Delineation of *Paraburkholderia tuberum* sensu stricto and description of *Paraburkholderia podalyriae* sp. nov. nodulating the South African legume *Podalyria calyptrata*

Lazarus Mavima[a], Chrizelle W. Beukes[a], Marike Palmer[b], Sofie E. De Meyer[c,d], Euan K. James[e], Marta Maluk[e], Muthama A. Muasya[f], Juanita R. Avontuur[a], Wai Yin Chan[a,g], Stephanus N. Venter[a,*], Emma T. Steenkamp[a]

[a]*Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa*

[b]*School of Life Sciences, University of Nevada Las Vegas, Las Vegas, NV, United States of America*

[c]*MALDIID Pty Ltd, Murdoch, Western Australia, Australia*

[d]*Laboratory of Microbiology, Department Biochemistry and Microbiology, Ghent University, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium*

[e]*The James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK*

[f] *Department of Biological Sciences, University of Cape Town, Cape Town, South Africa*

[g]*National Institute for Communicable Disease, National Health Laboratory Service, Johannesburg, South Africa*

* Corresponding Author:
S.N. Venter; Private Bag X20, Hatfield, 0028, South Africa;
e-mail: fanus.venter@up.ac.za; tel: + 27 (0) 12 420 4100

Declarations of interest: none

Abstract

Since the discovery of *Paraburkholderia tuberum*, an indigenous South African species and one of the first beta-rhizobia described, several other South African rhizobial *Paraburkholderia* species have been recognized. In this study, we investigate the taxonomic status of 31 rhizobial isolates from the root nodules of diverse South African legume hosts in the Core Cape Subregion, which were initially identified as *P. tuberum*. These isolates originate from the root nodules of genera in the Papilionoideae (e.g., *Cyclopia*, *Virgilia*, *Hypocalyptus*, *Podalyria* and *Aspalathus*) as well as *Vachellia karroo*, from the mimosoid clade in the subfamily Caesalpinioideae. Genealogical concordance analysis (using sequences of the loci *pab*, *gltB*, *rpoB*, *acnA* and 16S rRNA) was used to delineate the isolates into putative species clusters. Our approach separated the isolates into two distinct clusters. Clusters A included *P. tuberum*

STM678^T, the type strain of the species, suggesting that this monophyletic group represents *P. tuberum* sensu stricto. Cluster B grouped sister to *P. tuberum* and included isolates from the Paarl Rock Nature Reserve in the Western Cape Province. Average Nucleotide Identity (ANI) analysis further confirmed that isolates of Cluster A shared high genome similarity with *P. tuberum* STM678^T compared to Cluster B and other *Paraburkholderia* species. The cluster representing *P. tuberum* sensu stricto included isolates from multiple host species, while members of Cluster B associated with a single species of *Podalyria*, *P. calyptrata*. For this new taxon we accordingly propose the name *Paraburkholderia podalyriae* sp. nov., with the type strain WC7.3b^T (= LMG 31413^T; SARCC 750^T).

Keywords:
symbiovar, beta-rhizobia, symbiotic loci, *nodA*, *nifH*, genealogical concordance

Introduction

The Core Cape Subregion (CCR) represents a plant speciose area [69] that is recognized as a global hotspot for legume (Fabaceae) diversity [43,63]. After the Asteraceae, legumes constitute the second largest family of flowering plants in the region [69]. More than 760 legume species occur in the CCR, most of which belonging to the tribes Podalyrieae, Crotalarieae, Indigofereae and Psoraleeae in the subfamily Papilionoideae [67,69]. All legume species previously studied in the CCR engage in symbiosis with rhizobial bacteria that are capable of fixing atmospheric nitrogen in specialized plant organs known as nodules [63]. These bacteria are well-adapted to local conditions [12,17,51,63] which, in the case of the CCR, is a semi-arid climate with soils that are acidic, saline and nutrient-poor [69]. The nitrogen-fixing symbiosis thus allows legumes to thrive in the CCR's harsh conditions [25].

During the past few decades, diverse and apparently endemic bacterial assemblages have been documented to associate with CCR legumes [12,33,63]. Accordingly, the known diversity of rhizobia has been substantially expanded, especially within the so-called "beta-rhizobia" (i.e., rhizobia forming part of the class Betaproteobacteria, a taxon that was recently suggested to represent the new order "Betaproteobacteriales" or "Burkholderiales" [85,86]). In fact, *Paraburkholderia tuberum,* originally isolated from the root nodules of *Aspalathus carnosa*, a papilionoid species indigenous to the CCR [77,110], was one of the first beta-rhizobia described. Several more species of rhizobial *Paraburkholderia* were subsequently isolated from legumes endemic to the region. These include the validly published taxa *P. sprentiae* [32], *P.*

kirstenboschensis [104], *P. dipogonis* [66,101], *P. dilworthii* [30], *P. rhynchosiae* [29], *P. strydomiana* and *P. steynii* [14]. Several root-nodule endophytes that seem incapable of nodulation due to the absence of relevant symbiotic loci [32], have also been reported from the CCR. These loci encode nodulation (*nod*) and nitrogen fixation (*nif* and *fix*) genes needed for establishing the symbiosis with legumes [72]. Examples of species lacking these loci include *P. fynbosensis* and *P. graminis* isolated from root nodules of *Lebeckia ambigua* [32], and *P. aspalathi* isolated from *Aspalathus abietina* [73]. To date, the CCR, and in particular the Fynbos biome, have been the source of numerous other *Paraburkholderia* isolates, many of which likely represent novel species requiring formal description [12,63,64].

Paraburkholderia species are among the predominant symbionts of legumes in the CCR [33,63,64], where they mostly associate with legumes in the papilionoid tribes Podalyrieae, Crotalarieae and Indigofereae [12,33,41,63]. Rhizobial *Paraburkholderia* occurring in the CCR all encode symbiotic loci with evolutionary histories that differ markedly from those of *Paraburkholderia* from other parts of the world [12,17], which seems to be a common feature of rhizobial groups in the region [12,33,63,64]. The symbiotic loci of CCR *Paraburkholderia* share a more recent ancestry with those of so-called "alpha-rhizobia" (i.e., rhizobial species in the class Alphaproteobacteria) than *Paraburkholderia* species found outside this region [31]. Furthermore, recent evidence suggests that *Paraburkholderia* represent the ancestral rhizobial symbiont of CCR legumes, and that the members of this genus mostly co-evolved with their legume hosts [33].

Despite having been described almost two decades ago, the taxonomy of *P. tuberum* is still largely unresolved. Initially, several isolates from the Fynbos biome, along with isolates from the South American Caatinga and Cerrado biomes [17,64], were identified as *P. tuberum* or as representing *P. tuberum*-like strains, based on genealogies inferred from 16S ribosomal RNA (rRNA) and *recA* gene sequences [12,63,64]. Although several of these *P. tuberum*-like strains from South America were recently accommodated in two new species (i.e., *P. youngii* [74] and *P. atlantica* [90]), the taxonomic status of *P. tuberum* and strains resembling it from South Africa remains unclear. To complicate matters more, biovars or symbiovars have been introduced for distinguishing *P. tuberum* and *P. tuberum*-like strains from the different geographic locations and hosts [76,91]. For example, initially biovar (bv.) mimosa, later named symbiovar (sv.) atlantica, was proposed for *P. tuberum* or *P. tuberum*-like strains originating

from South America where they nodulate mimosoid legumes [76,91]. By contrast, those from the CCR were initially designated as bv. papilionoideae, because they appeared to nodulate papilionoid species. However, the nodulation ability of *P. tuberum* or *P. tuberum*-like strains from the CCR span diverse papilionoid legumes (e.g., *Cyclopia*, *Virgilia*, *Hypocalyptus*, *Amphithalea*, *Lebeckia*, *Rhynchosia*, *Podalyria* and *Aspalathus*), as well as the mimosoid legume *Vachellia karroo* (formerly *Acacia karroo*) [9,12,51,63,64], resulting in the later designation as sv. africana [91].

The goal of this study was therefore to resolve the taxonomy of *P. tuberum* or *P. tuberum*-like strains originating from the CCR and to determine how they fit within the symbiovar scheme developed by Paulitsch et al. [64], especially those isolates associated with the mimosoid legume, *V. karroo*. For this purpose, we used a polyphasic approach to study a collection of South African isolates that were previously designated as *P. tuberum* or *P. tuberum*-like. Potential species were delineated using genealogical concordance analysis based on DNA sequence information for five housekeeping loci [111]. Support for the putative species was then investigated with whole genome sequence comparisons and a range of phenotypic traits. For sv. assignment, we used phylogenetic analyses of two symbiotic loci for strains previously identified as *P. tuberum* and *P. tuberum*-like from both South Africa and South and Central America. Finally, the ability to nodulate mimosoid and papilionoid legumes was evaluated for each putative species and sv. delineated among our collection of CCR strains.

Materials and Methods

Rhizobial strains and culturing conditions

Thirty-one *Paraburkholderia* strains, isolated in previous studies [12,14,32,51,59,64], were used in this study (Table 1). Resuscitation from cryopreservation at -70 °C in 15% glycerol was achieved by streaking the bacteria on Yeast Mannitol Agar (YMA) medium followed by incubation at 28 °C for 3 to 5 days [52]. The same medium and conditions were used for routine cultivation of the bacteria. The type strain designated for the novel species proposed in this study (see below) was submitted to the Belgian Coordinated Collections of Microorganisms (BCCM; Ghent University, Belgium) and the South African Rhizobium Culture Collection (SARCC; Agricultural Research Council, Plant Health and Protection Institute, South Africa).

Table 1 *Paraburkholderia* trains which were investigated in this study, their hosts, geographical origins and nodulation test results.

Isolate [a]	Original plant host	Legume used in nodulation tests [b]			Location [c]	Reference
		Mimosa pudica	***Lebeckia ambigua***	***Vachellia karroo***		
Cluster B (*P. podalyriae* sp. nov.)						
WC7.3b^{T}*	*Podalyria calyptrata*	ND	E	ND	Paarl Rock Nature Reserve, WC	[12]
WC7.3d	*Podalyria calyptrata*	ND	ND	ND	Paarl Rock Nature Reserve, WC	[12]
WC7.3g	*Podalyria calyptrata*	ND	ND	ND	Paarl Rock Nature Reserve, WC	[12]
Cluster A (*P. tuberum* sensu stricto						
31.1*	*Vachellia karroo*	I	ND	E	Hermanus, WC	[14]
32	*Vachellia karroo*	I	E	E	Alexander Bay, WC	[14]
35.1	*Vachellia karroo*	ND	E	ND	Kamieskroon, NC	[14]
40	*Vachellia karroo*	I	E	E	Stellenbosch, WC	[14]
CI2	*Cyclopia intermedia*	N	E	ND	Dennehoek, WC	[12]
CI3	*Cyclopia intermedia*	N	ND	ND	Dennehoek, WC	[12]
Clong3	*Cyclopia longifolia*	N	ND	ND	Thornhill, WC	[12]
Cpub6	*Cyclopia pubescens*	N	ND	ND	Port Elizabeth, WC	[12]
HC1.1bd	*Hypocalyptus sophoroides*	N	E	ND	Old du Toit's Kloof Pass, WC	[12]
HC6.4b	*Hypocalyptus oxalidifolius*	N	E	ND	Fernkloof Nature Reserve, WC	[12]
Kb1A	*Virgilia oroboides*	N	E	ND	Kirstenbosch Botanical Gardens, WC	[12]
MM5384	*Podalyria sericea*	N	E	ND	Langebaan, WC	[64]
MM5477-R1	*Aspalathus callosa*	ND	ND	ND	Cape Point Nature Reserve, WC	[64]
MM5477-R3	*Aspalathus callosa*	N	E	ND	Cape Point Nature Reserve, WC	[64]
MM5482-R1	*Amphithalea ericifolia*	N	ND	ND	Cape Point Nature Reserve, WC	[64]
MM5482-R2	*Amphithalea ericifolia*	ND	ND	ND	Cape Point Nature Reserve, WC	[64]
MM5496-R1	*Aspalathus carnosa*	N	I	ND	Cape Point Nature Reserve, WC	[64]
MM6662-R1	*Rhynchosia capensis*	N	E	ND	Kogelberg Nature Reserve, WC	[64]
RAU6.4a	*Hypocalyptus oxalidifolius*	N	E	ND	Fernkloof Nature Reserve, WC	[12]
UCT2*	*Cyclopia genistoides*	ND	E	ND	Rein's Farms, WC	[12]
UCT31*	*Cyclopia sessiliflora*	N	E	ND	Grootvadersbosch, WC	[12]

Isolate [a]	Original plant host	Legume used in nodulation tests [b]			Location [c]	Reference
		Mimosa pudica	***Lebeckia ambigua***	***Vachellia karroo***		
UCT70	*Cyclopia maculata*	N	E	ND	Jonkershoek, WC	[12]
WSM4174	*Lebeckia ambigua*	N	ND	ND	Oorlogskloof Nature Reserve, NC	[51]
WSM4175	*Lebeckia ambigua*	N	ND	ND	Oorlogskloof Nature Reserve, NC	[51]
WSM4176*	*Lebeckia ambigua*	N	ND	ND	Oorlogskloof Nature Reserve, NC	[51]
WSM4177	*Lebeckia ambigua*	N	ND	ND	Oorlogskloof Nature Reserve, NC	[51]
WSM4179	*Lebeckia ambigua*	N	ND	ND	Oorlogskloof Nature Reserve, NC	[32]
WSM4180	*Lebeckia ambigua*	N	ND	ND	Oorlogskloof Nature Reserve, NC	[51]

[a] Strains indicated with * have whole genome sequences available.
[b] Results of the various nodulation tests are indicated as follows: E=Effective nodules were formed; I=Ineffective nodules were formed; N=No nodule formation; and ND=No Data (these strains where not included in the nodulation tests). All results are from tests done in this study.
[c] All isolates originate in South Africa. Abbreviations of the provinces are as follows: WC=Western Cape Province and NC=Northern Cape Province.

PCR and sequencing

From 3 to 5-day old YMA cultures, DNA was extracted using the Quick-gDNA™ MiniPrep kit (Zymo Research, USA) and quantified using the NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). These DNAs were then used as template to amplify five housekeeping loci that are variable in nature and found in the core genome of the broader *Burkholderia* sensu lato [13,38]. These included the 16S rRNA gene [107], as well as parts of the genes *pab* (encoding anthranilate synthase or para-amino-benzoate) [74], *gltB* (encoding glutamate synthase large subunit) [103], *rpoB* (encoding RNA polymerase subunit beta) [104] and *acnA* (encoding aconitate hydratase A) [74]. See Suppl. Table S1 for primer sequences and PCR cycling conditions. PCR was performed as described previously [74], after which amplicons were purified enzymatically. The latter involved incubation at 37 °C for 15 minutes in the presence 20 U/µl Exonuclease 1 (Thermo Fisher Scientific) and 1 U/µl of Alkaline phosphatase (Thermo Fisher Scientific), followed by inactivation of the enzymes by incubation at 80 °C for 15 minutes. Cleaned amplicons were then sequenced using the original PCR primers, the ABI PRISM Big Dye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems) and an ABI 377 Automated Capillary DNA Sequencer (Applied Biosystems, USA). ChromasLite v2.6.2 (Technelysium, Australia) and BioEdit v7.2.5 [47] were used to manually curate the DNA chromatograms and sequences.

For the delineation of putative species, portions of the protein coding genes *pab* (500 bp), *gltB* (650 bp), *rpoB* (1000 bp) and *acnA* (850 bp) were sequenced from all 31 isolates included in this study. However, the 16S rRNA gene was sequenced only for isolate HC1.1bd as its near full-length 16S rRNA gene sequence (1100 bp) was not available in GenBank. All sequences generated during this study were submitted to the European Nucleotide Archive (https://www.ebi.ac.uk) and are publicly available with the following accession numbers: LR694439 [16S rRNA gene], LR694315-LR694345 [*pab*], LR694346-LR694376 [*gltB*], LR694408-LR694438 [*rpoB*] and LR694134-LR694164 [*acnA*].

Datasets, taxon selection and phylogenetic analysis

Single gene nucleotide datasets for the five loci were compiled using the sequences generated during this study, as well as those previously determined [12,14,32,51,63,64]. The latter included the 16S rRNA gene sequences for most of the isolates, as well as the sequences for the loci that were extracted from the whole genomes of 31.1, UCT2, UCT31, WC7.3b^T and *P. tuberum* WSM4176 (see below; Suppl. Table 2). Our datasets also included sequences for *P.*

atlantica JPY251 [74], *P. atlantica* CCGE1002 [74,90] and *Parburkholderia* sp. H160 [82]. For comparative purposes, these datasets also contained sequences for the type strains of closely related *Paraburkholderia* species included in the List of Prokaryotic Names with Standing in Nomenclature (LPSN; www.bacterio.net [39,87,88]). The datasets accordingly included the sequences for all species with 16S rRNA sequences showing >95% blastn-based similarity [1] to those in the GenBank [11] database of the National Centre for Biotechnology Information (NCBI; https://blast.ncbi.nlm.nih.gov/Blast.cgi). For outgroup purposes, we used sequences for *P. kururiensis* JCM 10599^T [115], *P. caballeronis* LMG 26416^T [70], *P. mimosarum* LMG 23256^T [21] and *Caballeronia calidae* LMG 29321^T [36,93].

Datasets containing sequenced portions of genes *pab*, *gltB*, *rpoB* and *acnA* were manually aligned based on their inferred amino acid sequences using BioEdit. The 16S rRNA dataset was aligned with MAFFT (Multiple Alignment using Fast Fourier Transformation; https://mafft.cbrc.jp/alignment/server/) [56] by accounting for secondary structure using the Q-INS-i strategy [57]. These nucleotide datasets and their best-fit substitution models were subjected to maximum likelihood (ML) phylogenetic analyses in MEGA6 [109]. For all datasets, we used gamma correction (based on five discrete categories) to account for among-site rate variation [113]. The 16S rRNA dataset also included a proportion of invariable sites and used the Hasegawa-Kishino-Yano model (HKY; [48]). The Tamura 3-parameter model (T92; [108]) was employed for all of the protein-coding genes. After automatic generation of an initial tree with the BioNJ neighbor-joining algorithm, tree space was searched using the heuristic Nearest Neighbor Interchange (NNI) method based on a strong 'branch swap filter'. Branch support was estimated using bootstrap analysis of 1000 pseudoreplicates and the same model parameters [40].

For multilocus sequence analysis (MLSA), a partitioned nucleotide dataset was constructed from the five single gene alignments using FASconCAT-G [60]. This dataset was then subjected to ML analysis in IQ-TREE (2.0.6) [79], where each gene partition employed independent parameter estimation and the best-fit substitution model, as indicated by ModelFinder [55]. Branch support for the MLSA tree was estimated in IQ-TREE using these same parameters, together with ultrafast bootstrap (UFBoot) analysis of 1000 pseudoreplicates [49]. All phylogenetic trees were visualized with FigTree (1.4.4) (http://tree.bio.ed.ac.uk/) and edited in Inkscape (0.92) (https://www.inkscape.org).

For comparative purposes, we also constructed a phylogeny based on the nucleotide sequences for the gene encoding *recA*. Sequences were aligned with the AUTO strategy in MAFFT and then subjected to ML phylogenetic analysis in IQ-TREE as described above. Sequences for the analysis were obtained from GenBank or extracted from relevant genome available in NCBI. The dataset consisted of the published sequences for *Paraburkholderia* strains reported by Lemaire et al [65] and Zilli et al [116], and the available *recA* data for known *Paraburkholderia* species. We also included *recA* data that were extracted from the genome sequences of South African *P. tuberum* sensu lato generated here or that were previously determined.

Sv. identification for the taxa included in this study, was performed using two additional datasets containing sequences for the genes *nodA* (encoding N-acyltransferase needed for Nod factor formation) and *nifH* (encoding dinitrogenase reductase that is part of the nitrogenase enzyme complex) [6,20]. The taxon selection in these datasets were broad (Suppl. Table 2), because we attempted to include as many species and strains previously identified as belonging to bv. mimosae and papilionoideae proposed by Mishra et al. [76], and sv. atlantica, piptadeniae, tropicalis, mimosae and africana proposed by Paulitsch et al. [91]. The *nodA* and *nifH* gene sequences were translated in BioEdit, after which the inferred amino acid sequences were aligned using the AUTO strategy in MAFFT. The two alignments were then used for constructing ML trees in IQ-TREE as described above.

Whole genome sequence analyses

The genome sequences for *P. tuberum* WSM4176 and *P. tuberum* STM678^T were obtained from GenBank (Suppl. Table S2), while those for strains 31.1, UCT2, UCT31 and WC7.3b^T were determined during this study (Suppl. Table S2) and their assemblies have been deposited in GenBank with accession numbers GCA_014397645.1, GCA_014397665.1, GCA_014397655.1 and GCA_014397785.1. The sequencing was done by MicrobesNG (University of Birmingham, UK) according to Estrada-de los Santos et al. [38]. The obtained sequence reads were assembled with SPAdes (version 3.12.0) [10,80]. Whole genome assemblies produced were then used to evaluate their overall genome relatedness, as well as their relatedness to the genomes of close relatives. These included *P. sprentiae* WSM5005^T, *P. youngii* JPY169^T, *P. atlantica* CNPSo 3155^T, *P. atlantica* CCGE1002 and *Paraburkholderia* sp. H160. These analyses involved pair-wise estimation of Average Nucleotide Identity using the ANIb algorithm [3,44] in JSpecies [96]. Digital DNA-DNA hybridization (dDDH) [75] values were calculated with the genome-to-genome distance calculator (https://ggdc.dsmz.de/)

using strain WC7.3b^T as the reference. Finally, percentage G+C content was determined in JSpecies.

Phenotypic characterization

All strains representing the new species delineated here (see below) were subjected to analyses of cell morphology and motility, as well as growth ranges for temperature, pH and salinity as described previously [74]. Gram staining was performed according to O'Hara et al. [81]. The strains were also subjected to a range of biochemical tests using the Biolog GN2 MicroPlate^{TM} system (Biolog, USA) and API 20NE strips (BioMerieux, USA). These tests were carried out according to the manufacturer's instructions, with cultures incubated at 28 °C for three days. Additionally, the catalase test was performed using 3% hydrogen peroxide (https://www.asm.org/Protocols/Catalase-Test-Protocol; [95]), and the oxidase test was performed according to MacFaddin [68].

Nodulation tests

A set of 26 strains were chosen to test their ability to nodulate one or more of the mimosoid species *Mimosa pudica* and *V. karroo*, and the papilionoid species *L. ambigua* (Table 1). Nodulation tests for *M. pudica* and *V. karroo* were performed using the growth systems described by Bontemps et al. [17] and Beukes et al. [12], respectively. Nodulation tests performed with *L. ambigua* were accomplished using an axenic sand-culture system [114].

Results

Delineation of putative species

Comparisons of 16S rRNA gene sequences to those in the NCBI database showed that the 31 South African isolates shared high similarity with *P. tuberum* STM678^T (≥ 98.6%; [110]), *P. youngii* JPY169^T (≥ 98.7%; [74]) and *P. atlantica* CNPSo 3155^T (≥ 98.5%; [90]). They also shared 96.8 to 98.5% 16S rRNA gene sequence similarity with *P. ribeironis* STM7296^T [19], *P. ginsengisoli* NBRC100965^T [58], *P. susongensis* LMG29540^T [45], *P. monticola* JC2948^T [7], *P. sprentiae* WSM5005^T [28] and *Paraburkholderia* sp. H160 [82]. Their similarity to the outgroup taxa (*P. kururiensis* JCM 10599^T, *P. caballeronis* LMG 26416^T and *P. mimosarum* LMG 23256^T) were lower (i.e., 95.9-96.9%). Accordingly, all our datasets also included the sequences for these additional taxa.

Genealogies inferred from the five gene regions consistently grouped the 31 strains investigated, into two clusters (clusters A and B) (Fig. 1a-e). The type strain of *P. tuberum* (STM678^T) was included in Cluster A that contained 28 of our strains, with the remaining three included in Cluster B. Apart from Cluster A containing STM678^T, both of the clusters excluded other species or strains of *Paraburkholderia*. The same was also observed in the MLSA tree (Fig. 2), where Clusters A and B were fully supported (Fig. 2). This five-gene concatenated phylogeny also showed that Cluster B was more closely related to *P. sprentiae* WSM5005^T than to Cluster A containing *P. tuberum* STM678^T (Fig. 2).

Following the principles of genealogical concordance [111], the 31 isolates examined here likely represent two distinct species (Fig. 1a-e). The first species hypothesis is represented by Cluster A, which includes type strain STM678^T [110], likely indicating conspecificity with *P. tuberum.* It includes the six strains reported by Howieson et al [51] and De Meyer et al [32], 15 strains reported by Beukes et al [12,14] and seven strains reported by Lemaire et al [64]. The second species hypothesis is represented by Cluster B that contains strains WC7.3b^T, WC7.3d and WC7.3g, previously isolated from *Podalyria calyptrata* [12].

Phylogenetic analysis of *recA* was used to determine how Clusters A and B relate to the *Paraburkholderia* symbionts of *Podalyria calyptrata* reported by Lemaire et al [65] and symbionts of *Calliandra* species (mimosoid clade, subfamily Caesalpinioideae) reported by Zilli et al [116]. The results showed that some of the *P. calyptrata* symbionts appeared to be closely related to *P. tuberum*, *P. sprentiae*, *P. monticola* and *P. podalyriae* sp. nov., although none of them grouped with WC7.3b^T representing Cluster B (Suppl. Fig. 1). Also, none of the isolates from *Calliandra* grouped with either Cluster A or B. The *recA* data thus supported the uniqueness of Cluster B.

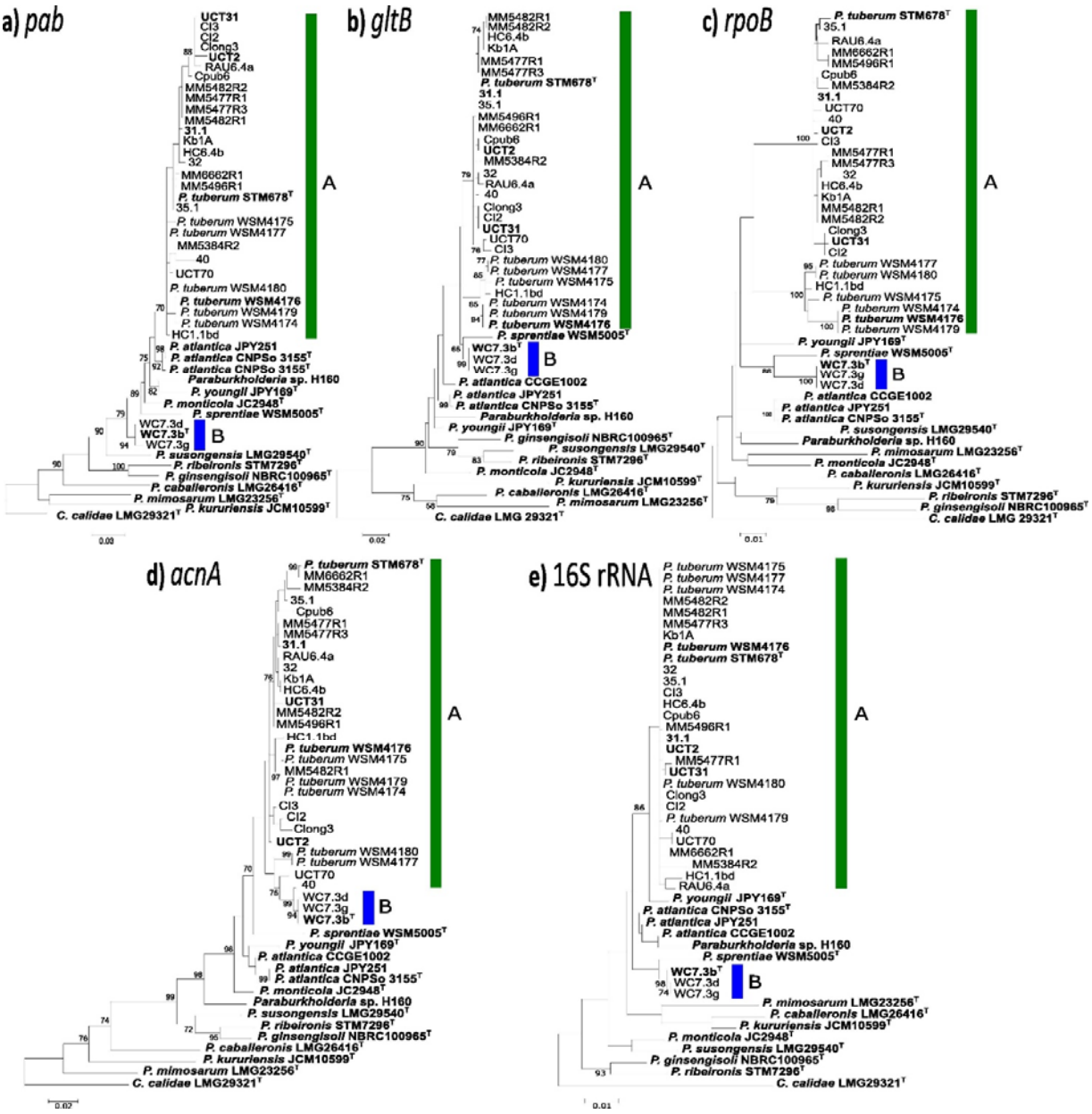

Fig. 1. Maximum likelihood phylogenies inferred from nucleotide sequence data for the 31 strains investigated and associated reference sequences. These included sequences for the four protein-coding genes *pab* (a), *gltB* (b), *rpoB* (c) and *acnA* (d) and the 16S rRNA gene (e). The scale bars indicate the number of nucleotide changes per site for each phylogeny, and bootstrap support values of >70% are indicated at the nodes. Isolates indicated in bold represent those with genome sequence data available. The DNA sequences for *C. calidae* LMG 29321^T [36,93], *P. kururiensis* JCM 10599^T [115], *P. caballeronis* LMG 26416^T [70] and *P. mimosarum* LMG 23256^T [21] were used for outgroup purposes, with the phylogenies rooted with *C. calidae* LMG 29321^T. While those for *P. ribeironis* STM 7296^T [19], *P. ginsengisoli* NBRC 100965^T [58], *P. susongensis* LMG 29540^T [42], *P. sprentiae* WSM5005^T [28], *P. tuberum* STM678^T [110], *P. youngii* JPY169^T [74], *P. atlantica* JPY251 [74], *P. atlantica* CCGE1002 [74], *P. atlantica* CNPSo 3155^T [90] and *Paraburkholderia* sp. H160 [112] were used for reference purposes. GenBank accession numbers are listed in (Suppl. Table S2).

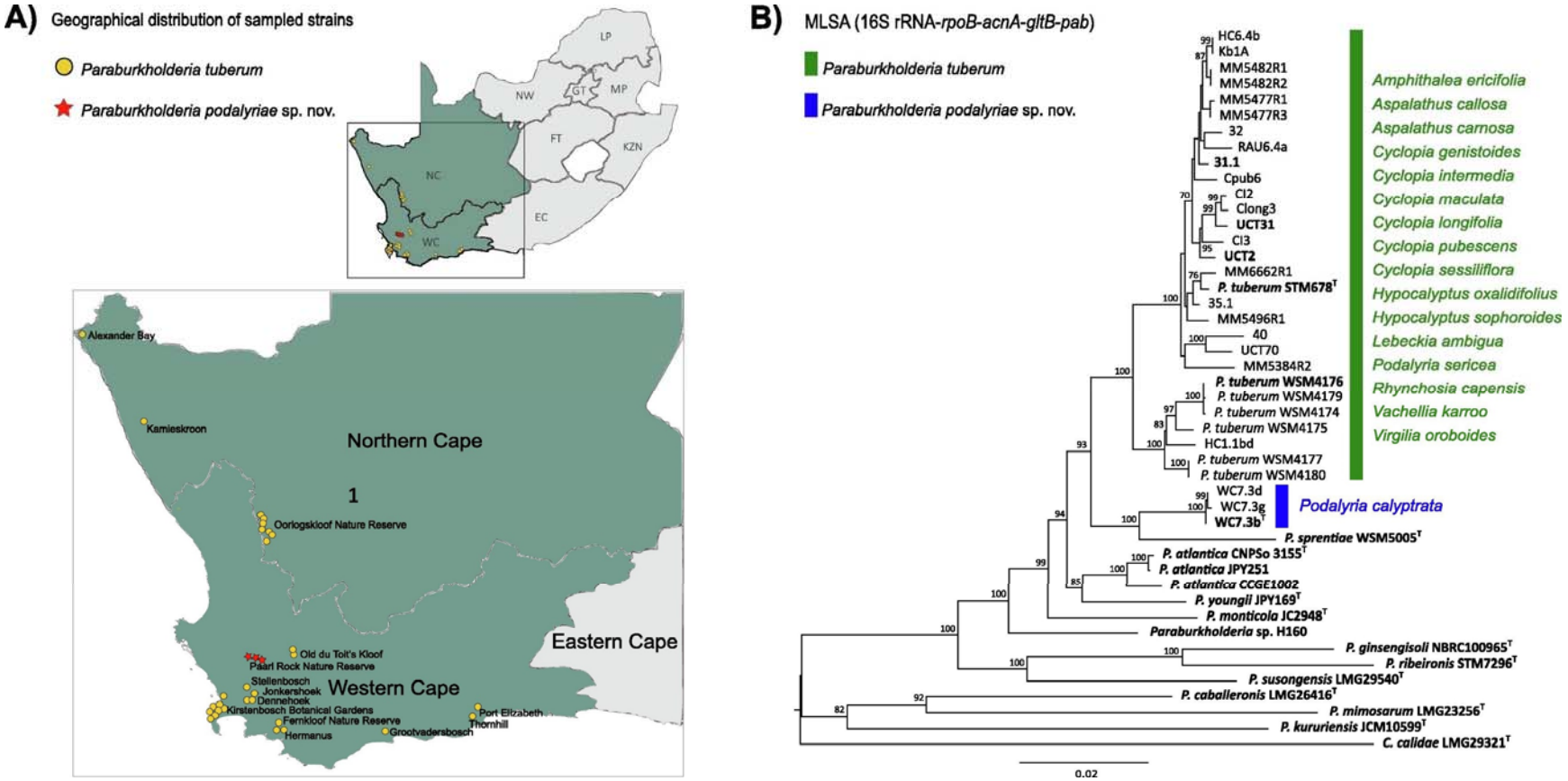

Fig. 2. Geographic origin, legume hosts and phylogeny of *P. tuberum* (Cluster A) and *P. podalyriae* sp. nov. (Cluster B). A) Map of the Western Cape and Northern Cape Provinces of South Africa showing the sampling locations for the two *Paraburkholderia* species. The map was drawn using DIVA-GIS 7.5.0 (https://www.diva-gis.org). B) An MLSA phylogeny based on the concatenated sequences for five genes (16S rRNA, *rpoB*, *acnA*, *gltB* and *pab*), showing the CCR legumes (indicated in green and blue) that the two species are known to nodulate. Bootstrap support values >60% are indicated at the nodes, and the scale bar indicates the number of nucleotide changes per site. Strains in bold have whole genome sequences available in NCBI.

Whole genome sequence-based support for putative species

The genome sequences for *Paraburkholderia* strains 31.1, UCT2, UCT31 and WC7.3b^T were determined in this study (Table 2). The sequencing depth for strains 31.1 and WC7.3b^T agreed with the most recently proposed minimal standard (≥ 50X) for using genome data in taxonomic studies [26,27], while the sequencing depth for strains UCT2 and UCT31 were 45X and 34X, respectively. The G+C content of these genomes ranged from 61.4 to 65.0 % (Table 2), which corresponds to what is known for other *Paraburkholderia* species [98]. Additionally, the sizes of the genomes broadly corresponded with those for other *Paraburkholderia*, with the representative of Cluster B (i.e., WC7.3b^T) having a genome of 9.36 Mb, and genome size for the representatives of Cluster A (i.e., 31.1, UCT2, UCT31, WSM4176 and *P. tuberum* STM678^T) ranging between 7.78-9.07 Mb (Table 2). By comparison, the genome of *P. sprentiae* strain WSM5005^T is 7.8 Mb in size and those of *P. atlantica* (i.e., CNPSo 3155^T, JPY303, JPY251, CCGE1002) range between 7.88-8.95 Mb [74], while the genomes of *P. youngii* strains (i.e., JPY169^T, JPY432, JPY454 and JPY418) range from 8.82 to 9.68 Mb [74].

All of the genomes sequenced in this study were also included in our ANI analyses (Fig. 3). The ANI-values obtained for representatives of Cluster A were all above 95.8%, which is in line with the suggested threshold of 95 to 96% ANI similarity for a bacterial species [3,44,96]. By contrast, all comparisons involving strain WC7.3b^T, the representative for Cluster B, produced ANI values below 92.8% (Fig. 3). The highest ANI values obtained for comparisons of strain WC7.3b^T with *P. tuberum* STM678^T, *P. sprentiae* WSM5005^T, *P. atlantica* CNPSo 3155^T, *P. youngii* JPY169^T and *Paraburkholderia* sp. H160 were 91.3%, 92.8%, 89.5%, 89.8% and 89.3%, respectively (Fig. 3). These ANI-values thus support recognition of our two clusters as separate species, and also confirm that Cluster B represents a new taxon. Similar conclusions were drawn from the inferred dDDH-values (Fig. 3). Consistent with the proposed dDDH guidelines [26], values well below 70% were obtained for comparisons involving strain WC7.3b^T (Fig. 3).

Table 2 Statistics for whole genomes of *Paraburkholderia* strains sequenced or used in this study.

Statistics	*P. podalyriae* sp. nov.	*P. tuberum* sensu stricto				
	WC7.3b^{T}	31.1	UCT2	UCT31	WSM4176*	STM678^{T}*
Sequencing coverage	229.55	184.44	34.77	44.65	ND	64
No. of contigs	382	297	199	238	66	222
Largest contig (bases)	514,144	254,651	377,118	288,179	ND	ND
N50**	162,521	136,771	134,479	128,747	430,419	118,037
GC content (%)	62.74	63.37	63.38	62.98	62.11	63.10
Genome size (Mb)	9.36	7.78	7.91	8.87	9.07	8.79
NCBI BioProject ID	PRJNA573588	PRJNA573578	PRJNA573582	PRJNA573585	PRJNA169686	PRJEB35318
Assembly Accession Number	GCA_014397785.1	GCA_014397645.1	GCA_014397665.1	GCA_014397655.1	GCA_000372945.1	GCA_902833905.1

* Strains whose genomes were downloaded from GenBank
** Length of the shortest contig that accumulatively show ≥50% the genome size [26].
ND=No data

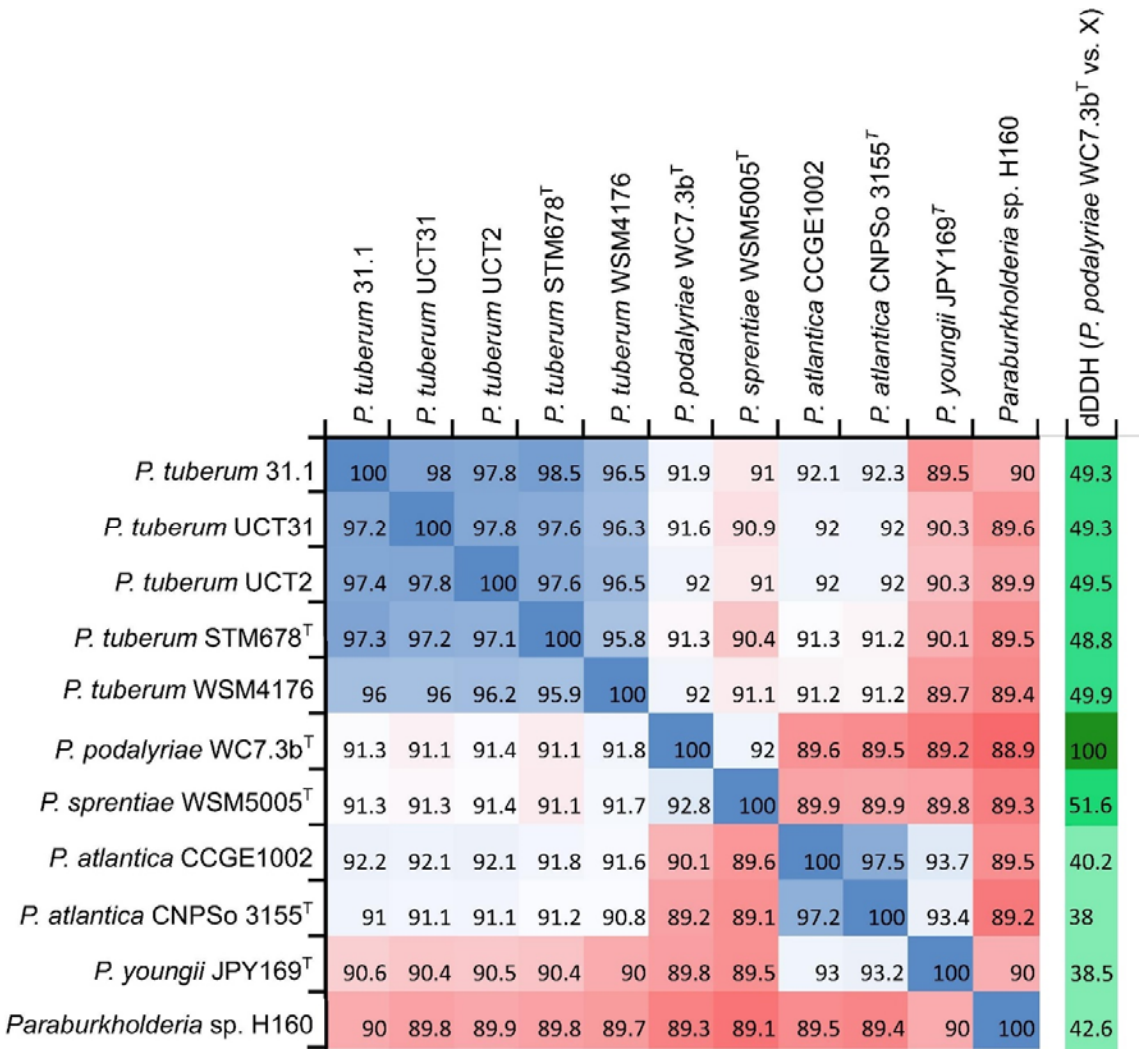

Fig. 3. Genome-based relatedness among *P. tuberum* sensu stricto (Cluster A represented by strains 31.1, UCT31, UCT2, WSM4176 and STM678^T), *P. podalyriae* sp. nov. (Cluster B represented by WC7.3b^T), and their closely related known *Paraburkholderia* species and strains. The matrix colour-coded in shades of blue and red represents the results of pair-wise ANIb analyses, while the column in shades of green represents dDDH-values obtained for comparisons with strain WC7.3b^T.

Phenotypic support for putative species

All of the strains in Cluster B were Gram-negative, motile, rod-shaped, and capable of growth at temperatures of 10 °C to 45 °C, across a pH range of 4 to 8 and with 0% to 1% of NaCl (Table 3, Suppl. Table S3 and Suppl. Fig 2). Based on the rate of colony or cell-biomass formation on YMA, the optimum conditions for growth of these isolates is at 25 °C to 37 °C, pH in the range of 5 to 7 and 0.1% to 0.5% of NaCl. Analysis with the Biolog GN2 MicroPlate^{TM} and API 20NE systems further revealed that Cluster B strains have many biochemical traits in common (e.g., the production of at least two enzymes, the assimilation of at least five carbon sources, and the fermentation of at least 44 different compounds) (Suppl. Table S3). Although, strains in Cluster A were not subjected to extensive phenotypic tests,

Table 3 Phenotypic characteristics of *Paraburkholderia podalyriae* sp. nov. compared to closely related *Paraburkholderia* species.

Phenotypic characteristics [a]	***Paraburkholderia* species** [b]										
	1	2 [#, ά, Θ]	3*	4*	5 [β, ψ, Θ]	6 [#, ϒ, Θ]	7 [#, Θ]	8 [#, ϒ, Θ]	9 [#, Θ]	10 [β, Θ]	11 [#, Θ]
Isolation source	RN	RN	RN	RN	RN	RN	RN	RN	RN	RN	RN
Nitrate reduction	+	-	+	+	-	V	+	+	+	+	+
Motility	+	+	+	+	+	ND	ND	ND	ND	+	ND
Growth conditions:											
Temperature (°C)	10-40	28	10-40	10-40	10-40	28,30,37	ND	ND	28	15-40	28,30,37
pH	4-8	ND	4-8	4-8	4.5-9	ND	ND	ND	ND	4-8	ND
Salinity (NaCl%)	0-1	ND	0-1	0-1	0-10	ND	ND	ND	ND	0-1	ND
Activity of:											
Arginine dihydrolase	-	-	-	-	+	-	-	ND	-	-	-
Tryptophan deaminase	-	-	-	-	w+	ND	+	ND	-	ND	ND
Urease	V	V	V	V	-	V	V	+	-	+	+
β-Galactosidase	+	+	+	+	+	-	V	+	+	+	+
β-Glucosidase	-	ND	-	-	ND	ND	ND	ND	ND	-	ND
Catalase	+	ND	+	+	+	+	-	ND	ND	+	+
Oxidase	+	+	+	+	+	+	+	ND	+	+	+
Assimilation of:											
D-Glucose	+	+	+	+	+	+	+	+	+	+	+
L-Arabinose	+	+	+	+	+	+	+	+	+	+	+
D-Mannose	+	+	+	+	+	V	+	+	+	+	+
D-Mannitol	+	+	+	+	+	+	+	+	+	+	+
N-Acetyl-Glucosamine	+	+	+	+	+	+	+	+	+	+	+
Potassium gluconate	+	+	+	+	+	V	+	-	+	ND	+
D-Maltose	-	V	-	-	ND	-	-	+	-	-	-
Adipic acid	-	V	V	V	-	-	+	+	-	-	-
Malic acid	-	V	-	V	+	+	+	+	-	+	+
Trisodium citrate	-	V	-	V	+	-	V	+	-	+	+
Capric acid	-	V	-	-	-	-	V	+	+	-	+
Fermentation of:											

Adonitol	+	V	w+	+	ND	-	+	+	+	+	+
L-Arabinose	+	+	+	+	-	+	+	+	+	+	+
D-arabitol	+	+	+	+	ND	+	+	+	+	+	+
D-cellobiose	-	V	w+	+	ND	-	V	+	-	+	+
D-fructose	+	+	+	+	ND	+	+	+	+	+	+
L-fucose	+	+	+	+	ND	+	+	+	+	ND	+
α-D-lactose	-	-	-	V	ND	-	-	-	-	-	-
Maltose	-	-	-	-	ND	-	-	-	-	ND	-
D-melibiose	-	-	-	-	-	-	-	-	+	-	-
D-raffinose	-	-	-	-	ND	-	-	-	-	+	-
L-rhamnose	+	+	+	+	-	-	+	+	+	+	+
D-Sorbitol	+	+	+	+	-	+	+	+	+	+	+
Sucrose	-	-	-	-	-	-	V	-	+	+	-
D-trehalose	-	-	-	w+	ND	-	+	+	-	+	+
Xylitol	-	V	-	+	ND	-	-	+	+	ND	-
D-Glucose	+	V	+	+	-	-	-	-	-	-	-
D-Mannitol	w+	V	+	w+	-	+	-	+	+	+	+
Inositol	w+	V	+	+	-	+	-	+	ND	ND	w+
Whole genome data											
G+C content (mol%)	62.7	62.8	61.22	63-63.3	61.6	64.8	61.8	62.8	62.1	63-65	64.5

[a] Origins of the phenotypic characteristics used for comparison: * Mavima et al. [74]; [#] Steenkamp et al. [104]; [ψ] Baek et al. [7]; [ϒ] Weber et al. [112]; [ά] De Meyer et al. [28]; [Θ] Mthombeni et al. [78]. RN, root nodule; RH, rhizosphere; +, all strains positive; w+, weakly positive; -, all strains negative; V, variable; ND, no data available.

[b] Isolates and strains are presented as follows: 1, *P. podalyriae* sp. nov. (n=3; this study); 2, *P. tuberum* (n=14) [78,110]; 3, *P. youngii* (n=19) [74]; 4, *P. atlantica* (n=4) [90]; 5, *P. sprentiae* (n=7) [28]; 6, *P. mimosarum* LMG 23256^T [21]; 7, *P. kirstenboschensis* (n=12) [104]; 8, *P. nodosa* LMG 23741^T [22]; 9, *P. phymatum* LMG 21445^T [110; 10, *P. diazotrophica* (n=5) [100]; 11, *P. sabiae* LMG 24235^T [23].

results for Biolog GN2 MicroPlate™ and API 20NE tests on 13 strains (WSM4180, WSM4176, HC6.4b, Kb1A, Cpub6, Clong3, CI2, CI3, UCT31, UCT2, UCT70, RAU6.4a and HC1.1bd) were available from a previous study [78]. As was the case for Cluster B, numerous of these biochemical traits were common among Cluster A strains (Suppl. Table S3), and these largely also corresponded with the traits reported for strain STM678^T [110]. However, the traits determined using Biolog GN2 MicroPlate™ and API 20NE systems generally lacked informativeness in distinguishing members of Cluster B from Cluster A and other closely related species (Table 3). Nevertheless, by being able to ferment D-glucose, members of Cluster B differed from its closest known relative, *P. sprentiae*. Cluster B strains also differed from *P. sprentiae* by being unable to assimilate malic acid and trisodium citrate, and their ability to ferment D-sorbitol.

Delineation of symbiovars

Although the *nodA* and *nifH* dataset differed in taxon selection (i.e., the *nifH* dataset also included free-living diazotrophs), both included strains that were designated as belonging to biovars *sensu* Mishra et al. [76] and symbiovars *sensu* Paulitsch et al. [91]. In both cases, we also included strains initially designated as *P. tuberum* or *P. tuberum*-like and later shown to represent *P. youngii* and *P. atlantica* [74], as well as all relevant South African *Paraburkholderia* species and strains for which the data were available. The resulting *nodA* dataset consisted of 195 aligned amino acid residues and 54 taxa, while the *nifH* dataset consisted of 185 aligned amino acid residues and 68 taxa.

ML analysis of the *nodA* and *nifH* datasets robustly recovered five clusters containing *Paraburkholderia* strains, which represented sv. atlantica, tropicalis, piptadeniae, mimosae and africana [91]. As expected, however, all strains previously suggested to be affiliated with *P. tuberum* were placed in one of two clusters [16,17]. One contained South African strains and the other contained strains from South and Central America (Fig. 4). The South African cluster included all strains previously designated as members of bv. papilionoideae *sensu* Mishra et al. [76], although the entire cluster subsequently has been renamed sv. africana [91]. This cluster has also now been shown to also contain South African symbionts associating with the mimosoid clade, and therefore no longer only associates with papilionoid hosts. The South and Central American cluster included known strains of *P. youngii* and *P. atlantica*, as well as two sets of strains for which exact species identities are not yet available. These are JPY807, JPY690, JPY682, JPY804 and JPY697 from the study by Bontemps et al. [18] and strains

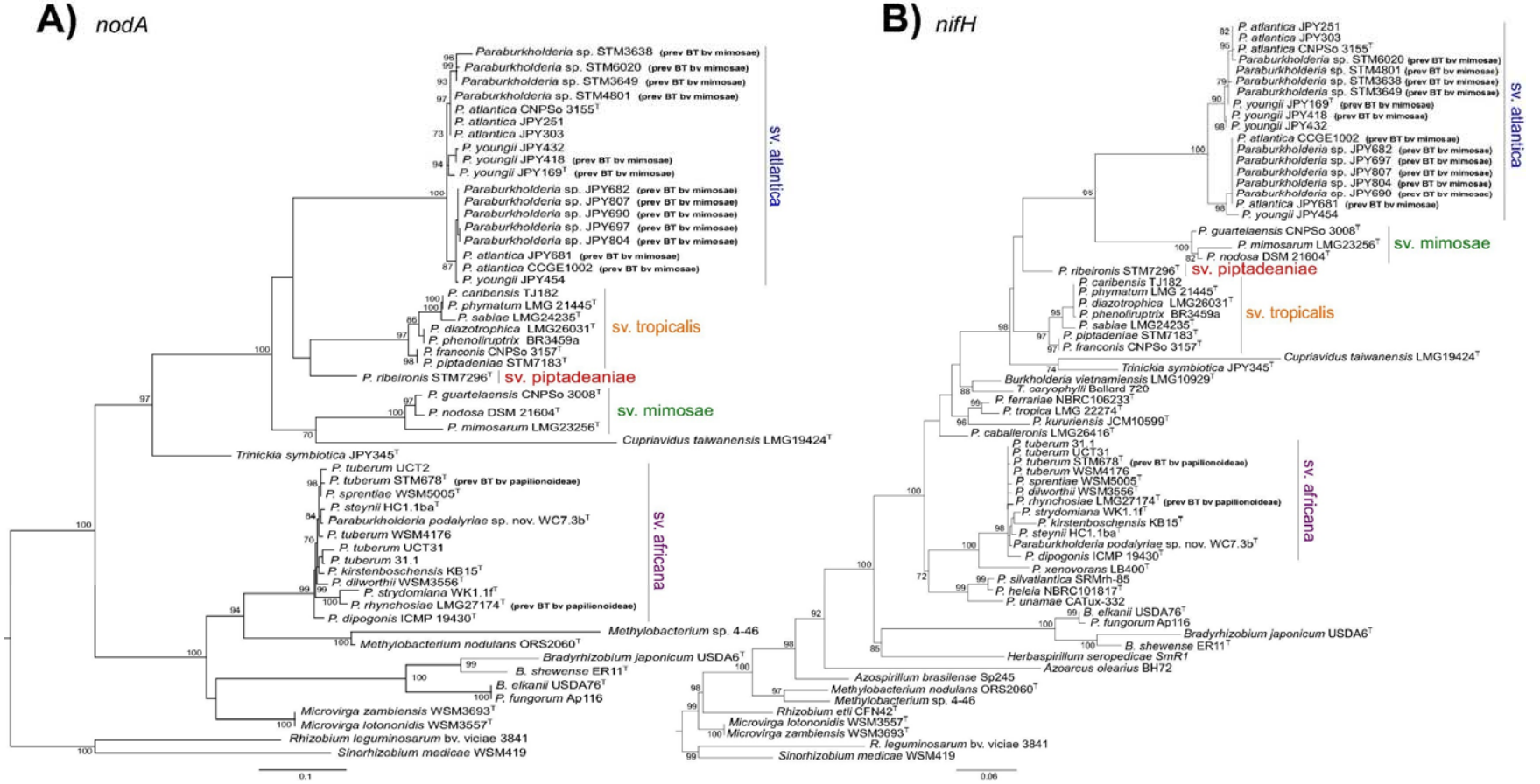

Fig. 4. Maximum likelihood phylogenies for *nodA* (A) and *nifH* (B) datasets showing the symbiovar (sv.) designations (*sensu* Paulitsch et al. [91]) for all strains and species of *Paraburkholderia* included. Strains previously classified as *P. tuberum* biovar mimosae and papilionoideae [76] are indicated as: prev BT bv mimosae and prev BT bv papilionoideae, respectively. The strains *Sinorhizobium medicae* WSM419 [97] and *Rhizobium leguminosarum* biovar viciae 3841 [50] were used for outgroup purposes. Bootstrap support values >70% are indicated at the nodes, and the scale bar indicates the number of nucleotide changes per site. GenBank accession numbers and references are listed in (Suppl. Table S2).

STM3638, STM6020, STM3649 and STM4801 from the study by Mishra et al. [76]. All of these, together with some of the *P. youngii* and *P. atlantica* strains were previously designated as members of bv. mimosae *sensu* Mishra et al. [76], but the assemblage has since been renamed sv. atlantica [91]. Therefore, following this system, the new species delineated here, together with *P. tuberum* sensu stricto, *P. sprentiae*, *P. steynii*, *P. kirstenboschensis*, *P. dilworthii*, *P. strydomiana*, *P. rhynchosiae* and *P. dipogonis* represent members of sv. africana (Fig. 4). Additionally, all South and Central American strains previously associated with *P. tuberum* represent members of sv. atlantica (Fig. 4).

Nodulation

Of the 15 Cluster A isolates tested, 14 were able to form effective nitrogen-fixing root nodules (i.e., with pink interiors suggestive of leghaemoglobin production) on *L. ambigua* (Table 1). Among the Cluster A strains, RAU6.4a is known to be capable of nodulating the papilionoid legumes *Hypocalyptus coluteoides*, *Vigna unguiculata*, and *Macroptilium atropurpureum*, while Kb1A can also nodulate the latter [12]. One of the Cluster B strains (WC7.3b^{T}) was also able to form effective nodules on *L. ambigua*, while another (WC7.3d) is known to be capable of nodulating *Ma. atropurpureum* [12]. However, none of the Cluster A strains tested could induce effective nodules on *M. pudica*, although ineffective nodules were induced by strains originating from *V. karroo* (i.e., 31.1, 32, 35.1 and 40). By comparison, when inoculated onto their original host, the latter strains (except for 35.1) were capable of effectively nodulating *V. karroo* (Table 1). Taken together, these data thus show that strains originating from hosts indigenous to the CCR (i.e., sv. africana strains) can nodulate various papilionoid legumes, and in some cases *V. karroo*, but not the South American native, *M. pudica*.

Discussion

Findings presented here represent an important milestone in the resolution of *P. tuberum* sensu lato. It follows on the recent work of Mavima et al. [74] and Paulitsch et al. [90] and ties up the remaining loose ends with respect to *P. tuberum* or *P. tuberum*-like strains occurring in South Africa and South and Central America. Where Mavima et al. [74] and Paulitsch et al. [90] showed that the strains from South and Central America represent distinct taxa and introduced the species *P. atlantica* and *P. youngii* to accommodate them, our study addressed the taxonomy of the strains initially identified as *P. tuberum* or *P. tuberum*-like from South Africa, particularly the CCR. We showed that they represent two distinct taxa, one representing *P. tuberum* sensu stricto and the other a new species for which we propose the name *P. podalyriae*

Table 4 Description of *Paraburkholderia podalyriae* sp. nov.

Genus name	*Paraburkholderia*
Species name	*Paraburkholderia podalyriae*
Specific epithet	*podalyriae*
Species status	sp. nov.
Species etymology	po.da.ly'ri.ae. N.L. fem. gen. n. *podalyriae*, of *Podalyria*, the host plant from which the species was isolated
Description of the new taxon and diagnostic traits	Cells are Gram-negative, motile, rod-shaped with an average length and width of 1.57 µm and 0.63 µm, respectively. Growth occurs on YMA enriched with 0 to 1% of NaCl, at a temperature of 10 °C to 45 °C (optimum 25 °C to 37 °C), and at a pH of 4.0 to 8.0 (optimum pH 5 to 7). Colonies are white to creamy, with circular or irregular form, convex or raised, and entire or curled margins. Isolates are positive for the activity of β-galactosidase. Positive for the reduction of nitrate to nitrite. Positive for the assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, potassium gluconate and N-acetyl-glucosamine. Positive for the hydrolysis of Tween 40, Tween 80, D-glucose, D-mannitol, D-sorbitol, I-arabinose, inositol, N-acetyl-D glucosamine, adonitol, L-arabinose, D-arabitol, D-fructose, L-fucose, D-galactose, D-mannose, L-rhamnose, D-sorbitol, pyruvic acid methyl ester, succinic acid mono-methyl ester, acetic acid, cis-aconitic acid, citric acid, D-galactonic acid, D-gluconic acid, glucosaminic acid, β-hydroxybutyric acid, D-L-lactic acid, propionic acid, quinic acid, D-saccharic acid, succinic acid, bromosuccinic acid, succinamic acid, D-alanine, L-alanine, asparagine, L-aspartic acid, L-glutamic acid, glycyl-L glutamic acid, L-histidine, hydroxy-L-proline, L-phenylalanine, L-proline, L-pyroglutamic acid, L-serine, L-threonine, urocanic acid and glycerol. Negative reactions are observed for the assimilation of D-maltose, capric, adipic acid, malic acid, trisodium citrate. Additionally, isolates are negative for the hydrolysis of xylitol, α-cyclodextrin, dextrin, glycogen, I-erythritol, gentiobiose, maltose, D-melibiose, D-cellobiose, β-methyl-D-glucoside, D-raffinose, sucrose, D-trehalose, turanose, α-hydroxybutyric acid, γ-hydroxybutyric acid, α-ketobutyric acid, α-ketovaleric, glycyl-L-aspartic acid, L-leucine L-ornithine, D-serine, D-L-carnitine, γ-aminobutyric acid, uridine, thymidine, phenylethyl-amine, 2-aminoethanol, 2-3-butanediol and α-D-glucose-1-phosphate.
Country of origin	South Africa
Region of origin	Western Cape Province
Source of isolation	Legume root nodules
Sampling date (dd/mm/yyyy)	01/09/2004
16S rRNA gene accession nr.	HF674696
Genome accession number	NCBI = GCA_014397785.1
Genome status	Complete
Genome size	9,362 Kbp
GC mol%	62.7
Number of strains in study	31
Source of isolation of non-type strains	Root nodules of *Podalyria calyptrata*
Designation of the Type Strain	WC7.3b^T
Strain Collection Numbers	WC7.3b^T = LMG 31413^T = SARCC 750^T

sp. nov. (see Table 4 for the protologue) Although a number of *P. tuberum* sensu lato strains still require taxonomic treatment (e.g., those reported from French Guiana as well as Mexico [18,76]), it should be relatively straightforward to accurately place them within the robust systematic framework for *P. tuberum* sensu lato that has emerged during the last two years.

Based on our findings, *P. tuberum* sensu stricto is a diverse taxon that can be encountered across the CCR. It can also associate with various papilionoid legumes native to southern Africa [16,32,51,64]. These include *Amphithalea* spp. [63,64], *Aspalathus* spp. [64,77], *Cyclopia* spp. [12,37,64], *Dipogon lignosus* [66], *Hypocalyptus* spp. [12,64], *Indigofera* spp. [64,65], *L. ambigua* [51,64], *Liparia* spp. [64], *Macroptilium* spp. [2,62,64], *Podalyria* spp. [64,65], *Rhynchosia* spp. [64,65], *Stirtonanthus taylorianus* [64], *Virgilia oroboides* [12,64] and *Xiphotheca fruticosa* [64]. The current study also demonstrated that *P. tuberum* sensu stricto associates with the southern African mimosoid *V. karroo* [14]. Notably, however, none of the strains originating from South and Central America formed part of *P. tuberum* sensu stricto. They were initially lumped with this species because identifications were based on 16S rRNA and *recA* gene sequences, but it is now known that neither gene (analyzed alone or together) contain sufficient phylogenetic signal to unambiguously resolve *P. tuberum* sensu lato and its lineages. Use of the species name '*tuberum*' for *Paraburkholderia* is thus discouraged until conspecificity with *P. tuberum* STM678^T can be established using multiple protein-coding genes or whole genome sequence information. We also suggest correction of erroneous species designations as '*B. tuberum*' for strains used in GenBank, to limit future confusion regarding *P. tuberum* sensu stricto and its diversity and distribution.

Paraburkholderia podalyriae sp. nov. represents a cohesive and exclusive new taxon that associates with the native South African papilionoid legume *P. calyptrata.* In the laboratory setting it can also form effective nodules on *Lebeckia* species, a trait that has so far been exhibited only by *Paraburkholderia* strains indigenous to southern Africa [51,64]. This points to the notion that *P. podalyriae* sp. nov. is native to the region, and that it might be a symbiont of other native legumes (particularly those growing under geographical or edaphic conditions similar that of the Paarl rock area), as interactions between South African papilionoids and native rhizobia are generally not highly specific [16,34,65]. Additionally, the discovery of *P. podalyriae* sp. nov. extends the number of nodulating *Paraburkholderia* species in the CCR to nine, with these species mainly from the Fynbos biome. These include *P. tuberum* sensu stricto

[110], *P. sprentiae* [28], *P. kirstenboschensis* [104], *P. dilworthii* [30], *P. rhynchosiae* [29], *P. dipogonis* [66,101], *P. strydomiana* [15] and *P. steynii* [15].

Although a robust species tree based on genome data for *Paraburkholderia* is still lacking, available MLSA data suggest that species belonging to *P. tuberum* sensu lato diverged relatively recently. The closest known relative of *P. podalyriae* sp. nov. is *P. sprentiae*, and together they share a most recent common ancestor with *P. tuberum* sensu stricto. The South and Central American species, *P. atlantica* and *P. youngii*, are each other's closest relatives, and they share a most recent common ancestor with the cluster containing the three South African species [74,90]. The distinct biogeographies within this five-species cluster thus appear to be the result of a single ancestral bacterium or lineage that "seeded" (via an as-yet-unknown mechanism) the respective regions with progenitors of the extant species. Furthermore, this ancestor did not necessarily have rhizobial or symbiotic properties, because the cluster corresponding to our five-species cluster were previously shown to be closely related to and nested among various non-symbiotic or environmental species such as *P. monticola* and *P. ginsengisoli* within the broader *Paraburkholderia* phylogeny [13]. During their subsequent divergence, the rhizobial *Paraburkholderia* species could thus have acquired their nodulation abilities via independent horizontal transfer events [31,91,92], as evident from nodulating gene phylogenies showing distinct continental patterns [64]. Our future research will seek to investigate the plausibility of this and hypotheses for explaining the geographic distribution of *P. podalyriae* sp. nov., *P. tuberum* sensu stricto, *P. sprentiae*, *P. atlantica* and *P. youngii*.

The continental split among members of the five-species cluster is reflected in their host affinities. For example, the two South and Central American species nodulate diverse mimosoid legumes (mainly *Mimosa* species), but some also nodulate the promiscuous papilionoids *Phaseolus vulgaris* and *Ma. atropurpureum* that are both native to the region [74,90,91,92]. It is also likely that these rhizobial species are capable of nodulating *Calliandra* species, as *Paraburkholderia* symbionts of this legume genus can nodulate the South America native species *Mimosa pudica* and carry *nod* genes that are closely related to those of *Mimosa* symbionts [102,116]. However, none of the *P. atlantica* and *P. youngii* strains studied by Mavima et al. [74] could nodulate the CCR native legume, *L. ambigua*. By comparison, none the strains tested for *P. podalyriae* sp. nov., *P. tuberum* sensu stricto, *P. sprentiae* can nodulate nodulate *M. pudica* [16,91]. However, they can nodulate various papilionoid species [16,29], which in the case of *P. tuberum* sensu stricto now also includes the South African mimosoid

legume *V. karroo*, as we show for the first time. This promiscuity is thus in-line with previous suggestions of reflecting adaptation to an environment characterized by high legume diversity [33].

The proposed hypothesis for the origin of the cluster containing *P. podalyriae* sp. nov., *P. tuberum* sensu stricto, *P. sprentiae*, *P. atlantica* and *P. youngii*, and the host specificities of these bacteria are congruent with the phylogenetic origins for their *nodA* and *nifH* genes. All South African species and strains of nodulating *Paraburkholderia*, including *P. podalyriae* sp. nov. and *P. tuberum* sensu stricto, encode *nodA* sequences that are apparently derived from those of alpha-rhizobia (i.e., *Methylobacterium*, *Bradyrhizobium* and *Microvirga*), while their *nifH* sequences are more closely related to that of the diazotroph *P. xenovorans*. By contrast, *nodA* and *nifH* sequences from *P. atlantica* and *P. youngii* group with those of nodulating *Paraburkholderia* from various other parts of the world, including South and Central America. These distribution patterns are similar to those reported previously [16,17,30,63,64], and reflects the two centres of nodulating *Paraburkholderia* diversity in South Africa and South and Central America [15,16,46,65]. Accordingly, the symbiotic loci of these continentally separated groups of rhizobia were acquired independently, early during the evolution of *Paraburkholderia* species in the respective centres [31,91,92].

Following the system proposed by Paulitsch et al. [91], strains representing *P. tuberum* sensu lato form part of sv. africana and atlantica. In the case of sv. africana, all South African species and strains were encompassed, while sv. atlantica contained strains of *P. atlantica* and *P. youngii*, as well as unnamed strains from the Mishra et al. [76] and Bontemps et al. [18] studies. All other mimosoid nodulating species from outside South Africa were included in sv. mimosae, tropicalis and piptadeniae. Apart from sv. africana, these groupings therefore did not reflect their geographic origins and, to some extent, also not their host range. The sv. system is based on the lineages recovered from the phylogenies of symbiotic gene sequences, which is similar to the system that has been developed for *Bradyrhizobium*. In this rhizobial genus, however, the system is well-developed, and *nod* and *nif* genealogies are invaluable resources for studying the evolutionary processes underpinning the geographic distribution of these rhizobia (e.g., [83,84,105,105]). Therefore, further development of the *Paraburkholderia* sv. system would undoubtedly complement taxonomic studies and improve our understanding of the functional diversity and ecology of these bacteria [91,92].

Acknowledgements
The authors would like to acknowledge the National Research Foundation (NRF), the Department of Science and Innovation (DSI) of South Africa and the German Academic Exchange Service (DAAD) for financial support. The authors also extend their gratitude to the Laboratory for Microscopy and Microanalysis Unit of the University of Pretoria for providing electron microscopy infrastructure and guidance. Whole genome sequencing was provided by MicrobesNG (http://www.microbesng.uk), which is supported by the BBSRC (grant number BB/L024209/1). The authors would also like to acknowledge Dr B. Lemaire for his isolation of several strains featured in this study.

References

1. Altschul, S.F., Gish, W., Miller, W., Meyers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
2. Angus, A.A., Lee, A., Lum, M.R., Shehayeb, M., Hessabi, R., Fujishige, N.A., Yerrapragada, S., Kano, S., Song, N., Yang, P., Estrada de los Santos, P., de Faria, S.M., Dakora, F.D., Weinstock, Hirsch, A.M. (2013) Nodulation and effective nitrogen fixation of *Macroptilium atropurpureum* (siratro) by *Burkholderia tuberum*, a nodulating and plant growth promoting beta-proteobacterium, are influenced by environmental factors. Plant. Soil. 369, 543–562.
3. Arahal, D. (2014) Whole-genome analyses: average nucleotide identity. Method. Microbiol. 41, 103–122.
4. Ardley, J.K., Parker, M.A., De Meyer, S.E., Trengove, R.D., O'Hara, G.W., Reeve, W.G., Yates, R.J., Dilworth, M.J., Willems, A., Howieson, J.G. (2012) *Microvirga lupini* sp. nov., *Microvirga lotononidis* sp. nov. and *Microvirga zambiensis* sp. nov. are alphaproteobacterial root-nodule bacteria that specifically nodulate and fix nitrogen with geographically and taxonomically separate legume hosts. Int. J. Syst. Evol. Microbiol. 62, 2579–2588.
5. Aserse, A.A., Woyke, T., Kyrpides, N.C., Whitman, W.B., Lindstrom, K. (2017) Draft genome sequences of *Bradyrhizobium shewense* sp. nov. ERR11^{T} and *Bradyrhizobium yuanmingense* CCBAU 10071^{T}. Stand. Genomic. Sci. 12, 74.
6. Atkinson, E.M., Palcic, M.M., Hindsgaul, O., Long, S.R. (1994) Biosynthesis of *Rhizobium meliloti* lipooligosaccharide Nod factors: NodA is required for an N-acyltransferase activity. Proc. Natl. Acad. Sci. 91, 8418–8422.
7. Baek, I., Seo, B., Lee, I., Yi, H., Chun, J. (2015) *Burkholderia monticola* sp. nov., isolated from mountain soil. Int. J. Syst. Evol. Microbiol. 65, 504–509.

8. Baldani, J.I., Baldani, V.L.D., Seldin, L., Döbereiner, J. (1986) Characterization of *Herbaspirillum seropedicae* gen. nov., sp. nov., a root-associated nitrogen-fixing bacterium. Int. J. Syst. Bacteriol. 36, 86–93.
9. Banfi, E., Galasso, G. (2008) New combinations in *Vachellia* Wight & Arn., formerly *Acacia* Mill. ss (Fabaceae). In: Proceedings of the Atti della Società Italiana di Scienze Naturali e del Museo Civico di Storia Naturale diMilano, Vol. 149, Milan, pp. 149–150.
10. 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., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A. (2012) Spades: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.
11. Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W. (2017) GenBank. Nucleic Acids Res. 45, D37–D42.
12. 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.
13. Beukes, C.W., Palmer, M., Manyaka, P., Chan, W.Y., Avontuur, J.R., van Zyl, E., Huntemann, M., Clum, A., Pillay, M., Palaniappan, K., Varghese, N., Mikhailova, N., Stamatis, D., Reddy, T.B.K., Daum, C., Shapiro, N., Markowitz, V., Ivanova, N., Kyrpides, N., Woyke, T., Blom, J., Whitman, W.B., Venter, S.N., Steenkamp, E.T. (2017) Genome data provides high support for generic boundaries in *Burkholderia* sensu lato. Front. Microbiol. 8, 1154.
14. 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, 1195.
15. 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., Venter, S.N. (2019b) *Paraburkholderia strydomiana* sp. nov. and *Paraburkholderia steynii* sp. nov.: rhizobial symbionts of the fynbos legume *Hypocalyptus sophoroides*. Anton. Leeuw. Int. J. G. 112, 1369–1385.
16. Beukes, C.W., Venter, S.N., Steenkamp, E.T., (2021) The history and distribution of nodulating *Paraburkholderia*, a potential inoculum for Fynbos forage species. Grass Forage Sci. 76, 10–32.

17. Bontemps, C., Elliott, G.N., Simon, M.F., Dos Reis Júnior, B.F., Gross, E., Lawton, R.C., Neto N.E., de Fátima Loureiro, M., de Faria, S.M., Sprent, J.I, James, E.K., Young, J.P.W. (2010) *Burkholderia* species are ancient symbionts of legumes. Mol. Ecol. 19, 44–52.
18. Bontemps, C., Rogel, M.A., Wiechmann, A., Mussabekova, A., Moody, S., Simon, M.F., Moulin, L., Elliott, G.N., Lacercat-Didier, L., Dasilva, C., Grether,R., Camargo-Ricalde, S.L., Chen, W., Sprent, J.I., Martínez-Romero, E., Young, J.P.W., James, E.K. (2016) Endemic *Mimosa* species from Mexico prefer alphaproteobacterial rhizobial symbionts. New Phytol. 209, 319–333.
19. Bournaud, C., Moulin, L., Cnockaert, M., Faria, S., Prin, Y., Severac, D., Vandamme, P. (2017) *Paraburkholderia piptadeniae* sp. nov. and *Paraburkholderia ribeironis* sp. nov., two root-nodulating symbiotic species of *Piptadenia gonoacantha* in Brazil. Int. J. Syst. Evol. Microbiol. 67, 432–440.
20. Brill, W.J. (1980) Biochemical genetics of nitrogen fixation. Microbiol. Rev. 44, 449-467.
21. Chen W-M., James, E.K., Coenye, T. (2006) *Burkholderia mimosarum* sp. nov., isolated from root nodules of *Mimosa* spp. from Taiwan and South America. Int. J. Syst. Evol. Micr. 56, 1847–51.
22. Chen, W-M., de Faria, S.M., James, E.K., Elliott, G.N., Lin, K-Y., Chou, J-H., Sheu, S-Y., Cnockaert, M., Sprent, J.I., Vandamme, P. (2007) *Burkholderia nodosa* sp. nov., isolated from root nodules of the woody Brazilian legumes *Mimosa bimucronata* and *Mimosa scabrella*. Int. J. Syst. Evol. Microbiol. 57, 1055–1059.
23. Chen, W-M., de Faria, S.M., Chou, J-H., James, E.K., Elliott, G.N., Sprent, J.I., Bontemps, C., Young, J.P.W., Vandamme, P. (2008) *Burkholderia sabiae* sp. nov., isolated from root nodules of *Mimosa caesalpiniifolia*. Int. J. Syst. Evol. Microbiol. 58, 2174–2179.
24. Chen, M.H., Sheu, S.Y., James, E.K., Young, C.C., Chen, W.M. (2013) *Azoarcus olearius* sp. nov., a nitrogen-fixing bacterium isolated from oil-contaminated soil. Int. J. Syst. Evol. Microbiol. 63, 3755–3761.
25. Chimphango, S.B., Potgieter, G., Cramer, M.D. (2015) Differentiation of the biogeochemical niches of legumes and non-legumes in the Cape Floristic Region of South Africa. Plant. Ecol. 216, 1583–1595.
26. Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, R.D., da Costa, M.S., Rooney, A.P., Yi, H., Xu, X-W., De Meyer, S., Trujillo, M.E. (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int. J. Syst. Evol. Microbiol. 68, 461–466.

27. de Lajudie, P.M., Andrews, A., 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.
28. De Meyer, S.E., Cnockaert, M., Ardley, J.K., Maker, G., Yates, R., Howieson, J.G., Vandamme, P. (2013a) *Burkholderia sprentiae* sp. nov., isolated from *Lebeckia ambigua* root nodules. Int. J. Syst. Evol. Microbiol. 63, 3950–3957.
29. De Meyer, S.E., Cnockaert, M., Ardley, J.K., Trengove, R.D., Garau, G., Howieson, J.G., Vandamme, P. (2013b) *Burkholderia rhynchosiae* sp. nov., isolated from *Rhynchosia ferulifolia* root nodules. Int. J. Syst. Evol. Microbiol. 63, 3944–3949.
30. De Meyer, S.E., Cnockaert, M., Ardley, J.K., Van Wyk, B-E., Vandamme, P.A., Howieson, J.G. (2014) *Burkholderia dilworthii* sp. nov., isolated from *Lebeckia ambigua* root nodules. Int. J. Syst. Evol. Microbiol. 64, 1090–1095.
31. De Meyer, S.E., Briscoe, L., Martinez-Hidalgo, P., Agapakis, C.M., De los Estrada Santos, P., Seshadri, R., Reeve, W., Weinstock, G., O΄Hara, G., Howieson, J.G., Hirsch, A. (2016) Symbiotic *Burkholderia* species show diverse arrangements of *nif/fix* and *nod* genes and lack typical high-affinity cytochrome *cbb3* oxidase genes. Mol. Plant-Microbe Interact. 29, 609–619.
32. De Meyer, S.E., Cnockaert, M., Moulin, L., Howieson, J.G., Vandamme, P. (2018) Symbiotic and non-symbiotic *Paraburkholderia* isolated from South African *Lebeckia ambigua* root nodules and the description of *Paraburkholderia fynbosensis* sp. nov. Int. J. Syst. Evol. Microbiol. 68, 2607–2614.
33. Dludlu, M.N., Chimphango, S.B.M., Walker, G., Stirton, C.H., Muasya, A.M. (2018a) Horizontal gene transfer among rhizobia of the Core Cape Subregion of southern Africa. S. Afr. J. Bot. 118, 342–352.
34. Dludlu, M.N., Chimphango, S.B.M., Stirton, C.H., Muasya, A.M. (2018b) Differential preference of *Burkholderia* and *Mesorhizobium* to pH and soil types in the Core Cape Subregion, South Africa. Genes. 9, 2.
35. Dobritsa, A.P., Samadpour, M. (2016) Transfer of eleven species of the genus *Burkholderia* to the genus *Paraburkholderia* and proposal of *Caballeronia* gen. nov. to accommodate twelve species of the genera *Burkholderia* and *Paraburkholderia*. Int. J. Syst. Evol. Microbiol. 66, 2836–2846.

36. Dobritsa, A.P., Linardopoulou, E.V., Samadpour, M. (2017) Transfer of 13 species of the genus *Burkholderia* to the genus *Caballeronia* and reclassification of *Burkholderia jirisanensis* as *Paraburkholderia jirisanensis* comb. nov. Int. J. Syst. Evol. Microbiol. 67, 3846–3853.
37. Elliott, G.N., Chen, W.M., Bontemps, C., Chou, J.H., Young, J.P.W., Sprent, J.I., James, E.K. (2007) Nodulation of *Cyclopia* spp. (Leguminosae, Papilionoideae) by *Burkholderia tuberum*. Ann. Bot. 100, 1403–1411.
38. Estrada-de los Santos, P., Palmer, M., Chávez-Ramírez, B., Beukes, C., Steenkamp, E.T., Briscoe, L., Khan, N., Maluk, M., Lafos, M., Humm, E., Arrabit, M., Crook, M., Gross, E., Simon, M.F., dos Reis Junior, F.B., Whitman, W.B., Shapiro, N., Poole, P.S., Hirsch, A.M., Venter, S.N., James, E.K. (2018) Whole genome analyses suggests that *Burkholderia* sensu lato contains two additional novel genera (*Mycetohabitans* gen. nov., and *Trinickia* gen. nov.): implications for the evolution of diazotrophy and nodulation in the *Burkholderiaceae*. Genes. 9, 2073–4425.
39. Euzéby, J.P. (1997) List of bacterial names with standing in nomenclature: a folder available on the internet. Int. J. Syst. Bacteriol. 47, 591–591.
40. Felsenstein, J. (1985) Confidence-limits on phylogenies—an approach using the bootstrap. Evol. 39, 783–791.
41. Garau, G., Yates, R.J., Deiana, P., Howieson, J.G. (2009) Novel strains of nodulating *Burkholderia* have a role in nitrogen fixation with papilionoid herbaceous legumes adapted to acid, infertile soils. Soil Biol. Biochem. 41, 125–134.
42. Gillis, M., Van, T.V., Bardin, R., Goor, M., Hebbar, P., Willems, A., Segers, P., Kersters, K., Heulin, T., Fernandez, M.P. (1995) Polyphasic taxonomy in the genus *Burkholderia* leading to an emended description of the genus and proposition of *Burkholderia vietnamiensis* sp. nov. for N_2-fixing isolates from rice in Vietnam. Int. J. Syst. Bacteriol. 45, 274–289.
43. Goldblatt, P., Manning, J.C. (2002) Plant diversity of the Cape region of southern Africa. Ann. Mo. Bot. Gard. 89, 281–302.
44. 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.
45. Gu, J.Y., Zang, S.G., Sheng, X.F., He, L.Y., Huang, Z., Wang, Q. (2015) *Burkholderia susongensis* sp. nov., a mineral-weathering bacterium isolated from weathered rock surface. Int. J. Syst. Evol. Microbiol. 65, 1031–1037.

46. Gyaneshwar, P., Hirsch, A.M., Moulin, L., Chen, W.-M., Elliott, G.N., Bontemps, C., Estrada-de los Santos, P., Gross, E., dos Reis, F.B., Jr., Sprent, J.I., Young, J.P.W., James, E.K. (2011) Legume-nodulating betaproteobacteria: diversity, host range, and future prospects. Mol. Plant Microbe Interact. 24, 1276–1288.
47. Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
48. Hasegawa, M., Kishino, K., Yano, T. (1985) Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.
49. 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.
50. Hördt, A., Lopez, M.G., Meier-Kolthoff, J.P., Schleuning, M., Weinhold, L.M., Tindall, B.J., Gronow, S., Kyrpides, N.C., Woyke, T., Goker, M. (2020) Analysis of 1,000+ Type-strain genomes substantially improves taxonomic classification of *Alphaproteobacteria*. Front. Microbiol. 11, 468.
51. Howieson, J.G., De Meyer, S.E., Vivas-Marfisi, A., Ratnayake, S., Ardley, J.K., Yates, R.J. (2013) Novel *Burkholderia* bacteria isolated from *Lebeckia ambigua* – a perennial suffrutescent legume of the fynbos. Soil. Biol. Biochem. 60, 55–64.
52. Hungria, M., O'Hara, G.W., Zilli, J.E., Araujo, R.S., Deaker, R., Howieson, J.G. (2016) Isolation and growth of rhizobia. In: Howieson, J.G., Dilworth, M.J. (Eds.) (2016) Working with rhizobia, Australian Centre for International Agricultural Research (ACIAR), Australia, pp. 50.
53. Jordan, D.C. (1982) Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* sp. nov., a genus of slow-growing, root nodule bacteria from leguminous plants. Int. J. Syst. Bacteriol. 32, 136–139.
54. Jourand, P., Giraud, E., Bena, G., Sy, A., Willems, A., Gillis, M., Dreyfus, B., de Lajudie, P. (2004) *Methylobacterium nodulans* sp. nov., for a group of aerobic, facultatively methylotrophic, legume root-nodule-forming and nitrogen-fixing bacteria. Int. J. Syst. Evol. Microbiol. 54, 2269–2273.
55. 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.
56. Katoh, K., Misawa, K., Kuma, K., Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids. Res. 30, 3059–3066.

57. Katoh, K., Toh, H. (2008) Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. BMC Bioinform. 9, 212.
58. Kim, H.B., Park, M.J., Yang, H.C., An, D.S., Jin, H.Z., Yang, D.C. (2006) *Burkholderia ginsengisoli* sp. nov., a β-glucosidase-producing bacterium isolated from soil of a ginseng field. Int. J. Syst. Evol. Microbiol. 56, 2529–2533.
59. Kock, M.M. (2003) Diversity of root nodulating bacteria associated with *Cyclopia* species. PhD thesis, Pretoria: University of Pretoria.
60. Kück, P., Meusemann, K. (2010) FASconCAT: Convenient handling of data matrices. Mol. Phylogen. Evol. 56, 1115–1118.
61. Kuykendall, L.D., Saxena, B., Devine, T.E., Udell, S.E. (1982) Genetic diversity in *Bradyrhizobium japonicum* Jordan 1982 and a proposal for *Bradyrhizobium elkanii* sp. nov. Can. J. Microbiol. 38, 501–505.
62. Lardi, M., Bolzan de Campos, S., Purtschert, G., Eberl, L., Pessi, G. (2017) Competition experiments for legume infection identify *Burkholderia phymatum* as a highly competitive β-rhizobium. Front. Microbiol. 8, 1527.
63. Lemaire, B., Dlodlo, O., Chimphango, S., Stirton, C.H., Schrire, B.D., Boatwright, J.S., Honnay, O., Smets, E., Sprent, J.I., Euan, K.J., 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.
64. Lemaire, B., Chimphango, S.B.M., Stirton, C., Rafudeen, S., Honnay, O., Smets, E., Chen, W.-M., Sprent, J., James, E.K., Muasya, A.M. (2016a) Biogeographical patterns of legume-nodulating *Burkholderia* spp.: from African fynbos to continental scales. Appl. Environ. Microbiol. 82, 5099 –5115.
65. Lemaire, B., Van Cauwenberghe, J., Verstraete, B., Chimphango, S., Stirton, C., Honnay, O., Smets, E., Sprent, J., James, E.K., Muasya, M.A. (2016b) Characterization of the papilionoid–*Burkholderia* interaction in the Fynbos biome: The diversity and distribution of beta-rhizobia nodulating *Podalyria calyptrata* (Fabaceae, Podalyrieae). Syst. Appl. Microbiol. 39, 41–48
66. Liu, W.Y.Y., Ridgway, H.J., James, T.K., James, E.K., Chen, W-M., Sprent, J.I., Young, J.P.W., Andrews, M. (2014) *Burkholderia* sp. induces functional nodules on the South African invasive legume *Dipogon lignosus* (Phaseoleae) in New Zealand soils. Microb. Ecol. 68, 542–555.

67. LPWG (2017) A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. International Association for Plant Taxonomy (IAPT).
68. MacFaddin, J.F. (2000) Gram-negative bacteria, Biochemical tests for identification of medical bacteria. Third Edition. Philadelphia: Lippincott Williams and Wilkins, pp. 363–7.
69. Manning, J., Goldblatt, P. (2012) Plants of the Greater Cape Floristic Region 1: the Core Cape flora, Strelitzia. 29. South African National Biodiversity Institute, Pretoria.
70. Martínez-Aguilar, L., Salazar-Salazar, C., Méndez, R.D., Caballero-Mellado, J., Hirsch, A.M., Vásquez-Murrieta, M.S., Estrada-de los Santos, P. (2013) *Burkholderia caballeronis* sp. nov., a nitrogen fixing species isolated from tomato (*Lycopersicon esculentum*) with the ability to effectively nodulate *Phaseolus vulgaris*. Anton. Leeuw. Int. J. G. 104, 1063–1071.
71. Marx, C.J., Bringel, F., Chistoserdova, L., Moulin, L., Farhan, U.l., Haque, M., Fleischman, D.E., Gruffaz, C., Jourand, P., Knief, C., Lee, M.C., Muller, E.E., Nadalig, T., Peyraud, R., Roselli, S., Russ, L., Goodwin, L.A., Ivanova, N., Kyrpides, N., Lajus, A., Land, M.L., Medigue, C., Mikhailova, N., Nolan, M., Woyke, T., Stolyar, S., Vorholt, J.A., Vuilleumier, S. (2012) "Complete genome sequences of six strains of the genus *Methylobacterium*". J. Bacteriol. 194, 4746–4748.
72. Masson-Boivin, C., Sachs, J.L. (2018) Symbiotic nitrogen fixation by rhizobia—the roots of a success story. Curr. Opin. Plant. Biol. 44, 7–15.
73. Mavengere, N.R., Ellis, A.G., Le Roux, J.J. (2014) *Burkholderia aspalathi* sp. nov., isolated from root nodules of the South African legume *Aspalathus abietina*. Thunb. Int. J. Syst. Evol. Microbiol. 64, 1906–1912.
74. 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., Venter, S.N., Steenkamp, E.T. (2021) *Paraburkholderia youngii* sp. nov. and *Paraburkholderia atlantica* – Brazilian and Mexican *Mimosa*-associated rhizobia that were previously known as *Paraburkholderia tuberum* sv. Mimosa. Syst. Appl. Microbiol. 44: 126152.
75. Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P., Göker, M. (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bionformatics. 14, 60.
76. Mishra, R.P., Tisseyre, P., Melkonian, R., Chaintreuil, C., Miché, L., Klonowska, A., Gonzalez, S., Bena, G., Laguerre, G., Moulin, L. (2012) Genetic diversity of *Mimosa pudica* rhizobial symbionts in soils of French Guiana: investigating the origin and diversity of *Burkholderia phymatum* and other beta-rhizobia. FEMS Microbiol. Ecol. 79, 487–503.

77. Moulin, L., Munive, A., Dreyfus, B., Boivin-Masson, C. (2001) Nodulation of legumes by members of the beta-subclass of *Proteobacteria*. Nature. 411, 948–950.
78. Mthombeni, L.S. (2012) Characterization of *Burkholderia* species associated with root nodules of legumes indigenous to South Africa. MSc dissertation, Pretoria: University of Pretoria, South Africa.
79. 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.
80. Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., McLean, J., Lasken, R., Clingenpeel, S.r., Woyke, T., Tesler, G., Alekseyev, M.A., Pevzner, P.A. (2013) Assembling genomes and mini-metagenomes from highly chimeric reads. In: Deng, M., Jiang, R., Sun, F., Zhang, X. (eds) Research in Computational Molecular Biology. RECOMB 2013. Lecture Notes in Computer Science, vol 7821. Springer, Berlin, Heidelberg.
81. O'Hara, G.W., Zilli, J.E., Poole, P.S., Hungria, M. (2016) Taxonomy and physiology of rhizobia. In: Howieson, J.G., Dilworth, M.J. (Eds.) (2016) Working with rhizobia, Australian Centre for International Agricultural Research (ACIAR), Australia, pp. 133.
82. Ormeño-Orrillo, E., Rogel, M.A., Chueire, L.M., Tiedje J.M., Martínez-Romero, E., Hungria, M. (2012) Genome sequences of *Burkholderia* sp. strains CCGE1002 and H160, isolated from legume nodules in Mexico and Brazil. J. Bacteriol. 194, 6927.
83. Parker, M.A. (2012) Legumes select symbiosis island sequence variants in *Bradyrhizobium*. Mol. Ecol. 21, 1769–1778.
84. Parker, M.A., Rousteau, A. (2014) Mosaic origins of *Bradyrhizobium* legume symbionts on the Caribbean island of Guadeloupe. Mol. Phylogenet. Evol. 77, 110–115.
85. 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.
86. Parks, D.H., Chuvochina, M., Chaumeil, P.A., Rinke, C., Mussig, A.J., Hugenholtz, P. (2020) A complete domain-to-species taxonomy for Bacteria and Archaea. Nat. Biotechnol. 38, 1079–1086.
87. Parte, A.C. (2013) LPSN – list of prokaryotic names with standing in nomenclature. Nucleic. Acids. Res. 42, D613–D616.

88. Parte, A.C. (2018) LPSN – List of prokaryotic names with standing in nomenclature (bacterio.net), 20 years on. Int. J. Syst. Evol. Microbiol. 68, 1825–1829.
89. Paulitsch, F., Dall'Agnol, R.F., Marçon Delamuta, J.R., Augosto Ribeiro, R., daSilva Batista, J.S., Hungria, M. (2019) *Paraburkholderia guartelaensis* sp. nov., a nitrogen-fixing species isolated from nodules of *Mimosa gymnas* in an eco-tone considered as a hotspot of biodiversity in Brazil. Arch. Microbiol. 201, 1435–1446.
90. Paulitsch, F., Dall'Agnol, R.F., Marçon Delamuta, J.R., Augosto Ribeiro, R., da Silva Batista, J.S., Hungria, M. (2020a) *Paraburkholderia atlantica* sp. nov. and *Paraburkholderia franconis* sp. nov., two new nitrogen-fixing nodulating species isolated from Atlantic forest soils in Brazil. Arch. Microbiol. 202, 1369–1380.
91. Paulitsch, F., Marçon Delamuta, J.R., Augosto Ribeiro, R., da Silva Batista, J.S., Hungria, M. (2020b) Phylogeny of symbiotic genes reveals symbiovars within legume-nodulating *Paraburkholderia* species. Syst. Appl. Microbiol. 43, 126151.
92. Paulitsch, F., dos Reis Jr, F.B., Hungria, M. (2021) Twenty years of paradigm-breaking studies of taxonomy and symbiotic nitrogen fixation by beta-rhizobia, and indication of Brazil as a hotspot of *Paraburkholderia* diversity. Arch. Microbiol.
93. Peeters, C., Meier-Kolthoff, J.P., Verheyde, B., De Brandt, E., Cooper, V.S., Vandamme, P. (2016) Phylogenomic study of *Burkholderia glathei*-like organisms, proposal of 13 novel *Burkholderia* species and emended descriptions of *Burkholderia sordidicola*, *Burkholderia zhejiangensis*, and *Burkholderia grimmiae*. Front. Microbiol. 7, 877.
94. Ramírez, M.D.A., España, M., Aguirre, C., Kojima, K., Ohkama-Ohtsu, N., Sekimoto, H., Yokoyama, T. (2019) *Burkholderia* and *Paraburkholderia* are predominant soybean rhizobial genera in Venezuelan soils in different climatic and topographical regions. Microbes. Environ. 34, 43–58.
95. Reiner, K. (2010) Catalase test protocol. American Society for Microbiology.
96. Richter, M., Rosselló-Móra, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. U. S. A. 106, 19126–19131.
97. Rome, S., Fernandez, M.P., Brunel, B., Normand, P., Cleyet-Marel, J.C. (1996) *Sinorhizobium medicae* sp. nov., isolated from annual Medicago spp. Int. J. Syst. Bacteriol. 46, 972–980.
98. Sawana, A., Adeolu, M., Gupta, R.S. (2014) Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. Front. Genet. 5, 429.

99. Segovia, L., Young, J.P., Martinez-Romero, E. (1993) Reclassification of American *Rhizobium leguminosarum* biovar phaseoli type I strains as *Rhizobium etli* sp. nov. Int. J. Syst. Bacteriol. 43, 374–377.

100. Sheu, S-Y., Chou, J-H., Bontemps, C., Elliott, G.N., Gross, E., dos Reis Junior, F.B., Melkonian, R., Moulin, L., James, E.K., Sprent, J.I., Young, J.P.W., Chen, W-M. (2013) *Burkholderia diazotrophica* sp. nov., isolated from root nodules of *Mimosa* spp. Int. J. Syst. Evol. Microbiol. 63, 435–441.

101. Sheu, S-Y., Chen, M-H., Liu, W.Y.Y., Andrews, M., James, E.K., Ardley, J.K., De Meyer, S.E., James, T.K., Howieson, J.G., Coutinho, B.G., Chen, W-M. (2015) *Burkholderia dipogonis* sp. nov., isolated from root nodules of *Dipogon lignosus* in New Zealand and Western Australia. Int. J. Syst. Evol. Microbiol. 65, 4716–4723.

102. Silva, V.C., Alves, P.A.C., Rhem, M.F.K., dos Santos, J.M.F., James, E.K., Gross, E. (2018) Brazilian species of *Calliandra* Benth. (tribe Ingeae) are nodulated by diverse strains of *Paraburkholderia*, Syst. Appl. Microbiol. 41, 241–250.

103. Spilker, T., Baldwin, A., Bumford, A., Dowson, C.G., Mahenthiralingam, E., LiPuma, J.J. (2009) Expanded multilocus sequence typing for *Burkholderia* species. J. Clin. Microbiol. 47, 2607–2610.

104. Steenkamp, E.T., van Zyl, E., Beukes, C.W., Avontuur, J.R., Chan, W.Y., Palmer, M., Mthombeni, L.S., Phalane, F.L., Sereme, K.T., Venter, S.N. (2015) *Burkholderia kirstenboschensis* sp. nov. nodulates papilionoid legumes indigenous to South Africa. Syst. Appl. Microbiol. 38, 545–54.

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

106. Stępkowski, T., Banasiewicz, J., Granada, C.E., Andrews, M., Passaglia, L.M.P. (2018) Phylogeny and phylogeography of rhizobial symbionts nodulating legumes of the tribe Genisteae. Genes. 9, 163.

107. Suau, A., Bonnet, R., Sutren, M., Godon, J.J., Gibson, G.R., Collins, M.D., Doré, J. (1999) Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol. 65, 4799–807.

108. Tamura, K. (1992) "Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C content biases". Mol. Biol. Evol. 9, 678–687.

109. Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

110. Vandamme, P., Goris, J., Chen, W-M., de Vos, P., Willems, A. (2002) *Burkholderia tuberum* sp. nov. and *Burkholderia phymatum* sp. nov., nodulate the roots of tropical legumes. Syst. Appl. Microbiol. 25, 507–512.

111. Venter, S.N., Palmer, M., Beukes, C.W., Chan, W-Y., Shin, G., Zyl, E., Seale, T., Coutinho, T.A., Steenkamp, E.T. (2017) Practically delineating bacterial species with genealogical concordance. Anton. Leeuw. Int. J. G. 110, 1311–1325.

112. Weber, C.F., King, G.M. (2017) Volcanic soils as sources of novel CO-oxidizing *Paraburkholderia* and *Burkholderia*: *Paraburkholderia hiiakae* sp. nov., *Paraburkholderia metrosideri* sp. nov., *Paraburkholderia paradise* sp. nov., *Paraburkholderia peleae* sp. nov., and *Burkholderia alpine* sp. nov. a member of the *Burkholderia cepacia* complex. Front. Microbiol. 8, 207.

113. Yang, Z. (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J. Mol. Evol. 39, 306–314.

114. Yates, R.J., Howleson, J.G., Reeve, W.G., Nandasena, K.G., Law, I.J., Bräu, L., Ardley, J.K., Nistelberger, H.M., Real, D., O'Hara, G.W. (2007) *Lotononis angolensis* forms nitrogen fixing, lupinoid nodules with phylogenetically unique, fast-growing, pink-pigmented bacteria, which do not nodulate *L. bainesii* or *L. listii*. Soil Biol. Biochem. 39, 1680–1688.

115. Zhang, H., Hanada, S., Shigematsu, T., Shibuya, K., Kamagata, Y., Kanagawa, T., Kurane, R. (2000) *Burkholderia kururiensis* sp. nov., a trichloroethylene (TCE)-degrading bacterium isolated from an aquifer polluted with TCE. Int. J. Syst. Evol. Micr. 50, 743–749.

116. Zilli, J.É., de Moraes Carvalho, C.P., de Matos Macedo, A.V., de Barros Soares, L.H., Gross, E., James, E.K., Simon, M.F., de Faria, S.M. (2021) Nodulation of the neotropical genus *Calliandra* by alpha or betaproteobacterial symbionts depends on the biogeographical origins of the host species. *Braz. J. Microbiol.*