *Mesorhizobium senso stricto* on GTDB Release 08-RS214 (28th April 2023) available at: https://gtdb.ecogenomic.org/, were included in the maximum likelihood analyses. All type strains are indicated with ^{T} and such names that are currently not validly published under the ICNP are indicated in quotation marks. The strains identified in this study are indicated in bold and colour coded with the representative genomes indicated with ^{Ts}. Branch support for the UBCG phylogeny was based on 1000 repetitions (%) and only ultrafast bootstrap approximation support values below 100 % are indicated on the phylogeny at the node of the supported branch. All SH-aLRT values were ≥ 0.8 for each bipartition and are not indicated on the figure. Black circles on the branches signifies bipartitions that are supported in all gene trees (n = 92), notably, this is exclusively applied to branches where the strains isolated in this study and the closest sister taxon are terminal taxa. The scale bar of the phylogeny represents the number of nucleotide substitutions per site. Growth characteristics included temperature (°C), pH and NaCl (w/v %) ranges as colour bars with downward-facing triangles demonstrating the minima and maxima growth ranges of the respective species clusters.

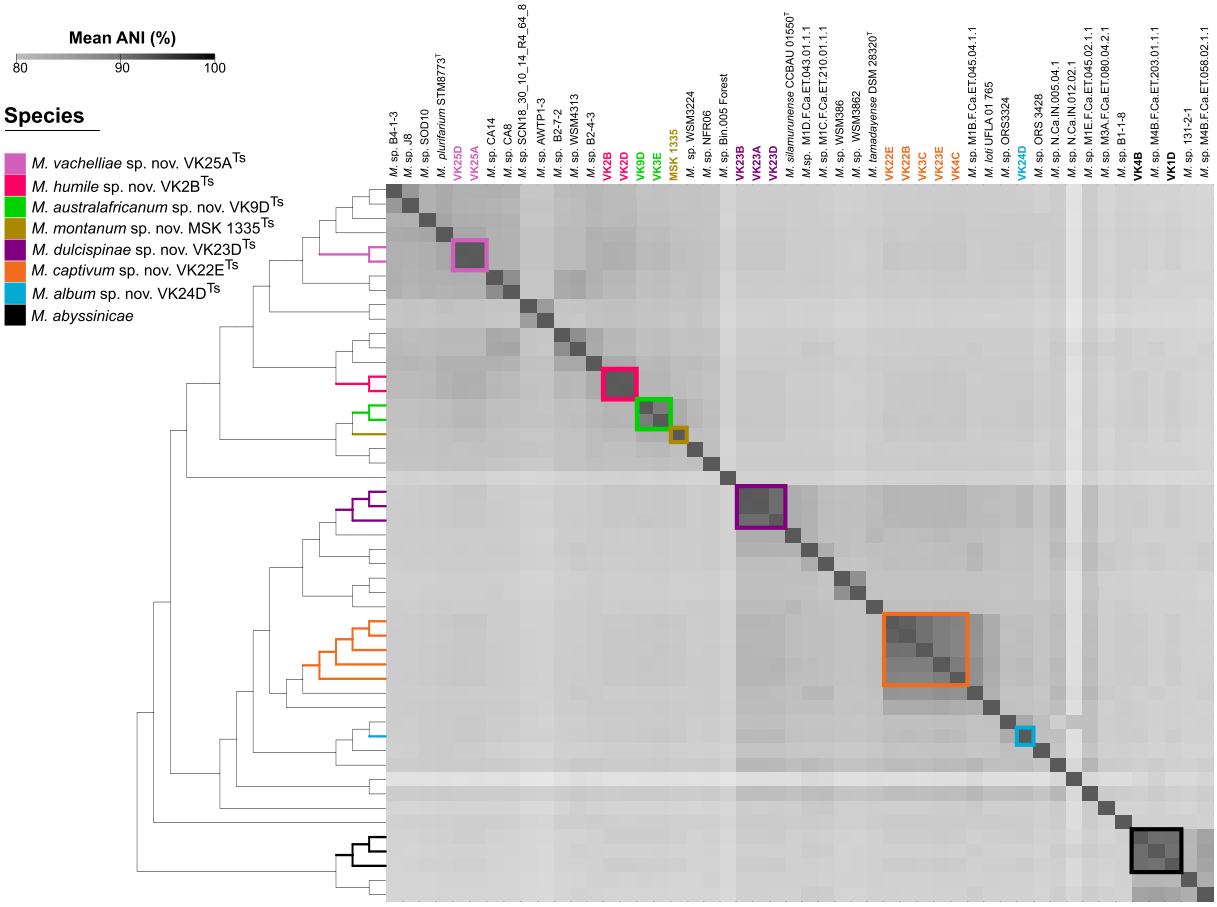

Fig. 3. The depicted grayscale heatmap demonstrates the matrix of the mean Average Nucleotide Identities (ANI) values based on the nucleotide BLAST search as implemented in JSpecies version 1.2.1 (Richter and Rosselló-Móra, 2009), with upper and lower triangular pairwise ANI values considered. A dendogram of the condensed UBCG phylogeny demonstrates the phylogenetic grouping of the strains as observed in Fig. 2 and Supplementary Figure S2. Coloured blocks represent the species clusters and representative genomes are indicated with ^{Ts}. All genome accession numbers are listed in Supplementary Table S1 and the full result of the ANI analyses can be found in Supplementary Table S3.

VK4B and strain AC98c^{T} all shared the ability to grow at pH ranges of 4 to 9, but VK1D and VK4B could tolerate higher temperatures and NaCl concentrations, ≥35 °C and ≥ 0.5 %, respectively (Degefu et al., 2013). Although different methods were used to characterize *M. abyssinicae* AC98c^{T}, our results further suggested that this species likely also includes substantial phenotypic diversity as many of the notable traits previously reported (Degefu et al., 2013) appears to be variable. These include phenotypic traits such as the use of formic acid, propionic acid and p-hydroxy-phenylacetic acid as sole carbon sources (see Table 9 and information by Degefu et al., 2013).

Nodulation confirmation and symbiovar identification

We first confirmed that the genomes for ten representative strains (MSK1335^{Ts}, VK1D^{Ts}, VK4B, VK25A^{Ts}, VK2B^{Ts}, VK3E, VK24D^{Ts}, VK23D^{Ts}, VK4C, VK22E^{Ts}, and VK23E) examined in this study, do encode common *nod* and *nif* genes (Table 1). Then the nodulation capability of ten strains were confirmed on the original host, *V. karroo*. All the representative strains were able to induce effective nodules, evident from the pink interior of the root nodule seven weeks post-inoculation. Following sequence analyses of either *atpD* or *rpoB*, the nodule-isolated bacteria were confirmed as identical to those originally used as inocula.

By making use of *nodC* sequences, we further investigated whether

Table 1
Statistics for the draft genomes [a] of representative strains for the seven new *Mesorhizobium* species delineated and described in this study.

Genome statistics [b,c]	*M. vachelliae* (VK25A^{Ts})	*M. humile* (VK2B^{Ts})	*M. australafricanum* (VK9D^{Ts})	*M. montanum* (MSK 1335^{Ts})	*M. dulcispinae* (VK23D^{Ts})	*M. captivum* (VK22E^{Ts})	*M. album* (VK24D^{Ts})
GenBank genome accession [α, ε]	GCA_033977325.1	GCA_033977215.1	GCA_033977265.1	GCA_033977145.1	GCA_033977345.1	GCA_033977165.1	GCA_033977205.1
NCBI SRA accession [α]	SRR26261387	SRR26261391	SRR26261389	SRR26261393	SRR26261378	SRR26261381	SRR26261376
SeqCode registry no. [α]	32828	32827	32829	32826	32830	32831	32771
Total length (bp) [δ]	7,055,270	6,749,911	6,462,116	6,869,492	6,698,899	6,954,195	7,217,979
Contig no. [δ]	83	77	74	76	78	142	176
Largest contig size (bp) [δ]	706,620	613,800	726,971	501,236	953,217	409,668	381,941
N50 [δ]	364,428	234,275	348,747	220,403	352,659	124,007	123,883
N75 [δ]	191,540	144,717	153,509	148,237	164,634	69,743	65,048
L50 [δ]	7	10	7	11	6	18	19
L75 [δ]	14	19	14	20	13	36	39
Completeness (%) [δ]	99.1	99.1	99.1	99.1	99.1	99.1	99.1
Contamination (%) [δ]	1.9	0.9	1.9	0.9	0.9	0.9	0.9
GC mole % [δ]	63.61	63.56	63.15	62.94	63.48	63.30	63.25
Average read coverage [α]	174.885	230.757	153.078	127.153	232.521	225.423	144.491
Percent mapped [δ]	98.841	99.115	99.065	99.829	99.556	99.645	98.626
16S rRNA gene accession [α, β]	OR735554	OR735550	OR735552	OR735540	OR735547	OR735544	OR735549
Fraction of most complete 16S rRNA gene fragment (%) [δ]	100	100	100	100	100	100	100
No. of standard amino acids with detected tRNA elements [δ]	20	20	20	20	20	20	20
Closest relative reference [γ]	GCF_000824785.1	n/a	GCF_003289945.1	GCF_003289945.1	GCF_003952385.1	GCF_003952465.1	n/a
nodA GenBank accession [θ]	OR800572	OR800571	OR800570	OR800576	OR800574	OR800575	OR800573
nodC GenBank accession [θ]	PP430827	PP430828	PP430829	PP430830	PP430831	PP430832	PP430833
nifH GenBank accession [θ]	OR800580	OR800577	OR800578	OR800583	OR800582	OR800581	OR800579

[a] Statistics are based on assemblies built from contigs of ≥ 500 base pairs (bp).
[b] In the table, [α] indicates required information of the designated genome to be named under the SeqCode, [β] indicates requirements for validation of a type strain name under the ICNP. Information that is recommended, but optional for name validation under the SeqCode and the ICNP are indicated with [δ]. The table also indicates information regarding the minimum requirements ([ε]) and recommendations ([θ]) for the description of rhizobia as determined by the Subcommittee on Taxonomy of Rhizobia and Agrobacteria, along with all the requirements as set out by the ICNP. The closest reference genome as determined by GTDB-Tk version 1.7.0 is indicated with [γ] and all genomes were classified as *Mesorhizobium senso stricto* with ANI values shared between the two genomes all < 96 % and instances where the genome was not assigned to the closest species as it falls outside the pre-defined ANI radius is indicated with n/a (full result available at https://kbase.us/n/159176/19/).
[c] Abbreviations: NCBI = National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov); SRA = Sequence Read Archive; N50 and N75 = contig lengths where 50% and 75% of the genome is contained in the shortest contigs of at least that length; L50 and L75 = number of contigs needed to cover 50% and 75% of the genome.

the bacteria investigated were related to *Mesorhizobium* strains with characterized legume specificities. This is because the correlation of *nodC* gene sequences with legume-host range can be used for delineating symbiovars in *Mesorhizobium* (Peix et al., 2015; León-Barrios et al., 2021). However, phylogenetic analysis showed that all the South African strains examined grouped together, but distinct from other known symbiovars (Supplementary Figure S9).

Discussion

Here we described seven novel *Mesorhizobium* species from South Africa using the SeqCode (Hedlund et al., 2022). Since introduction, it has been used to validly name numerous unculturable strains that would have been otherwise classified as *Candidatus* taxa with names having no standing under the ICNP (e.g., Nguyen et al., 2023, Eddie et al., 2017). While the SeqCode was primarily developed to bring stability to taxa for which pure cultures are typically not available, it also provides an effective solution for validly naming cultivated organisms for which the deposition of strains into different culture collections is difficult (Konstantinidis and Rosselló-Móra, 2015). For example, the nitrate-reducing bacteria isolated and cultured from a lagoon in India could not be validated under the ICNP due to restrictions on the export of live cultures, but were later validated as *Roseiconus nitratireducens* and *Ro. lacunae* under the SeqCode (SeqCode Register list: seqco.de/r:v04mpu86; Kumar et al., 2021). A more recent example is the cultured species *Salinibacter pepae* that was named and validated under the SeqCode as the fastidious nature of the bacterium prevented deposition of a viable axenic culture in a second culture collection (SeqCode Register list: seqco.de/r:b5vsvzg3; Viver et al., 2023).

The 18 rhizobial symbionts considered in this study were delineated to species level using genome data and the principles of genealogical concordance (Rahimlou et al., 2021; Sutcliffe et al., 2021; Venter et al., 2017). Previously, such studies would have employed only a handful of housekeeping genes (e.g., Avontuur et al., 2022; Claassens et al., 2023; Mavima et al., 2021), a practice that often precludes comprehensive taxonomic resolution (Hördt et al., 2020; Young et al., 2021). The use of

Table 2
Protologue description of *M. montanum* sp. nov.

Species name	*Mesorhizobium montanum*
Species status	sp. nov.
Specific epithet	*montanum*
Species etymology	mon.ta'num. L. neut. adj. *montanum,* pertaining to mountains, in particular the Kamiesburg mountains, surrounding Kamieskroon, the region where soil was sampled from.
Diagnostic traits and description of novel taxon	Cells are Gram-negative, motile rods. On YM agar, following 5 days of incubation at 28 °C, the colonies are circular, cream in colour, translucent with entire margins and convex elevations with mucoid consistency due to the excessive production of exopolysaccharides. The strain was not able to tolerate a pH of ≤ 6, but could grow at a pH of 10 and a NaCl concentration of 0.5 % to 1 %. The strain was able to grow at 15 °C to 37 °C. The strain tested positive for the activity of nitrate reduction to nitrite, arginine and urea hydrolysis and the assimilation of 4-nitrophenyl-β,D-galactopyranoside and D-glucose. The strain could utilise the following sugar carbon sources: dextrin, D-maltose, D-trehalose, D-cellobiose, gentiobiose, sucrose, D-turanose, stachyose, D-raffinose, α-D-lactose, D-melibiose, β-methyl-D-glucoside, D-salicin, N-acetyl-D-glucosamine, N-acetyl-β-D-mannosamine, N-acetyl-D-galactosamine, N-acetyl neuraminic acid, α-D-glucose, D-mannose, D-fructose, D-galactose, 3-methyl glucose, D-fucose, L-fucose, L-rhamnose, inosine. Similarly, the strain could utilise D-sorbitol, D-mannitol, myo-inositol, glycerol, D-arabitol, D-aspartic acid, D-serine, D-glucose6-PO_4, D-fructose6-PO_4, glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, L-pyroglutamic acid, L-serine, pectin, D-galacturonic acid, L-galactonic acid lactone, D-gluconic acid, D-glucuronic acid, glucuronamide, mucic acid, quinic acid, D-saccharic acid, p-hydroxy-phenylacetic acid, methyl pyruvate, D-lactic acid methyl ester, L-lactic acid, citric acid, α-keto-glutaric acid, D-malic acid, L-malic acid, bromo-succinic acid, Tween 40, γ-amino-butryric acid, α-hydroxy-butyric acid, β-hydroxy-D,L-butyric acid, α-keto-butyric acid, acetoacetic acid, propionic acid, acetic acid and formic acid as carbon sources. The strain was able to form effective symbiosis with *V. karroo*.
Designated Genome	MSK 1335^{Ts}
Genome accession number	GCA_033977145.1^{Ts}
SRA accession	SRR26261393
Country of origin	South Africa
Region of origin	Soil sampled from Kamieskroon, Northern Cape Province
Isolation source	The root nodules of *Vachellia karroo.*
Date of isolation	2013
16S rRNA gene accession	OR735540
Genome status	Draft genome
Genome size (bp)	6 869 492
Size of largest contig (bp)	501 236
GC mol %	62.94
N50	220 403
Number of contigs	76
Read coverage	127.153
Number of strains in species cluster [a]	n/a
SeqCode registry URL	32826

[a] The number of strains in the species cluster and excludes the strain for which the genome was designated as the nomenclatural type.

Table 3
Protologue description of *M. humile* sp. nov.

Species name	*Mesorhizobium humile*
Species status	sp. nov.
Specific epithet	*humile*
Species etymology	hu.mi'le. L. neut. adj. *humile,* humble, referring to the type strain's small colony size on YM agar and limited carbon source utilisation.
Diagnostic traits and description of novel taxon	Cells are Gram-negative, motile rods. On YM agar, following 5 days of incubation at 28 °C, the colonies are small, circular, white, and opaque with entire margins and flat elevations with a dry consistency. The strain was able to grow in the pH range of 6 to 9 and tolerate NaCl concentrations of 0.3 % to 1.5 %. The strain was able to grow at 4 °C to 35 °C. The strain tested positive for urea and esculin ferric citrate hydrolysis. The strain could assimilate 4-nitrophenyl-β,D-galactopyranoside, D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, potassium gluconate, adipic acid and malic acid. The strain was able to form effective symbiosis with *V. karroo*.
Designated Genome	VK2B^{Ts}
Genome accession number	GCA_033977215.1^{Ts}
SRA accession	SRR26261391
Country of origin	South Africa
Region of origin	Soil sampled from Van Rhyn's Pass, Western Cape Province
Isolation source	The root nodules of *Vachellia karroo*
Date of isolation	2013
16S rRNA gene accession	OR735550
Genome status	Draft genome
Genome size (bp)	6 749 911
Size of largest contig (bp)	613 800
GC mol %	63.56
N50	234 275
Number of contigs	77
Read coverage	230.757
Number of strains in species cluster [a]	1
SeqCode registry URL	32827

[a] The number of strains in the species cluster and excludes the strain for which the genome was designated as the nomenclatural type.

genome-based taxonomy has proven much more reliable in delineating other rhizobial taxa, such as *Bradyrhizobium* (Ormeño-Orrillo and Martínez-Romero, 2019) and other genera of the family *Rhizobiaceae* (Ma et al., 2023). In the current study, putative species boundaries were determined using phylogenies inferred from 92 core genes to find genealogically concordant clusters of strains. This task was simplified using the UBCG pipeline which provides an indication of how many of the 92 genes supported particular clusters via its gene support index (Na et al., 2018). Therefore, in addition to being an invaluable and easy-to-use tool for genome-wide phylogenetic analyses as previously shown for the *Rhizobiaceae* (Ma et al., 2023), the UBCG pipeline also advances the functional principle of delineating bacterial taxa using genealogical concordance by directly comparing the number of core genes supporting an inferred lineage. Undoubtedly, the UBCG pipeline and its recent successor, UBCG2 (Kim et al., 2021) will continue to assist taxonomists with the delineation of objective species hypotheses for evaluation using a polyphasic approach (Venter et al., 2017).

Given that genomes for type strains of certain *Mesorhizobium* species are not available, a multipronged strategy was used to ensure the novelty of our seven new species. The first step was to identify species lacking type strains with sequenced genomes by comparing information from the GTDB and the LPSN (Parks et al., 2022; Parte et al., 2020). To deal with the taxa missing in our 92-gene dataset, sequences previously generated for MLSA and available in public nucleotide databases were utilised. This allowed for the identification of two *V. karroo* symbionts as putative members of *M. abyssinicae*, the type strain (AC98c^{T}) of which lacks a publicly available genome sequence. This species was first isolated from the root nodules of *Vachellia abyssinica* and *V. tortilis* in Ethiopia (Degefu et al., 2011; Degefu et al., 2013), although strains of *M. abyssinicae* have also been identified in the root nodules of other legumes in Ethiopia (Aserse et al., 2013; Tena et al., 2017). Additionally, our 92-gene analysis showed that *M. abyssinicae* likely also includes *Mesorhizobium* sp. M4B.F.Ca.ET.203.01.1.1 which was originally

Table 4
Protologue description of *M. vachelliae* sp. nov.

Species name	*Mesorhizobium vachelliae*
Species status	sp. nov.
Specific epithet	*vachelliae*
Species etymology	va.chel.li'ae. N.L. gen. n. *vachelliae,* of *Vachellia,* referring to the host plant from which this organism was recovered.
Diagnostic traits and description of novel taxon	Cells are Gram-negative, motile rods. On YM agar, following 5 days of incubation at 28 °C, the colonies are circular, cream, translucent with entire margins and convex elevations with viscid consistency. The strain was able to grow in the pH range of 5 to 10 and tolerate a NaCl concentration of 0.3 % to 2.5 %. The strain was able to grow at 4 °C to 37 °C. The strain tested positive for the activity of nitrate reduction to nitrite, urea and esculin ferric citrate hydrolysis. The strain could assimilate 4-nitrophenyl-β,D-galactopyranoside, D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose and malic acid. The strain could utilize dextrin, D-maltose, D-trehalose, D-cellubiose, gentiobiose, sucrose, D-turanose, stachyose, D-raffinose, D-melibiose, β-methyl-D glucoside, D-salicin, N-acetyl-D glucosamine, N-acetyl- β -D mannosamine, N-acetyl-D galactosamine, α-D-glucose, D-mannose, D-fructose, D-galactose, 3-methyl glucose, D-fucose, L- fucose, L-rhamnose, D-sorbitol, D-mannitol, D-arabitol, myo-inositol, glycerol, D-fructose6-PO_4, Glycyl-L-proline, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, pectin, glucuronamide, methyl pyruvate, D-lactic acid methyl ester, L-lactic acid, citric acid, D-malic acid, L-malic acid, bromo-succinic acid, γ-amino-butyric acid, α-hydroxy-butyric acid, β-hydroxy-D-L-butyric acid, acetoacetic acid, propionic acid, acetic acid and formic acid as sources of carbon. The strain was able to form effective symbiosis with *V. karroo*.
Designated Genome	VK25A^{Ts}
Genome accession number	GCA_033977325.1^{Ts}
SRA accession	SRR26261387
Country of origin	South Africa
Region of origin	Soil isolated from Mookgophong, Limpopo Province
Isolation source	The root nodules of *Vachellia karroo*
Date of isolation	2013
16S rRNA gene accession	OR735554
Genome status	Draft genome
Genome size (bp)	7 055 270
Size of largest contig (bp)	706 620
GC mol %	63.61
N50	364 428
Number of contigs	83
Read coverage	174.885
Number of strains in species cluster [a]	1
SeqCode registry URL	32828

[a] The number of strains in the species cluster and excludes the strain for which the genome was designated as the nomenclatural type.

Table 5
Protologue description of *M. australafricanum* sp. nov.

Species name	*Mesorhizobium australafricanum*
Species status	sp. nov.
Specific epithet	*australafricanum*
Species etymology	aus.tral.a.fri.ca'num. L. neut. adj. *australis*, southern; L. neut. adj. *africanum*, African; N.L. neut. adj. *australafricanum*, pertaining to southern Africa.
Diagnostic traits and description of novel taxon	Cells are Gram-negative, motile rods. On YM agar, following 5 days of incubation at 28 °C, the colonies are circular, cream, translucent with entire margins and convex elevations with mucoid consistency. The strain was able to grow in the pH range of 5 to 10 and tolerate a NaCl concentration of 0.3 % to 2.5 %. The strain can grow at 10 °C to 35 °C. The strain tested positive for the activity of D-glucose fermentation and urea and esculin ferric citrate hydrolysis. The strain could assimilate 4-nitrophenyl-β,D-galactopyranoside, D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, adipic acid and malic acid. The strain could utilize dextrin, D-maltose, D-trehalose, D-cellubiose, gentiobiose, sucrose, D-turanose, stachyose, D-raffinose, α-D-lactose, D-melibiose, β-methyl-D glucoside, D-salicin, N-acetyl-D glucosamine, N-acetyl-β–D mannosamine, N-acetyl-D galactosamine, N-acetyl neuraminic acid, α-D-glucose, D-mannose, D-fructose, D-galactose, 3-methyl glucose, D-fucose, L- fucose, L-rhamnose, inosine, D-sorbitol, D-mannitol, D-arabitol, myo-inositol, glycerol, D-glucose6-PO_4, D-fructose6-PO_4, D-aspartic acid, gelatin, glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, L-pyroglutamic acid, pectin, D-galacturonic acid, L-galactonic acid lactone, D-gluconic acid, D-glucuronic acid, glucuronamide, D-saccharic acid, methyl pyruvate, D-lactic acid methyl ester, citric acid, α-keto-glutaric acid, L-lactic acid, D-malic acid, L-malic acid, bromo-succinic acid, Tween 40, γ-Amino-butyric acid, β-hydroxy-D-L-butyric acid, acetoacetic acid, propionic acid, acetic acid and formic acid as sources of carbon. The strain was able to form effective symbiosis with *V. karroo*.
Designated Genome	VK9D^{Ts}
Genome accession number	GCA_033977265.1^{Ts}
SRA accession	SRR26261389
Country of origin	South Africa
Region of origin	Soil isolated from Stutterheim, Eastern Cape province
Isolation source	The root nodules of *Vachellia karroo*
Date of isolation	2013
16S rRNA gene accession	OR735552
Genome status	Draft genome
Genome size (bp)	6 462 116
Size of largest contig (bp)	726 971
GC mol %	63.15
N50	348 747
Number of contigs	74
Read coverage	153.078
Number of strains in species cluster [a]	2
SeqCode registry URL	32829

[a] The number of strains in the species cluster and excludes the strain for which the genome was designated as the nomenclatural type.

isolated from chickpea rhizosphere soil in Ethiopia (Greenlon et al., 2019). The genome of this strain represents 23 other genomes in the GTDB, all of which were either reconstructed metagenomes or isolated from chickpea root nodules and rhizosphere soil in Ethiopia (Greenlon et al., 2019). Therefore, all of the previous research, combined with the available genome-based resources, would go a long way towards understanding the biology and ecology of this apparently widespread rhizobial species.

ANI represented a major line of evidence for the species hypotheses identified in this study (Richter and Rosselló-Móra, 2009). Although the 96 % ANI threshold is useful for distinguishing species in most natural populations (Gosselin et al., 2022; Rodriguez-R et al., 2021), this metric is often taxon-specific and regarded merely as an aid to taxonomic descriptions (Goris et al., 2007; Konstantinidis and Tiedje, 2005; Palmer et al., 2020). In the current study, for example, all except one of the *M. captivum* sp. nov. strains shared ANI values exceeding 96 %. However, we proposed the inclusion of the outlier strain (i.e., VK4C) based on its close phylogenetic grouping with all other strains of *M. captivum* sp. nov. Further research is needed to interrogate the taxonomic status of strain VK4C, and until such time we propose treating it as *M. captivum*. Nevertheless, all ANI values estimated for strains from the same species were above 96 %, reiterating a higher average genome identity between

Table 6
Protologue description of *M. album* sp. nov.

Species name	*Mesorhizobium album*
Species status	sp. nov.
Specific epithet	*album*
Species etymology	al'bum. L. neut. adj. *album,* white, referring to the white colonies of the type strain on YM agar.
Diagnostic traits and description of novel taxon	Cells are Gram-negative, motile rods. On YM agar, following 5 days of incubation at 28 °C, the colonies are circular, white, opaque with entire margins and convex elevations with viscid consistency. The strain was able to grow in the pH range of 6 to 9 and tolerate a NaCl concentration of 0.3 % to 2.5 %. The strain can grow at 15 °C to 40 °C. The strain tested positive for the ability to reduce nitrates to nitrites and nitrogen, and could hydrolyse arginine, urea, esculin ferric citrate and gelatin. The strain could assimilate 4-nitrophenyl-β,D-galactopyranoside, D-glucose, L-arabinose, potassium gluconate, malic acid and phenylacetic acid. The strain could utilize dextrin, D-maltose, D-trehalose, D-cellubiose, gentiobiose, sucrose, D-turanose, stachyose, D-raffinose, α-D-lactose, D-melibiose, β-methyl-D glucoside, α-D-glucose, D-mannose, D-fructose, D-galactose, 3-methyl glucose, D-fucose, L-fucose, L-rhamnose, inosine, Glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, L-pyroglutamic acid, L-serine, pectin, D-galacturonic acid, L-galactonic acid lactone, D-gluconic acid, glucuronamide, mucic acid, quinic acid, D-saccharic acid, methyl pyruvate, D-lactic acid methyl ester, L-lactic acid, D-malic acid, L-malic acid, bromo-succinic acid, Tween 40, β-hydroxy-D,L-butyric acid, acetoacetic acid and acetic acid as sources of carbon. The strain was able to form effective symbiosis with *V. karroo*.
Designated Genome	VK24D^{Ts}
Genome accession number	GCA_033977205.1^{Ts}
SRA accession	SRR26261376
Country of origin	South Africa
Region of origin	Soil isolated from Mookgophong, Limpopo Province
Isolation source	The root nodules of *Vachellia karroo*
Date of isolation	2013
16S rRNA gene accession	OR735549
Genome status	Draft genome
Genome size (bp)	7 217 979
Size of largest contig (bp)	381 941
GC mol %	63.25
N50	123 883
Number of contigs	176
Read coverage	144.491
Number of strains in species cluster [a]	n/a
SeqCode registry URL	32771

[a] The number of strains in the species cluster and excludes the strain for which the genome was designated as the nomenclatural type.

Table 7
Protologue description of *M. dulcispinae* sp. nov.

Species name	***Mesorhizobium dulcispinae***
Species status	sp. nov.
Specific epithet	*dulcispinae*
Species etymology	dul.ci.spi.na.e L. fem. adj. *dulcis,* sweet; L. fem. n. *spina,* thorn; N.L. gen. n. *dulcispinae,* of a sweet thorn, referring to the common name of *Vacellia karroo*, the host plant from which this organism was recovered.
Diagnostic traits and description of novel taxon	Cells are Gram-negative, motile rods. On YM agar, following 5 days of incubation at 28 °C, the colonies are circular, cream, translucent with entire margins and convex elevations with viscid consistency. The strain was able to grow in the pH range of 6 to 9 and tolerate a NaCl concentration of 0.3 % to 2.5 %. The strain can grow at 10 °C to 35 °C. The strain could hydrolyse urea and esculin ferric citrate and assimilate 4-nitrophenyl-β,D-galactopyranoside, D-glucose, L-arabinose, D-mannose, D-maltose and potassium gluconate. The strain could utilize dextrin, D-maltose, D-trehalose, D-cellubiose, gentiobiose, sucrose, D-turanose, stachyose, D-raffinose, α-D-lactose, D-melibiose, β-methyl-D glucoside, D-salicin, N-acetyl-D-glucosamine, N-acetyl-β-D-mannosamine, N-acetyl-D-galactosamine, α-D-glucose, D-mannose, D-fructose, D-galactose, 3-methyl glucose, D-fucose, L- fucose, L-rhamnose, D-sorbitol, D-mannitol, D-arabitol, myo-inositol, glycerol, D-glucose6-PO_4, D-fructose6-PO_4, glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, L-pyroglutamic acid, L-serine, pectin, D-galacturonic acid, L-galactonic acid lactone, D-gluconic acid, D-glucuronic acid, glucuronamide, mucic acid, quinic acid, D-saccharic acid, P-hydroxy-phenylacetic acid, L-lactic acid, D-lactic acid methyl ester, citric acid, α-keto-glutaric acid, L-lactic acid, D-malic acid, L-malic acid, bromo-succinic acid, Tween 40, γ-amino-butyric acid, β-hydroxy-D-L-butyric acid, acetoacetic acid, propionic acid and acetic acid as sources of carbon. The strain was able to form effective symbiosis with *V. karroo*.
Designated Genome	VK23D^{Ts}
Genome accession number	GCA_033977345.1^{Ts}
SRA accession	SRR26261378
Country of origin	South Africa
Region of origin	Soil isolated from Mookgophong, Limpopo Province
Isolation source	The root nodules of *Vachellia karroo*
Date of isolation	2013
16S rRNA gene accession	OR735547
Genome status	Draft genome
Genome size (bp)	6 698 899
Size of largest contig (bp)	953 217
GC mol %	63.48
N50	352 659
Number of contigs	78
Read coverage	232.521
Number of strains in species cluster [a]	3
SeqCode registry URL	32830

[a] The number of strains in the species cluster and excludes the strain for which the genome was designated as the nomenclatural type.

strains in a species cluster as opposed to those of genomes from different *Mesorhizobium* species (Goris et al., 2007; Ormeño-Orrillo et al., 2015; Yoon et al., 2017).

As expected, the incorporation of physiological and biochemical attributes enhanced our knowledge of the biological traits of these *Mesorhizobium* strains, while at the same time also demonstrating intra-species coherence and interspecies differentiation (Brady et al., 2010). For example, our strains generally showed a high degree of tolerance to environmental stressors such as salinity, temperature and pH, which corresponds what is known for many other *Mesorhizobium* type strains (Laranjo and Oliveira, 2011; Laranjo et al., 2014). In terms of intra-species cohesion, various shared commonalities (e.g., urea, esculin and gelatin hydrolysis, and the use of certain sugars and hexose phosphates as sole carbon source) were observed for *M. vachelliae* sp. nov., *M. australafricanum* sp. nov., *M. humile* sp. nov., *M. dulcispinae* sp. nov. and *M. captivum* sp. nov.. Some of these conserved characters differed among species (e.g., strains of *M. abyssinicae* and *M. vachelliae* sp. nov. grew across a wider range of temperature, salinity and pH than the other species). However, we also observed substantial within-species variation for many other traits. These included the capacity for utilising certain carbon sources, assimilating various substrates and for reducing nitrate to nitrogen. Such intra-species variation is not uncommon as the assays for scoring the relevant characters are non-reproducible in some cases (De Lajudie et al., 2019; Sutcliffe, 2015). The lack of phenotypic coherence in some cases could also be attributed to these metabolic

Table 8
Protologue description of *M. captivum* sp. nov.

Species name	*Mesorhizobium captivum*
Species status	sp. nov.
Specific epithet	*captivum*
Species etymology	cap.ti'vum L. neut. adj. *captivum,* captured or captive, referring to the capturing of this organism with a compatible rhizobial host.
Diagnostic traits and description of novel taxon	Cells are Gram-negative, motile rods. On YM agar, following 5 days of incubation at 28 °C, the colonies are circular, cream, translucent with entire margins and convex elevations with viscid consistency. The strain was able to grow in the pH range of 6 to 9 and tolerate a NaCl concentration of 0.3 % to 1.5 %. The strain can grow at 15 °C to 35 °C. The strain tested positive for urea and esculin ferric citrate hydrolysis. The strain could assimilate 4-nitrophenyl-β,D-galactopyranoside, D-glucose, L-arabinose, D-mannose, D-mannitol, D-maltose, potassium gluconate, adipic acid and trisodium citrate. The strain could utilize dextrin, D-maltose, D-trehalose, D-cellubiose, gentiobiose, sucrose, D-turanose, D-raffinose, α-D-lactose, D-melibiose, β-methyl-D glucoside, D-salicin, N-acetyl-D glucosamine, N-acetyl-β-D mannosamine, N-acetyl-D galactosamine, α-D-glucose, D-mannose, D-fructose, D-galactose, 3-methyl glucose, D-fucose, L-fucose, L-rhamnose, D-sorbitol, D-mannitol, D-arabitol, myo-inositol, glycerol, D-glucose6-PO_4, D-fructose6-PO_4, glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, pectin, D-galacturonic acid, L-galactonic acid lactone, D-gluconic acid, D-glucuronic acid, glucuronamide, mucic acid, D-lactic acid methyl ester, citric acid, L-lactic acid, D-malic acid, L-malic acid, bromo-succinic acid, Tween 40, γ-amino-butyric acid, β-hydroxy-D-L-butyric acid, acetoacetic acid and acetic acid as sources of carbon. The strain was able to form effective symbiosis with *V. karroo*.
Designated Genome	VK22E^{Ts}
Genome accession number	GCA_033977165.1^{Ts}
SRA accession	SRR26261381
Country of origin	South Africa
Region of origin	Soil isolated from Mookgophong, Limpopo Province
Isolation source	The root nodules of *Vachellia karroo*
Date of isolation	2013
16S rRNA gene accession	OR735544
Genome status	Draft genome
Genome size (bp)	6 954 195
Size of largest contig (bp)	409 668
GC mol %	63.30
N50	124 007
Number of contigs	142
Read coverage	225.423
Number of strains in species cluster [a]	4
SeqCode registry URL	32831

[a] The number of strains in the species cluster and excludes the strain for which the genome was designated as the nomenclatural type.

properties being found on mobile genetic elements, whose mobility could negate the use of such traits for taxonomic purposes (Ormeño-Orrillo and Martínez-Romero, 2013).

The genome-based definition of *Mesorhizobium senso stricto* as implemented by the GTDB was used in this study (Parks et al., 2022). The GTDB's splitting of *Mesorhizobium* based on relative evolutionary divergence (Parks et al., 2018) is consistent with previous suggestions that the genus is non-monophyletic (Hördt et al., 2020). The GTDB currently classifies species with *Mesorhizobium* type strains into seven genera (i.e., *Mesorhizobium sensu stricto* and six novel genera that are yet to be described). Species retained in *Mesorhizobium sensu stricto* includes 36 type strains as listed on the LPSN, while 107 unnamed novel species are recognised by the GTDB (Parks et al., 2018). Although this assemblage of bacteria is relatively diverse, most represent rhizobial symbionts of forage legumes (Nandasena et al., 2009), medicinal plants (Jarvis and Tighe, 1994; Marcos-García et al., 2017; Martínez-Hidalgo and Hirsch, 2017), crops (Laranjo et al., 2008) and other ecologically important legumes (Degefu et al., 2013; Zhu et al., 2015). A small number of *Mesorhizobium* type strains have been isolated from unusual sources such as water (Pedron et al., 2021) and compost (Lin et al., 2019), while the bulk of the unnamed *Mesorhizobium sensu stricto* species included in the GTDB originate from unique environments such as methylated amine enrichment medium (BioSample: SAMN04076519), activated sludge (BioSample: SAMN21211502) and bioreactors (BioSample: SAMN05660516). Implementation of a more holistic view based on both the LPSN and GTDB would thus provide for a more unbiased overview of the microbial diversity of this genus (Chuvochina et al., 2023).

The findings of this study emphasise the need for more studies on the diversity and taxonomy of *Mesorhizobium*. In our 92-gene phylogeny, three of the new species formed part of larger clades consisting of GTDB representatives of undescribed species. The clade containing *M. humile* sp. nov. also included undescribed strains from soil (BioSamples: SAMN12071710 and SAMN12071672) and rhizosphere soil of *Cicer arietinum* from Australia (BioSample: SAMN07540851). The closest known relative of the clade containing *M. australafricanum* sp. nov. and *M. montanum* sp. nov. is an undescribed strain obtained from root nodule of *Otholobium candicans* in South Africa (BioSample: SAMN02440597) and another unnamed strain deposited into GenBank without any metadata (BioSample: SAMN03159463). Similarly, the closest relative of *M. album* sp. nov. was an undescribed strain from a root nodule of *Vachellia seyal* (Diouf et al., 2015). Although the other larger clades containing the remainder of our new species included at least one known species, they also included GTDB-recognized unnamed species. The clade containing *M. vachelliae* sp. nov. shared close relatives of *M. plurifarium* from the root nodules of *Senegalia senegal* (de Lajudie et al., 1998), as well as undescribed strains originating from soil and root nodules of *Senegalia senegal* (Diouf et al., 2015; Vinuesa et al., 2005) and *Glycyrrhiza uralensis* in Japan (BioSample: SAMD00251586). The clade with *M. dulcispinae* sp. nov. included *M silamurunense* isolated from root nodules of *Astragalus* species in China (Zhao et al., 2012), while the clade with *M. captivum* sp. nov. included a *M. loti* strain isolated from a zinc minefield (BioSample: SAMN04351373), as well as an undescribed strain isolated from *C. arietinum* root nodules (Greenlon et al., 2019).

The seven new *Mesorhizobium* species descriptions provided here comply with the rules of the SeqCode (Hedlund et al., 2022), although we also attempted to accommodate guidelines outlined by the Subcommittee on Taxonomy of Rhizobia and Agrobacteria of the International Committee on Systematics of Prokaryotes (ICSP) (de Lajudie et al., 2019). The Subcommittee previously acknowledged the SeqCode's value (Mousavi and Young, 2023), but because the SeqCode is not endorsed by the ICSP (Göker et al., 2022), the Subcommittee can only consider names to be validly published if they are described in accordance with the ICNP's rules (de Lajudie et al., 2019; Mousavi and Young, 2023). Accordingly, in terms of the guidelines listed by the Subcommittee, we provided evidence of intra-species variation (where new species included multiple strains) with genomes of sufficient quality and completeness, documented phenotype information and demonstrated that members of the new species are capable of nodulating *V. karroo*, thereby satisfying Koch's postulates (de Lajudie et al., 2019). This study represents an initial endeavour to formally name the rhizobial symbionts of *V. karroo* that would otherwise be left in limbo under the ICNP's current stance on culture availability and significantly contributes to the understanding of the intricate interaction between *V. karroo* and a prominent rhizobial partner, *Mesorhizobium*. Additionally, by formally naming the symbionts of *V. karroo* using the SeqCode, this work constitutes and illustrative example of how genome sequence data can serve as nomenclatural types for naming bacteria in countries where CBD- and NP-related legislation precludes compliance with the ICNP.

Table 9
Phenotypic traits for the species delineated in this study.

Phenotypic trait[a]	*Mesorhizobium* species[b]							
	vac. (n = 2)	*hum.* (n = 2)[c]	*aus.* (n = 2)	*mon.* (n = 1)	*dul.* (n = 3)	*cap.* (n = 5)	*alb.* (n = 1)	*aby.* (n = 2)
Temperature growth range (°C) [α]	4–37	4–35	4–35	15–37	10–40	4–35	15–40	10–40
Salinity growth range (% w/v NaCl) [α]	0.3–2.5	0.3–2.5	0.3–2.5	0.5–1.0	0.3–2.5	0.3–2.5	0.3–2.5	0.3–2.5
pH growth range [α]	5–10	6–9	5–10	6–10	6–9	6–10	6–9	5–10
Nitrate reduction ($NO_3 \rightarrow NO_2$)	+	−	−	+	v	−	+	v
Nitrite reduction ($NO_2 \rightarrow N_2$)	−	−	−	−	v	v	+	−
Hydrolysis of:								
L-arginine, urea	v	v	v	+	v	v	+	v
Esculin ferric citrate, gelatin	v	v	v	−	v	v	+	v
Utilization of:								
4-Nitrophenyl-β, D-galactopyranoside, D-glucose	+	+	+	+	+	v	+	v
L-arabinose, potassium gluconate, malic acid, phenylacetic acid	v	v	v	−	v	v	+	v
D-Mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, captric acid, adipic acid, trisodium citrate	v	v	v	−	v	v	−	v
Carbon source: Sugars								
Dextrin, D-maltose, D-trehalose, D-cellobiose, gentiobiose, sucrose, D-turanose, stachyose	+	+	+	+	v	v	+	+
D-Raffinose, α-D-lactose, D-melibiose, β-methyl-D-glucoside	v	+	+	+	v	v	+	+
D-Salicin, N-acetyl-D-glucosamine, N-acetyl-β-D-mannosamine. N-acetyl-D-galactosamine, N-acetyl neuraminic acid	v	+	+	+	v	v	−	+
α-D-Glucose, D-mannose, D-fructose, D-galactose, 3-methyl glucose, D-fucose, L-fucose, L-rhamnose, inosine	v	+	+	+	v	v	+	+
Carbon sources: carboxylic acids, esters and fatty acids								
p-Hydroxy-phenylacetic acid, propionic acid, formic acid,	v	−	v	+	v	v	−	v
Tween 40, α-hydroxy-butyric acid, β-hydroxy-D, L-butyric acid, acetoacetic acid, acetic acid	v	+	+	+	+	v	+	+
Methyl pyruvate, D-lactic acid methyl ester, L-lactic acid, D-malic acid, L-malic acid, bromo-succinic acid	+	+	+	+	v	v	+	+
Citric aicd, α-keto-glutaric acid, γ-amino-butyric acid	v	+	+	+	v	v	−	v
α-Keto-glutaric acid	−	−	−	−	−	−	−	−
Carbon source: Amino acids								
Glycyl-L-proline, L-alanine, L-aspartic acid, L-glutamic acid, L-pyroglutamic acid	v	+	+	+	v	v	+	+
L-arginine, L-histidine, L-serine	v	−	v	+	v	v	+	+
Carbon source: Hexose-PO_4								
D-Glucose-6-PO_4, D-Fructose-6-PO_4	v	+	+	+	+	+	−	+
Carbon source: Hexose acids								
Pectin, D-galacturonic acid, L-galactonic acid lactone	v	+	+	+	v	v	+	+
D-Gluconic acid, D-glucuronic acid, glucuronamida, mucic acid, quinic acid, D-saccharic acid	v	−	v	+	v	v	+	+
Other carbon sources:								
D-sorbitol, D-mannitol, Myo-inositol, glycerol	+	+	+	+	+	+	−	+
D-Arabitol, D-aspartic acid, D-serine	v	−	v	+	v	v	−	v

[a] Phenotypic traits were determined and the results are indicated as follows; positive (+), negative (−) or variable (v) if all strains of a species did not display the same result, while [α] demonstrates minimum and maximum ranges of physiological growth characteristics. For the full repertoire of physiological characters see Supplementary Table S4, and for the biochemical attributes of the strains see Supplementary Table S5 and S6.

[b] Species are abbreviated as follows: *vac.* = *M. vachelliae* sp. nov., *hum.* = *M. humile* sp. nov., *aus.* = *M. australafricanum* sp. nov., *mon.* = *M. montanum* sp. nov., *dul.* = *M. dulcispinae* sp. nov., *cap.* = *M. captivum* sp. nov., and *alb.* = *M. album* sp. nov. The column labelled "*aby*." includes the results for strains of *M. abyssinicae* recovered in this study.

[c] In light of the minuscule size of the bacterial colonies on culture media and the insufficient availability of bacterial cells in solution, we opted to omit the Biolog GenIII Microplate™ results for VK2B *of M. humile*. Consequently, we relied solely on the Biolog GenIII Microplate™ results acquired from VK2D within the species cluster for informative analysis.

Funding

The authors would like to thank the National Research Foundation (NRF) through the Centre of Excellence in Plant Health Biotechnology (CPHB) and the University of Pretoria for student funding.

CRediT authorship contribution statement

Melandré van Lill: Formal analysis, Investigation, Writing - Original Draft, Visualisation, Conceptualization. **Stephanus N. Venter:** Writing – review & editing, Supervision, Resources, Conceptualization. **Esther K. Muema:** Resources, Writing – review & editing. **Marike Palmer:** Writing – review & editing. **Wai Y. Chan:** Software. **Chrizelle W. Beukes:** Writing – review & editing, Supervision, Conceptualization. **Emma T. Steenkamp:** Writing – review & editing, Supervision, Resources, 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

The authors would like to acknowledge the Centre of Excellence in Plant Health Biotechnology (CPHB) at the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, for the utilization of their infrastructure.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.

org/10.1016/j.syapm.2024.126504.

References

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. https://doi.org/10.1016/S0022-2836(05)80360-2.

Arkin, A.P., Cottingham, R.W., Henry, C.S., Harris, N.L., Stevens, R.L., Maslov, S., Dehal, P., Ware, D., Perez, F., Canon, S., Sneddon, M.W., Henderson, M.L., Riehl, W. J., Murphy-Olson, D., Chan, S.Y., Kamimura, R.T., Kumari, S., Drake, M.M., Brettin, T.S., Glass, E.M., Chivian, D., Gunter, D., Weston, D.J., Allen, B.H., Baumohl, J., Best, A.A., Bowen, B., Brenner, S.E., Bun, C.C., Chandonia, J.-M., Chia, J.-M., Colasanti, R., Conrad, N., Davis, J.J., Davison, B.H., DeJongh, M., Devoid, S., Dietrich, E., Dubchak, I., Edirisinghe, J.N., Fang, G., Faria, J.P., Frybarger, P.M., Gerlach, W., Gerstein, M., Greiner, A., Gurtowski, J., Haun, H.L., He, F., Jain, R., Joachimiak, M.P., Keegan, K.P., Kondo, S., Kumar, V., Land, M.L., Meyer, F., Mills, M., Novichkov, P.S., Oh, T., Olsen, G.J., Olson, R., Parrello, B., Pasternak, S., Pearson, E., Poon, S.S., Price, G.A., Ramakrishnan, S., Ranjan, P., Ronald, P.C., Schatz, M.C., Seaver, S.M.D., Shukla, M., Sutormin, R.A., Syed, M.H., Thomason, J., Tintle, N.L., Wang, D., Xia, F., Yoo, H., Yoo, S., Yu, D., 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. Nat. Biotechnol. 36, 566–569. https://doi.org/10.1038/nbt.4163.

Aserse, A.A., Räsänen, L.A., Aseffa, F., Hailemariam, A., Lindström, K., 2013. Diversity of sporadic symbionts and nonsymbiotic endophytic bacteria isolated from nodules of woody, shrub, and food legumes in Ethiopia. Appl. Microbiol. Biotechnol. 97, 10117–10134. https://doi.org/10.1007/s00253-013-5248-4.

Avontuur, J.R., Palmer, M., Beukes, C.W., Chan, W.Y., Tasiya, T., van Zyl, E., Coetzee, M. P.A., Stepkowski, T., Venter, S.N., Steenkamp, E.T., 2022. *Bradyrhizobium altum* sp. nov., *Bradyrhizobium oropedii* sp. nov. and *Bradyrhizobium acaciae* sp. nov. from South Africa show locally restricted and pantropical *nodA* phylogeographic patterns. Molecular Phylogenetics and Evolution 167, 107338. 10.1016/j.ympev.2021.107338.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Beukes, C.W., Venter, S.N., Law, I.J., Phalane, F.L., Steenkamp, E.T., 2013. South african papilionoid legumes are nodulated by diverse *burkholderia* with unique nodulation and nitrogen-fixation loci. PLoS One 8, e68406.

Beukes, C.W., Boshoff, F.S., Phalane, F.L., Hassen, A.I., Le Roux, M.M., Stępkowski, T., Venter, S.N., Steenkamp, E.T., 2019a. Both alpha- and Beta-rhizobia occupy the root nodules of *Vachellia karroo* in South Africa. Front. Microbiol. 10 https://doi.org/10.3389/fmicb.2019.01195.

Beukes, C.W., Steenkamp, E.T., Van Zyl, E., Avontuur, J., Chan, W.Y., Hassen, A.I., Palmer, M., Mthombeni, L.S., Phalane, F.L., Sereme, T.K., 2019b. *Paraburkholderia strydomiana* sp. nov. and *Paraburkholderia steynii* sp. nov.: rhizobial symbionts of the fynbos legume *Hypocalyptus sophoroides*. Antonie Van Leeuwenhoek 112, 1369–1385. https://doi.org/10.1007/s10482-019-01269-5.

Brady, C.L., Cleenwerck, I., Venter, S.N., Engelbeen, K., De Vos, P., Coutinho, T.A., 2010. Emended description of the genus *Pantoea*, description of four species from human clinical samples, *Pantoea septica* sp. nov., *Pantoea eucrina* sp. nov., *Pantoea brenneri* sp. nov. and *Pantoea conspicua* sp. nov., and transfer of *Pectobacterium cypripedii* (Hori 1911) Brenner et al. 1973 emend. Hauben et al. 1998 to the genus as *Pantoea cypripedii* comb. nov.. International journal of systematic and evolutionary microbiology 60, 2430–2440. 10.1099/ijs.0.017301-0.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 421. https://doi.org/10.1186/1471-2105-10-421.

Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dreano, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D., Pohl, T., Portetelle, D., Puhler, A., Purnelle, B., Ramsperger, U., Renard, C., Thebault, P., Vandenbol, M., Weidner, S., Galibert, F., 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proceedings of the National Academy of Sciences 98, 9877–9882. 10.1073/pnas.161294398.

Chan, P.P., Lowe, T.M., 2019. tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Methods in molecular biology (Clifton, N.J.) 1962, 1–14. 10.1007/978-1-4939-9173-0_1.

Chan, P.P., Lin, B.Y., Mak, A.J., Lowe, T.M., 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. 49, 9077–9096. https://doi.org/10.1093/nar/gkab688.

Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., Parks, D.H., 2022. GTDB-tk v2: memory friendly classification with the genome taxonomy database. Bioinformatics 38, 5315–5316. https://doi.org/10.1093/bioinformatics/btac672.

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D.R., da Costa, M.S., Rooney, A. P., Yi, H., Xu, X.-W., De Meyer, S., 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int. J. Syst. Evol. Microbiol. 68, 461–466. https://doi.org/10.1099/ijsem.0.002516.

Chuvochina, M., Rinke, C., Mussig, A., Chaumeil, P.-A., Parks, D., Hugenholtz, P., 2023. Genome taxonomy database and SeqCode: microbial taxonomy and nomenclature in the age of big sequence data. Biodiversity Information Science and Standards 7. https://doi.org/10.3897/biss.7.111135.

Claassens, R., Venter, S.N., Beukes, C.W., Stępkowski, T., Chan, W.Y., Steenkamp, E.T., 2023. *Bradyrhizobium xenonodulans* sp. nov. isolated from nodules of australian *Acacia* species invasive to South Africa. Syst. Appl. Microbiol. 46, 126452 https://doi.org/10.1016/j.syapm.2023.126452.

Cleenwerck, I., Vandemeulebroecke, K., Janssens, D., Swings, J., 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. https://doi.org/10.1099/ijs.0.02064-0.

Cowan, D.A., Valverde, A., Wingfield, M.J., Rybicki, E.P., Tuffin, M.I., 2013. Biodiversity : so much more than legs and leaves. S. Afr. J. Sci. 109, 1–9. https://doi.org/10.10520/EJC145692.

Cowan, D.A., Lebre, P.H., Amon, C.E.R., Becker, R.W., Boga, H.I., Boulangé, A., Chiyaka, T.L., Coetzee, T., de Jager, P.C., Dikinya, O., Eckardt, F., Greve, M., Harris, M.A., Hopkins, D.W., Houngnandan, H.B., Houngnandan, P., Jordaan, K., Kaimoyo, E., Kambura, A.K., Kamgan-Nkuekam, G., Makhalanyane, T.P., Maggs-Kölling, G., Marais, E., Mondlane, H., Nghalipo, E., Olivier, B.W., Ortiz, M., Pertierra, L.R., Ramond, J.-B., Seely, M., Sithole-Niang, I., Valverde, A., Varliero, G., Vikram, S., Wall, D.H., Zeze, A., 2022. Biogeographical survey of soil microbiomes across sub-Saharan Africa: structure, drivers, and predicted climate-driven changes. Microbiome 10, 131. https://doi.org/10.1186/s40168-022-01297-w.

Da Silva, M., Desmeth, P., Venter, S.N., Shouche, Y., Yurkov, A., 2023. How legislations affect new taxonomic descriptions. Trends Microbiol. 31, 111–114. https://doi.org/10.1016/j.tim.2022.10.010.

David Cooper, H., Noonan-Mooney, K., 2013. Convention on biological diversity. Encyclopedia of Biodiversity: Second Edition 306–319. https://doi.org/10.1016/B978-0-12-384719-5.00418-4.

Davies-Barnard, T., Friedlingstein, P., 2020. The global distribution of biological nitrogen fixation in terrestrial natural ecosystems. e2019GB006387 Global Biogeochem. Cycles 34. https://doi.org/10.1029/2019GB006387.

de Lajudie, P.M., Andrews, M., Ardley, J., Eardly, B., Jumas-Bilak, E., Kuzmanović, N., Lassalle, F., Lindström, K., Mhamdi, R., Martínez-Romero, E., Moulin, L., Mousavi, S. A., Nesme, X., Peix, A., Puławska, J., Steenkamp, E., Stępkowski, T., Tian, C.F., Vinuesa, P., Wei, G., Willems, A., Zilli, J., Young, P., 2019. Minimal standards for the description of new genera and species of rhizobia and agrobacteria. Int. J. Syst. Evol. Microbiol. 69, 1852–1863. https://doi.org/10.1099/ijsem.0.003426.

de Lajudie, P., Willems, A., Nick, G., Moreira, F., Molouba, F., Hoste, B., Torck, U., Neyra, M., Collins, M.D., Lindström, K., Dreyfus, B., Gillis, M., 1998. Characterization of tropical tree rhizobia and description of *mesorhizobium plurifarium* sp. nov. Int. J. Syst. Bacteriol. 48 Pt 2, 369–382. https://doi.org/10.1099/00207713-48-2-369.

Degefu, T., Wolde-meskel, E., Frostegård, Å., 2011. Multilocus sequence analyses reveal several unnamed *mesorhizobium* genospecies nodulating *Acacia* species and *Sesbania sesban* trees in southern regions of Ethiopia. Syst. Appl. Microbiol. 34, 216–226. https://doi.org/10.1016/j.syapm.2010.09.006.

Degefu, T., Wolde-Meskel, E., Liu, B., Cleenwerck, I., Willems, A., Frostegård, Å., 2013. *Mesorhizobium shonense* sp. nov., *mesorhizobium hawassense* sp. nov. and *mesorhizobium abyssinicae* sp. nov., isolated from root nodules of different agroforestry legume trees. Int. J. Syst. Evol. Microbiol. 63, 1746–1753. https://doi.org/10.1099/ijs.0.044032-0.

Dingaan, M., du Preez, P.J., 2017. Vachellia (acacia) karroo communities in South Africa: an overview. IntechOpen London, UK. https://doi.org/10.5772/intechopen.70456.

Diouf, F., Diouf, D., Klonowska, A., Le Queré, A., Bakhoum, N., Fall, D., Neyra, M., Parrinello, H., Diouf, M., Ndoye, I., Moulin, L., 2015. Genetic and genomic diversity studies of *Acacia* symbionts in Senegal reveal new species of *mesorhizobium* with a putative geographical pattern. PLoS One 10, e0117667.

Dludlu, M.N., Chimphango, S.B.M., Stirton, C.H., Muasya, A.M., 2018. Differential preference of *burkholderia* and *mesorhizobium* to pH and soil types in the Core cape subregion. South Africa. Genes 9. https://doi.org/10.3390/genes9010002.

Driver A., Sink, K.J., Nel, J.L., Holness, S., Van Niekerk, L., Daniels, F., Jonas, Z., Majiedt, P.A., Harris, L. & Maze, K. 2012. National Biodiversity Assessment 2011: An assessment of South Africa's biodiversity and ecosystems. Synthesis Report. South African National Biodiversity Institute and Department of Environmental Affairs, Pretoria.

Eddie, B.J., Wang, Z., Hervey, J.W., Leary, D.H., Malanoski, A.P., Tender, L.M., Lin, B., Strycharz-Glaven, S.M., 2017. Metatranscriptomics Supports the Mechanism for Biocathode Electroautotrophy by "Candidatus *Tenderia electrophaga*." mSystems 2, 10.1128/msystems.00002-17. 10.1128/msystems.00002-17.

Field, D., Garrity, G., Gray, T., Morrison, N., Selengut, J., Sterk, P., Tatusova, T., Thomson, N., Allen, M.J., Angiuoli, S.V., Ashburner, M., Axelrod, N., Baldauf, S., Ballard, S., Boore, J., Cochrane, G., Cole, J., Dawyndt, P., De Vos, P., DePamphilis, C., Edwards, R., Faruque, N., Feldman, R., Gilbert, J., Gilna, P., Glöckner, F.O., Goldstein, P., Guralnick, R., Haft, D., Hancock, D., Hermjakob, H., Hertz-Fowler, C., Hugenholtz, P., Joint, I., Kagan, L., Kane, M., Kennedy, J., Kowalchuk, G., Kottmann, R., Kolker, E., Kravitz, S., Kyrpides, N., Leebens-Mack, J., Lewis, S.E., Li, K., Lister, A.L., Lord, P., Maltsev, N., Markowitz, V., Martiny, J., Methe, B., Mizrachi, I., Moxon, R., Nelson, K., Parkhill, J., Proctor, L., White, O., Sansone, S.-A., Spiers, A., Stevens, R., Swift, P., Taylor, C., Tateno, Y., Tett, A., Turner, S., Ussery, D., Vaughan, B., Ward, N., Whetzel, T., San Gil, I., Wilson, G., Wipat, A., 2008. The minimum information about a genome sequence (MIGS) specification. Nat. Biotechnol. 26, 541–547. https://doi.org/10.1038/nbt1360.

Gao, J.L., Xue, J., Sun, Y.C., Xue, H., Wang, E.T., Yan, H., Tong, S., Wang, L.W., Zhang, X., Sun, J.-G., 2020. *Mesorhizobium rhizophilum* sp. nov., a 1-aminocyclopropane-1-carboxylate deaminase producing bacterium isolated from rhizosphere of maize in Northeast China. Antonie Van Leeuwenhoek 113, 1179–1189. https://doi.org/10.1007/s10482-020-01425-2.

Gaunt, M.W., Turner, S.L., Rigottier-Gois, L., Lloyd-Macgilp, S.A., Young, J.P., 2001. Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of

rhizobia. Int. J. Syst. Evol. Microbiol. 51, 2037–2048. https://doi.org/10.1099/00207713-51-6-2037.
Geddes, B.A., Kearsley, J., Morton, R., diCenzo, G.C., Finan, T.M., 2020. The genomes of rhizobia, Advances in Botanical Research. Elsevier Ltd. 10.1016/bs.abr.2019.09.014.
Gerding, M., O'Hara, G.W., Bräu, L., Nandasena, K., Howieson, J.G., 2012. Diverse *mesorhizobium* spp. with unique *nodA* nodulating the south african legume species of the genus *Lessertia*. Plant and Soil 358, 385–401. https://doi.org/10.1007/s11104-012-1153-3.
Godfray, H.C.J., Knapp, S., Oren, A., 2004. Prokaryote diversity and taxonomy: current status and future challenges. philosophical transactions of the Royal Society of London. Series b: Biological Sciences 359, 623–638. https://doi.org/10.1098/rstb.2003.1458.
Göker, M., Moore, E.R.B., Oren, A., Trujillo, M.E., 2022. Status of the SeqCode in the international journal of systematic and Evolutionary microbiology. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/ijsem.0.005754.
Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., Tiedje, J. M., 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. 57, 81–91. https://doi.org/10.1099/ijs.0.64483-0.
Gosselin, S., Fullmer, M.S., Feng, Y., Gogarten, J.P., 2022. Improving phylogenies based on average nucleotide identity, incorporating saturation correction and nonparametric bootstrap support. Syst. Biol. 71, 396–409. https://doi.org/10.1093/sysbio/syab060.
Greenlon, A., Chang, P.L., Damtew, Z.M., Muleta, A., Carrasquilla-Garcia, N., Kim, D., Nguyen, H.P., Suryawanshi, V., Krieg, C.P., Yadav, S.K., Patel, J.S., Mukherjee, A., Udupa, S., Benjelloun, I., Thami-Alami, I., Yasin, M., Patil, B., Singh, S., Sarma, B.K., Von Wettberg, E.J.B., Kahraman, A., Bukun, B., Assefa, F., Tesfaye, K., Fikre, A., Cook, D.R., 2019. Global-level population genomics reveals differential effects of geography and phylogeny on horizontal gene transfer in soil bacteria. Proceedings of the National Academy of Sciences of the United States of America 116, 15200–15209. 10.1073/pnas.1900056116.
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, 307–321. https://doi.org/10.1093/sysbio/syq010.
Guo, D., Arnolds, J.L., 2018. Biodiversity in South Africa. Selected Countries in Africa. CRC Press, Global Biodiversity.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Syposium Series 41, 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. 10.17159/sajs.2021/11765.
Hassen, A.I., Swanevelder, Z.H., Bopape, F.L., Lamprecht, C.S., 2020. Draft Genome Sequence of *Mesorhizobium* sp. Strain SARCCRB16n, an Effective Nodulating and Nitrogen-Fixing Symbiont of Rooibos [*Aspalathus linearis* (Burm. f.)] in South Africa. Microbiology resource announcements 16–17. 10.1128/mra.01187-19.
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., 2022. SeqCode: a nomenclatural code for prokaryotes described from sequence data. Nat. Microbiol. 7, 1702–1708. https://doi.org/10.1038/s41564-022-01214-9.
Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.
Hördt, A., López, M.G., Meier-Kolthoff, J.P., Schleuning, M., Weinhold, L.-M., Tindall, B. J., Gronow, S., Kyrpides, N.C., Woyke, T., Göker, M., 2020. Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of alphaproteobacteria. Front. Microbiol. 11 https://doi.org/10.3389/fmicb.2020.00468.
Howieson, J.G., Dilworth, M.J., 2016. Working with rhizobia. Australian centre for international agricultural research Canberra, Australia.
Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. 11, 119. https://doi.org/10.1186/1471-2105-11-119.
Jarvis, B.D.W., Tighe, S.W., 1994. Rapid identification of *Rhizobium* species based on cellular fatty acid analysis. In: Graham, P.H., Sadowsky, M.J., Vance, C.P. (Eds.), Symbiotic Nitrogen Fixation. Springer, Dordrecht. https://doi.org/10.1007/978-94-011-1088-4_4.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.
Katoh, K., Rozewicki, J., Yamada, K.D., 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. 20, 1160–1166. https://doi.org/10.1093/bib/bbx108.
Kim, M., Kim, W., Park, W., 2022. *Aquibium microcysteis* gen. nov., sp. nov., isolated from a *Microcystis aeruginosa* culture and reclassification of *mesorhizobium carbonis* as *aquibium carbonis* comb. nov. and *mesorhizobium oceanicum* as *aquibium oceanicum* comb. nov. Int. J. Syst. Evol. Microbiol. 72, 1–9. https://doi.org/10.1099/ijsem.0.005230.
Kim, J., Na, S.I., Kim, D., Chun, J., 2021. UBCG2: up-to-date bacterial core genes and pipeline for phylogenomic analysis. J. Microbiol. 59, 609–615. https://doi.org/10.1007/s12275-021-1231-4.
Knight, J., Flack-Davison, E., Engelbrecht, S., Visagie, R.G., Beukes, C.W., Mwale, M., 2023. Compliance with the Nagoya protocol in the southern african development community. S. Afr. J. Bot. 159, 302–318. https://doi.org/10.1016/j.sajb.2023.06.015.
Konstantinidis, K.T., Rosselló-Móra, R., 2015. Classifying the uncultivated microbial majority: a place for metagenomic data in the *candidatus* proposal. Syst. Appl. Microbiol. 38, 223–230. https://doi.org/10.1016/j.syapm.2015.01.001.
Konstantinidis, K.T., Tiedje, J.M., 2005. Genomic insights that advance the species definition for prokaryotes. Proc. Natl. Acad. Sci. 102, 2567–2572. https://doi.org/10.1073/pnas.0409727102.
Kück, P., Longo, G.C., 2014. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front. Zool. 11, 81. https://doi.org/10.1186/s12983-014-0081-x.
Kumar, D., Kumar, G., Uppada, J., Ahmed, S., Sasikala, C., Venkata Ramana, C., 2021. Descriptions of *roseiconus nitratireducens* gen. nov. sp. nov. and *roseiconus lacunae* sp. nov. Arch. Microbiol. 203, 741–754. https://doi.org/10.1007/s00203-020-02078-5.
Laranjo, M., Alexandre, A., Rivas, R., Velázquez, E., Young, J.P.W., Oliveira, S., 2008. Chickpea rhizobia symbiosis genes are highly conserved across multiple *mesorhizobium* species. FEMS Microbiol. Ecol. 66, 391–400. https://doi.org/10.1111/j.1574-6941.2008.00584.x.
Laranjo, M., Oliveira, S., 2011. Tolerance of *mesorhizobium* type strains to different environmental stresses. Antonie Van Leeuwenhoek 99, 651–662. https://doi.org/10.1007/s10482-010-9539-9.
Laranjo, M., Alexandre, A., Oliveira, S., 2014. Legume growth-promoting rhizobia: an overview on the *mesorhizobium* genus. Microbiol. Res. 169, 2–17. https://doi.org/10.1016/j.micres.2013.09.012.
Lemaire, B., Dlodlo, O., Chimphango, S., Stirton, C., Schrire, B., Boatwright, J.S., Honnay, O., Smets, E., Sprent, J., James, E.K., Muasya, A.M., 2015. Symbiotic diversity, specificity and distribution of rhizobia in native legumes of the Core cape subregion (South Africa). FEMS Microbiol. Ecol. 91, 1–17. https://doi.org/10.1093/femsec/fiu024.
León-Barrios, M., Flores-Félix, J.D., Pérez-Yépez, J., Ramirez-Bahena, M.H., Pulido-Suárez, L., Igual, J.M., Velázquez, E., Peix, A., 2021. Definition of the novel symbiovar canariense within mesorhizobium neociceri sp. nov., a new species of genus mesorhizobium nodulating Cicer canariense in the "caldera de taburiente" National Park (La Palma, Canary Islands). Syst. Appl. Microbiol. 44, 126237 https://doi.org/10.1016/j.syapm.2021.126237.
Letunic, I., Bork, P., 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296. https://doi.org/10.1093/nar/gkab301.
Lewis, M., 2012. Access and benefit-sharing in the wake of CBD COP10: the Nagoya protocol and its potential implications for South Africa. South African Journal of Environmental Law and Policy 17.
Lin, S.-Y., Hameed, A., Hsieh, Y.-T., Young, C.-C., 2019. *Mesorhizobium composti* sp. nov., isolated from compost. Antonie Van Leeuwenhoek 112, 1387–1398. https://doi.org/10.1007/s10482-019-01270-y.
Ma, T., Xue, H., Piao, C., Jiang, N., Li, Y., 2023. Phylogenomic reappraisal of the family *rhizobiaceae* at the genus and species levels, including the description of *ectorhizobium quercum* gen. nov., sp. nov. Front. Microbiol. 14, 1207256. https://doi.org/10.3389/fmicb.2023.1207256.
Marcos-García, M., Menéndez, E., Ramírez-Bahena, M.H., Mateos, P.F., Peix, Á., Velazquez, E., Rivas, R., 2017. *Mesorhizobium helmanticense* sp. nov., isolated from *Lotus corniculatus* nodules. Int. J. Syst. Evol. Microbiol. 67, 2301–2305. https://doi.org/10.1099/ijsem.0.001942.
Maroyi, A., 2017. *Acacia karroo* hayne: ethnomedicinal uses, phytochemistry and pharmacology of an important medicinal plant in southern Africa. Asian Pac. J. Trop. Med. 10, 351–360. https://doi.org/10.1016/j.apjtm.2017.03.021.
Martínez-Hidalgo, P., Hirsch, A.M., 2017. The nodule microbiome: N_2-fixing rhizobia do not live alone. Phytobiomes Journal 1, 70–82. https://doi.org/10.1094/PBIOMES-12-16-0019-RVW.
Mavima, L., Beukes, C.W., Palmer, M., De Meyer, S.E., James, E.K., Maluk, M., Gross, E., dos Reis Junior, F.B., Avontuur, J.R., Chan, W.Y., 2021. *Paraburkholderia youngii* sp. nov. and '*Paraburkholderia atlantica*'–brazilian and mexican Mimosa-associated rhizobia that were previously known as *Paraburkholderia tuberum* sv. *mimosae*. Syst. Appl. Microbiol. 44, 126152 https://doi.org/10.1016/j.syapm.2020.126152.
Mousavi, S.A. and Young, J.P.W., 2023. International Committee on Systematics of Prokaryotes Subcommittee on the Taxonomy of Rhizobia and Agrobacteria. Minutes of the closed annual meeting: video conference on 11 October 2022 followed by online discussion until 31 December 2022. International Journal of Systematic and Evolutionary Microbiology, 73(4), p.005856. 10.1099/ijsem.0.005856.
Muema, E.K., Steenkamp, E.T., Venter, S.N., 2022. Rhizosphere diazotrophs and other bacteria associated with native and encroaching legumes in the succulent Karoo biome in South Africa. Microorganisms 10, 216. https://doi.org/10.3390/microorganisms10020216.
Murray, A.E., Freudenstein, J., Gribaldo, S., Hatzenpichler, R., Hugenholtz, P., Kämpfer, P., Konstantinidis, K.T., Lane, C.E., Papke, R.T., Parks, D.H., 2020. Roadmap for naming uncultivated archaea and bacteria. Nat. Microbiol. 5, 987–994. https://doi.org/10.1038/s41564-020-0733-x.
Na, S.I., Kim, Y.O., Yoon, S.H., Ha, S., Baek, I., Chun, J., 2018. UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. J. Microbiol. 56, 280–285. https://doi.org/10.1007/s12275-018-8014-6.
Nandasena, K.G., O'Hara, G.W., Tiwari, R.P., Willems, A., Howieson, J.G., 2009. *Mesorhizobium australicum* sp. nov. and *mesorhizobium opportunistum* sp. nov., isolated from *Biserrula pelecinus* L. in Australia. Int. J. Syst. Evol. Microbiol. 59, 2140–2147. https://doi.org/10.1099/ijs.0.005728-0.
Ndlovu, J., Richardson, D.M., Wilson, J.R.U., Le Roux, J.J., 2013. Co-invasion of south african ecosystems by an australian legume and its rhizobial symbionts. J. Biogeogr. 40, 1240–1251. https://doi.org/10.1111/jbi.12091.

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. https://doi.org/10.1093/molbev/msu300.

Nguyen, V.H., Wemheuer, B., Song, W., Bennett, H., Webster, N., Thomas, T., 2023. Identification, classification, and functional characterization of novel sponge-associated acidimicrobiial species. Syst. Appl. Microbiol. 46, 126426 https://doi.org/10.1016/j.syapm.2023.126426.

Oren, A., Arahal, D., Göker, M., Moore, E., Rossello-Mora, R., Sutcliffe, I. (eds), 2023. International Code of Nomenclature of Prokaryotes (2022 revision). International Journal of Systematic and Evolutionary Microbiology. 10.1099/ijs.0.2008/001636.

Ormeño-Orrillo, E., Martínez-Romero, E., 2013. Phenotypic tests in Rhizobium species description: an opinion and (a sympatric speciation) hypothesis. Syst. Appl. Microbiol. 36, 145–147. https://doi.org/10.1016/j.syapm.2012.11.009.

Ormeño-Orrillo, E., Martínez-Romero, E., 2019. A genomotaxonomy view of the *bradyrhizobium* genus. Front. Microbiol. 10, 1334. https://doi.org/10.3389/fmicb.2019.01334.

Ormeño-Orrillo, E., Servín-Garcidueñas, L.E., Rogel, M.A., González, V., Peralta, H., Mora, J., Martínez-Romero, J., Martínez-Romero, E., 2015. Taxonomy of rhizobia and agrobacteria from the *Rhizobiaceae* family in light of genomics. Syst. Appl. Microbiol. 38, 287–291. https://doi.org/10.1016/j.syapm.2014.12.002.

Pallen, M.J., Telatin, A., Oren, A., 2021. The next million names for archaea and bacteria. Trends Microbiol. 29, 289–298. https://doi.org/10.1016/j.tim.2020.10.009.

Palmer, M., Steenkamp, E.T., Blom, J., Hedlund, B.P., Venter, S.N., 2020. All ANIs are not created equal: implications for prokaryotic species boundaries and integration of ANIs into polyphasic taxonomy. Int. J. Syst. Evol. Microbiol. 70, 2937–2948. https://doi.org/10.1099/ijsem.0.004124.

Palmer, M., Sutcliffe, I., Venter, S.N., Hedlund, B.P., 2022. It is time for a new type of type to facilitate naming the microbial world. New Microbes and New Infections 47. https://doi.org/10.1016/j.nmni.2022.100991.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W., 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25, 1043–1055. https://doi.org/10.1101/gr.186072.114.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., Hugenholtz, P., 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat. Biotechnol. 36, 996–1004. https://doi.org/10.1038/nbt.4229.

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. https://doi.org/10.1093/nar/gkab776.

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. https://doi.org/10.1099/ijsem.0.004332.

Pedron, R., Luchi, E., Albiac, M.A., Di Cagno, R., Catorci, D., Esposito, A., Bianconi, I., Losa, D., Cristofolini, M., Guella, G., Jousson, O., 2021. *Mesorhizobium comanense* sp. nov., isolated from groundwater. International journal of systematic and evolutionary microbiology 71. 10.1099/ijsem.0.005131.

Peix, A., Ramírez-Bahena, M.H., Velázquez, E., Bedmar, E.J., 2015. Bacterial associations with legumes. Crit. Rev. Plant Sci. 34 (1–3), 17–42. https://doi.org/10.1080/07352689.2014.897899.

Poole, P., Ramachandran, V., Terpolilli, J., 2018. Rhizobia: from saprophytes to endosymbionts. Nat. Rev. Microbiol. 16, 291–303. https://doi.org/10.1038/nrmicro.2017.171.

Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2 – approximately maximum-likelihood trees for Large alignments. PLoS One 5, e9490.

Rahi, P., 2021. Regulating access can restrict participation in reporting new species and taxa. Nat. Microbiol. 6, 1469–1470. https://doi.org/10.1038/s41564-021-01002-x.

Rahimlou, S., Bahram, M., Tedersoo, L., 2021. Phylogenomics reveals the evolution of root nodulating alpha- and beta-proteobacteria (rhizobia). Microbiol. Res. 250, 126788 https://doi.org/10.1016/j.micres.2021.126788.

Ramos, B.S., 2021. The Nagoya Protocol on Access To Genetic Resources and the Fair and Equitable Sharing of Benefits Arising From Their Utilization and Its Implementation in Spain. Revista Electronica de Estudios Internacionales 2021. 10.17103/reei.42.04.

Richter, M., Rosselló-Móra, R., 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proceedings of the National Academy of Sciences 106, 19126–19131. 10.1073/pnas.0906412106.

Rodriguez-R, L.M., Gunturu, S., Harvey, W.T., Rosselló-Mora, R., Tiedje, J.M., Cole, J.R., Konstantinidis, K.T., 2018. The microbial genomes atlas (MiGA) webserver: taxonomic and gene diversity analysis of archaea and bacteria at the whole genome level. Nucleic Acids Res. 46, W282–W288. https://doi.org/10.1093/nar/gky467.

Rodriguez-R, L.M., Jain, C., Conrad, R.E., Aluru, S., Konstantinidis, K.T., 2021. Reply to: "re-evaluating the evidence for a universal genetic boundary among microbial species". Nat. Commun. 12, 4060. https://doi.org/10.1038/s41467-021-24129-1.

Soumare, A., Diedhiou, A.G., Thuita, M., Hafidi, M., Ouhdouch, Y., Gopalakrishnan, S., Kouisni, L., 2020. Exploiting biological nitrogen fixation: a route Towards a sustainable agriculture. Plants. https://doi.org/10.3390/plants9081011.

Steenkamp, E.T., van Zyl, E., Beukes, C.W., Avontuur, J.R., Chan, W.Y., Palmer, M., Mthombeni, L.S., Phalane, F.L., Sereme, T.K., Venter, S.N., 2015. *Burkholderia kirstenboschensis* sp. nov. nodulates papilionoid legumes indigenous to South Africa. Syst. Appl. Microbiol. 38, 545–554. https://doi.org/10.1016/j.syapm.2015.09.003.

Sun, Y., Buhler, J., 2007. Designing patterns for profile HMM search. Bioinformatics 23, 36–43. https://doi.org/10.1093/bioinformatics/btl323.

Sutcliffe, I.C., 2015. Challenging the anthropocentric emphasis on phenotypic testing in prokaryotic species descriptions: rip it up and start again. Front. Genet. https://doi.org/10.3389/fgene.2015.00218.

Sutcliffe, I.C., Rosselló-Móra, R., Trujillo, M.E., 2021. Addressing the sublime scale of the microbial world: reconciling an appreciation of microbial diversity with the need to describe species. New Microbes and New Infections 43, 100931. https://doi.org/10.1016/j.nmni.2021.100931.

Suzuki, M.T., Giovannoni, S.J., 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 62, 625–630. https://doi.org/10.1128/aem.62.2.625-630.1996.

Taylor, C.L., Barker, N.P., 2012. Species limits in Vachellia (acacia) karroo (mimosoideae: leguminoseae): evidence from automated ISSR DNA "fingerprinting". S. Afr. J. Bot. 83, 36–43. https://doi.org/10.1016/j.sajb.2012.07.014.

Tena, W., Wolde-Meskel, E., Degefu, T., Walley, F., 2017. Genetic and phenotypic diversity of rhizobia nodulating chickpea (Cicer arietinum L.) in soils from southern and central Ethiopia. Can. J. Microbiol. 63, 690–707. https://doi.org/10.1139/cjm-2016-0776.

Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235. https://doi.org/10.1093/nar/gkw256.

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. https://doi.org/10.1007/s10482-017-0869-8.

Vincent, J.M., 1970. A manual for the practical study of the root-nodule bacteria. Published by Blackwell Scientific Pub, International Biological Program.

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. https://doi.org/10.1016/j.ympev.2004.08.020.

Viver, T., Conrad, R.E., Lucio, M., Harir, M., Urdiain, M., Gago, J.F., Suárez-Suárez, A., Bustos-Caparros, E., Sanchez-Martinez, R., Mayol, E., 2023. Description of two cultivated and two uncultivated new *Salinibacter* species, one named following the rules of the bacteriological code: *Salinibacter grassmerensis* sp. nov.; and three named following the rules of the SeqCode: *Salinibacter pepae* sp. nov., *Salinibacter abyssi* sp. nov., and *Salinibacter pampae* sp. nov. Systematic and Applied Microbiology 46, 126416. 10.1016/j.syapm.2023.126416.

William, S., Feil, H., Copeland, A., 2012. Bacterial genomic DNA isolation using CTAB. Joint Genome Institute Sigma 50, 6876. http://my.jgi.doe.gov/general/protocols/DNA_Isolation_Bacterial_CTAB_Protocol.doc.

Yoon, S.-H., Ha, S.-M., Lim, J., Kwon, S., Chun, J., 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110, 1281–1286. https://doi.org/10.1007/s10482-017-0844-4.

Young, J.M., Kuykendall, L.D., Martínez-Romero, E., Kerr, A., Sawada, H., 2001. A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi, R. undicola* and *R. vitis*. International Journal of Systematic and Evolutionary Microbiology 51, 89–103. 10.1099/00207713-51-1-89.

Young, J.P.W., Moeskjær, S., Afonin, A., Rahi, P., Maluk, M., James, E.K., Cavassim, M.I. A., Rashid, M.-H.-O., Aserse, A.A., Perry, B.J., 2021. Defining the *Rhizobium leguminosarum* species complex. Genes 12, 111. https://doi.org/10.3390/genes12010111.

Zhao, C.T., Wang, E.T., Zhang, Y.M., Chen, W.F., Sui, X.H., Chen, W.X., Liu, H.C., Zhang, X.X., 2012. *Mesorhizobium silamurunense* sp. nov., isolated from root nodules of *Astragalus* species. Int. J. Syst. Evol. Microbiol. 62, 2180–2186. https://doi.org/10.1099/ijs.0.031229-0.

Zhu, Y.J., Lu, J.K., Chen, Y.L., Wang, S.K., Sui, X.H., Kang, L.H., 2015. *Mesorhizobium acaciae* sp. nov., isolated from root nodules of *Acacia melanoxylon* R. Br. International Journal of Systematic and Evolutionary Microbiology 65, 3558. https://doi.org/10.1099/ijsem.0.000455.