Paraburkholderia solitsugae sp. nov. and Paraburkholderia elongata sp. nov., phenolic acid–degrading bacteria isolated from forest soil and emended description of Paraburkholderia…

Article *in* International Journal of Systematic and Evolutionary Microbiology · August 2020

DOI: 10.1099/ijsem.0.004387

CITATIONS
7

READS
396

6 authors, including:

Roland Wilhelm
Purdue University
40 PUBLICATIONS **695** CITATIONS
SEE PROFILE

K. Taylor Cyle
North Carolina State University
13 PUBLICATIONS **215** CITATIONS
SEE PROFILE

Jeffrey D Newman
Lycoming College
48 PUBLICATIONS **795** CITATIONS
SEE PROFILE

INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

TAXONOMIC DESCRIPTION
Wilhelm *et al.*, *Int. J. Syst. Evol. Microbiol.*
DOI 10.1099/ijsem.0.004387

MICROBIOLOGY SOCIETY

Paraburkholderia solitsugae sp. nov. and *Paraburkholderia elongata* sp. nov., phenolic acid-degrading bacteria isolated from forest soil and emended description of *Paraburkholderia madseniana*

Roland C. Wilhelm[1,*], K. Taylor Cyle[1], Carmen Enid Martinez[1], David C. Karasz[1], Jeffrey D. Newman[2] and Daniel H. Buckley[1]

Abstract

Two bacterial strains, 1N^T and 5N^T, were isolated from hemlock forest soil using a soluble organic matter enrichment. Cells of 1N^T (0.65×1.85 µm) and 5N^T (0.6×1.85 µm) are Gram-stain-negative, aerobic, motile, non-sporulating and exist as single rods, diplobacilli or in chains of varying length. During growth in dilute media (≤0.1× tryptic soy broth; TSB), cells are primarily motile with flagella. At higher concentrations (≥0.3× TSB), cells of both strains increasingly form non-motile chains, and cells of 5N^T elongate (0.57×~7 µm) and form especially long filaments. Optimum growth of 1N^T and 5N^T occurred at 25–30 °C, pH 6.5–7.0 and <0.5% salinity. Results of comparative chemotaxonomic, genomic and phylogenetic analyses revealed that 1N^T and 5N^T were distinct from one another and their closest related type strains: *Paraburkholderia madseniana* RP11^T, *Paraburkholderia aspalathi* LMG 27731^T and *Paraburkholderia caffeinilytica* CF1^T. The genomes of 1N^T and 5N^T had an average nucleotide identity (91.6 and 91.3%) and *in silico* DNA–DNA hybridization values (45.8%±2.6 and 45.5%±2.5) and differed in functional gene content from their closest related type strains. The composition of fatty acids and patterns of substrate use, including the catabolism of phenolic acids, also differentiated strains 1N^T and 5N^T from each other and their closest relatives. The only ubiquinone present in strains 1N^T and 5N^T was Q-8. The major cellular fatty acids were $C_{16:0}$, 3OH-$C_{16:0}$, $C_{17:0}$ cyclo, $C_{19:0}$ cyclo $\omega 8c$ and summed features 2 (3OH-$C_{14:0}$ / $C_{16:1}$ iso I), 3 ($C_{16:1}$ $\omega 6c/\omega 7c$) and 8 ($C_{18:1}$ $\omega 7c/\omega 6c$). A third bacterium, strain RL16-012-BIC-B, was isolated from soil associated with shallow roots and was determined to be a strain of *P. madseniana* (ANI, 98.8%; 16S rRNA gene similarity, 100%). Characterizations of strain RL16-012-BIC-B (DSM 110723=LMG 31706) led to proposed emendments to the species description of *P. madseniana*. Our polyphasic approach demonstrated that strains 1N^T and 5N^T represent novel species from the genus *Paraburkholderia* for which the names *Paraburkholderia solitsugae* sp. nov. (type strain 1N^T=DSM 110721^T=LMG 31704^T) and *Paraburkholderia elongata* sp. nov. (type strain 5N^T=DSM 110722^T=LMG 31705^T) are proposed.

INTRODUCTION

The genus *Paraburkholderia* was recently established from the division of *Burkholderia*, which is presently delineated into seven genera: *Burkholderia*, *Paraburkholderia*, *Caballeronia*, *Robbsia*, *Trinickia*, *Mycetohabitans* and *Pararobbsia* [1–5]. Currently, *Paraburkholderia* contains the second greatest number of described species (n=78, of which 71 are validly named) and has expanded rapidly, with a doubling of newly described species in the past 5 years. Nearly all *Paraburkholderia* have been isolated from soil, rhizosphere or plant root tissues (Table S1, available in the online version of this article). Many species are capable

Author affiliations: [1]School of Integrative Plant Sciences, Bradfield Hall, Cornell University, Ithaca, NY, 14853, USA; [2]Lycoming College, Williamsport, PA, 17701, USA.
***Correspondence:** Roland C. Wilhelm, rwilhelm@cornell.edu
Keywords: forest soil; filamentation; *Paraburkholderia*; phenolic acid-degrading.
Abbreviations: ANI, average nucleotide identity; BLASTp, basic local alignment search tool for proteins; DyP-type, dye-decolorizing peroxidase-type enzyme; FAME, fatty acid methyl esters; IAA, indole-3-acetic acid; MLSA, multi-locus sequence alignment; pobA, 3-hydroxybenzoate 4-monooxygenase gene; RAST, Rapid Annotation using Subsystem Technology; RL16, *Paraburkholderia madseniana* RL16-012-BIC-B; rrn, ribosomal RNA operon; rRNA, ribosomal RNA; RuBisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; SEM, scanning electron microscopy; SESOM, soil-extracted, solubilized organic matter; TSB, tryptic soy broth.
Genome accessions: WOEZ00000000 (*P. solitsugae* 1N^T), WOEY00000000 (*P. elongata* 5N^T) and WVHR00000000 (*P. madseniana* RL16-012-BIC-B). 16S rRNA gene accessions: MN723156, MN723157 and MK373450, respectively. NCBI taxonomy IDs: 2675748, 2675747 and 2694937, respectively.
Two supplementary figures and ten supplementary tables are available with the online version of this article.

of root nodulation [6–16], or exhibit an endophytic lifestyle [17–20]; and can perform symbiotic or asymbiotic dinitrogen fixation [9, 21]. Notably, several nodule isolates were incapable of nodulating the plant species from which they were isolated [22]. Our understanding of the ecology and evolution of *Paraburkholderia* can be improved by expanding the representation of species from this genus.

Most *Paraburkholderia* isolated from soil originate from forest ecosystems (Table S1), where members of the genus have been shown to fix dinitrogen [21], solubilize mineral phosphate [23, 24], degrade lignin [25] and enhance the degradation of soil organic matter [26, 27]. Two bacterial strains, 1N^T and 5N^T, were isolated from forest soil as part of an effort to characterize substrate preferences and uptake kinetics of soluble organic matter [28]. These strains belong to a clade containing both forest soil (*Paraburkholderia madseniana*) and root nodule (*Paraburkholderia aspalathi*) isolates [29]. We identified a third isolate from this clade, strain RL16-012-BIC-B (henceforth 'RL16'), which was isolated from the rhizosphere of a *Digitalis* species in a wooded city park [30]. Our characterization of these three strains provides an opportunity to expand understanding of the physiological and genomic traits within a clade comprising species isolated from both soil and roots.

ISOLATION AND ECOLOGY

Strains 1N^T and 5N^T were isolated from the upper 5 cm of the B horizon of a moderately well-drained Dystrudept soil (pH 4.3–4.5) from a hemlock stand at the Arnot research forest (Van Etten, NY;42.278611° N, –76.634361° W) [31, 32]. After serial dilution, a soil slurry was spread plated onto agar media prepared with soil-extracted, solubilized organic matter (SESOM) derived from the overlying Oa horizon. The chemical composition of SESOM (pH 3.55) was comprehensively characterized by Cyle *et al.* [28] and contained total organic carbon and nitrogen concentrations of 185.5 mg l^{-1} C and 11.3 mg l^{-1} N, respectively. Colonies appeared after 7–14 days of incubation at room temperature and strains 1N^T and 5N^T were streaked for isolation on SESOM agar. Additional details about the environmental source and growth attributes of strain 1N^T are provided in [28]. Strain RL16-012-BIC-B (also referred to as 'RL16-012-BSH-B') was isolated by Haeckl *et al.* [30] during the development of a genome-guided method for isolating *Burkholderia*. Strain RL16-012-BIC-B (henceforth 'RL16') was isolated from soil associated with the roots of understory *Digitalis* in a mixed deciduous woodland in a city park adjacent to Maple Place Towers in Burnaby, BC (49.269418° N, 122.94937° W). The strain was enriched using a base medium designed for isolating *Burkholderia* supplemented with L-sorbose, hydroxyproline (1 g l^{-1}) and antibacterials acriflavine and fusaric acid [30]. For all chemotaxonomic and growth characterizations, the strains were cultured on dilute tryptic soy broth (0.1× TSB; recipe in Supplementary Methods) at 25 °C, salinity 0.1% (w/v NaCl) and pH 7.0, unless otherwise specified.

PHYLOGENETIC AND GENOME FEATURES

Genomic DNA was extracted from strains 1N^T, 5N^T and RL16 according to the protocol of Griffiths *et al.* [33] and submitted to the Cornell University Sequencing Facility for sequencing using three multiplexed runs of Illumina MiSeq Nano (2×250 bp). Raw sequencing data was quality preprocessed with Trimmomatic (version 0.32) [34] and FastX Toolkit (version 0.7) [35] then assembled with SPAdes (version 3.10.1) [36]. Assemblies were scaffolded with Ragout using the *P. madseniana* RP11^T genome for a reference [37]. Raw sequencing data and genome assemblies were accessioned under the NCBI BioProject accession PRJNA590275. The phylogeny of strains were determined from a maximum-likelihood tree based on a multi-locus sequence alignment (MLSA) of 49 housekeeping genes (Table S2) using the KBase [38] application 'Insert Set of Genomes Into Species Tree' (version 2.1.10), dependent on FastTree2 [39]. Based on the MLSA phylogeny, genomes from 10 of the closest relatives to 1N^T and 5N^T were downloaded from the National Center for Biotechnological Information. A maximum-likelihood phylogenetic tree was reconstructed from full-length 16S rRNA genes from close relatives using MEGA X [40] with the Tamurai–Nei substitution model, a uniform substitution rate and 200 bootstraps for branch support. *Caballeronia glathei* DSM 50014^T (GCA_000698595.1) served as the outgroup for all phylogenetic analyses. The number of copies of the *rrn* operon was determined based on the ratio of average read depth for the consensus 16S rRNA gene versus single-copy genes identified using BUSCO [41]. Single nucleotide polymorphisms (SNPs) in the 16S rRNA gene were manually identified using read mapping and visualized using the Integrated Genome Viewer [42]. Genome G+C content and DNA–DNA hybridization values were predicted *in silico* based on genomic data using the Type (Strain) Genome Server [43]. Functional gene annotations were performed on open-reading frames predicted with Prodigal (version 2.6.2) [44] using RAST [45]. Specific catabolic genes were targeted with hmmsearch [46] using hidden Markov models supplied by [25] for laccases, aryl alcohol oxidases and dye-decolouring peroxidases. Secreted proteins were identified based on signal peptide sequence predictions from SignalP-5.0 [47] with a threshold of p_{other} <0.05. Plant growth-promoting genes characteristic of the endophyte, *P. phytofirmans* PsJN^T, and root nodulating species, *P. mimosarum* LMG 23256^T, were manually annotated with BLASTp (>40% similarity across >90% length), targeting nitrogenase, nodulation factors, 1-aminocyclopropane-1-carboxylate (ACC) deaminase [48, 49], gibberellin [50] and indole-3-acetic acid (IAA) synthesis pathways [51–53]. Annotations for all genes presented in Table 1 were manually verified based on homology with characterized enzymes.

Strains 1N^T and 5N^T were most closely related to type strain *P. madseniana* RP11^T, although strain 5N^T was most closely related to proposed strain *P. solitsugae* 1N^T on the basis of average nucleotide identity (ANI), DNA–DNA hybridization and functional gene content (Tables 1 and S3). Measures of ANI and DNA–DNA hybridization were below the respective thresholds (95 and 70%, respectively)

Table 1. Phylogenetic and genomic characteristics that differentiate strains 1N^T, 5N^T and RL16-012-BIC-B from their ten closest related type strains

Strains: 1, 1N^T; 2, *Paraburkholderia madseniana* RP11^T; 3, *Paraburkholderia madseniana* RL16-012-BIC-B; 4, 5N^T; 5, *Paraburkholderia aspalathi* LMG 27731^T; 6, *Paraburkholderia caffeinilytica* CF1^T; 7, *Paraburkholderia fungorum* LMG 16225^T; 8, *Paraburkholderia sediminicola* HU2-65W^T; 9, *Paraburkholderia phytofirmans* PsJN^T; 10, *Paraburkholderia aromaticivorans* BN5^T; 11, *Paraburkholderia xenovorans* LB400^T; 12, *Paraburkholderia bryophila* 1S18^T; 13, *Paraburkholderia rhynchosiae* WSM3937^T. In (A), all measures of phylogenetic relatedness use 1N^T as reference and columns were ordered by descending average nucleotide identity. In (B), all measures of phylogenetic relatedness are provided in reference to 5N^T and were also ordered by descending average nucleotide identity.

A	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Genome size (Mb)	11.1	10.1	9.6	9.7	9.89	8.32	8.7	7.31	8.21	8.91	9.73	8.01	8.03
Number of contigs	353	323	224	201	104	3	124	118	3	8	3	91	169
G+C content (mol%)	60.6	61.3	61.5	61.3	61.1	62.2	61.8	63.6	62.3	62.9	62.6	62.9	61.7
Average nucleotide identity value (%)	100	91.6	91.5	91.4	90.7	90.7	89.7	86.9	86.7	86.6	86.6	86.2	85
DNA–DNA hybridization value	100	45.8	45.2	45.5	42.4	41.9	38.9	31.7	31.1	31.4	31.1	29.9	28.6
16S rRNA gene similarity (%)	100	99.3	99.3	98.6	98.8	98.7	99.1	98.9	98.7	98.2	98.8	98.7	98.4
Aromatic degradation genes*	130	105	98	113	99	85	88	66	76	88	97	64	88
Total dioxygenase genes	55	52	44	64	53	45	38	30	33	45	59	24	40
Phthalate 4,5-dioxygenase	+	+	+	+	+	–	–	–	–	–	–	–	–
RuBisCO operon	+	–	–	+	–	–	–	+	–	+	+	–	+
Soluble methane monooxygenase	–	+	+	–	+	–	–	–	–	–	–	–	–
Nitrogenase iron protein (nifH)	–	–	–	–	–	–	–	–	–	+	+	–	+
B	**4**	**1**	**3**	**2**	**5**	**6**	**7**	**8**	**9**	**11**	**10**	**12**	**13**
Average nucleotide identity value (%)	100	91.4	91.1	91.1	90.4	90.3	89.1	86.6	86.4	86.4	86.4	86	85
DNA–DNA hybridization value	100	45.5	43.6	43.8	41.1	40.5	36.8	31.3	30.7	30.9	30.7	29.5	28.5
16S rRNA gene similarity (%)	100	98.6	99.2	99.2	99.1	97.9	98.8	98.9	98.8	98.2	98.3	97.8	98.1

*Total dereplicated RAST SEED subsystem feature counts.

for delineating new species [54, 55]. The MLSA phylogeny placed 1N^T and 5N^T in the same clade as *P. madseniana* and *P. aspalathi* (Fig. 1a). The 16S rRNA gene-based phylogeny broadly lacked branch support and was deemed unreliable (Fig. 1b), consistent with observations that the 16S rRNA gene is an unreliable indicator for delineating species of *Paraburkholderia* [1, 29, 56]. Strains 1N^T, 5N^T and RL16 each contained six copies of the 16S rRNA gene with sequence heterogeneity evident among copies (see Supplementary Methods). This heterogeneity offers one possible explanation for the poor resolving power of the 16S rRNA gene-based phylogeny, which may be an artefact of comparing varying consensus sequences.

The genome assembly for strain 1N^T was substantively larger than its closest relative, *P. madseniana* RP11^T and is among the largest of published *Paraburkholderia* genomes, totalling 11075000 bases (N50 value, 64800; read depth, 24×) with 10636 predicted open reading frames and the lowest G+C content of all 11 related strains (Table 1). The functional gene content of strain 1N^T differed from RP11^T (Table S4), including the presence of the complete ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) operon and the absence of a soluble methane monooxygenase (Table 1). The genome assembly for strain 5N^T was considerably smaller than 1N^T, its closest relative, with a total of 9667127 bases (N50 value, 86400; read depth, 21.3×) and 8930 open reading frames. Functional gene content of strain 5N^T also differed from 1N^T (Table S5), including the presence of genes encoding a toluene-4-monooxygenase and a type IV secretion system. Strain 5N^T encoded the greatest number of dioxygenases (n=64) of all *Paraburkholderia* genomes examined (Table 1) and also encoded the RuBisCO operon. Conversely, strain 5N^T encoded among the fewest secreted oxidases of any genome, while strain 1N^T encoded the greatest number, including laccases, an aryl alcohol oxidase and a DyP-type peroxidase (Table S6).

Strain RL16 was closely related to *P. madseniana* RP11^T in terms of 16S rRNA gene sequence (100 % identity), ANI (98.8%) and DNA–DNA hybridization values (90.3±2.1). The RL16 genome was slightly smaller than *P. madseniana* RP11^T, totalling 9594840 bases (N50 value, 127114; read depth, 42.9×) with 9620 predicted open reading frames (Table 1). The consensus 16S rRNA gene sequence of strain RL16 contained three single nucleotide polymorphisms

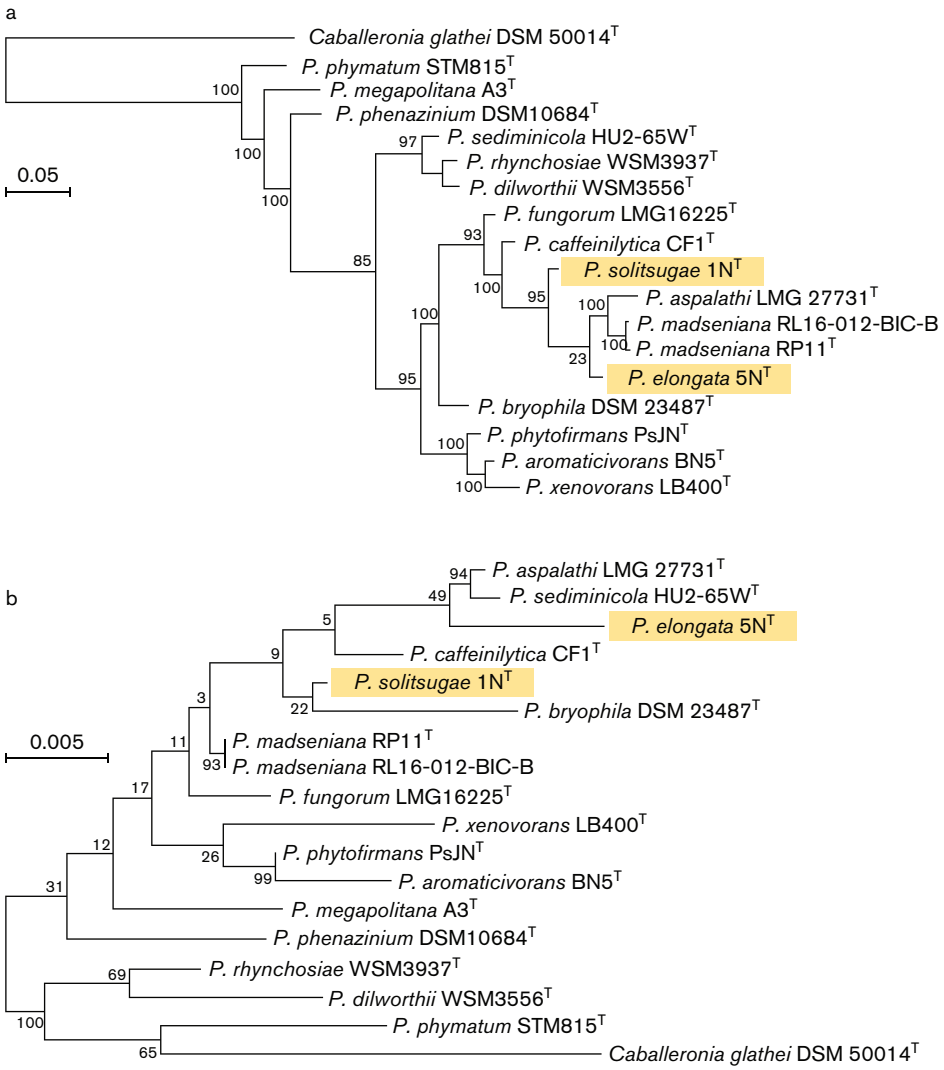

Fig. 1. The phylogenetic relationships of strains 1N^T, 5N^T and RL16-012-BIC-B with closely related species of *Paraburkholderia* according to maximum-likelihood phylogenetic trees based on a multi-locus sequence alignment of (a) 49 concatenated house-keeping genes and (b) the full-length 16S rRNA gene. The scale bar corresponds to substitutions per nucleotide position. Accessions for genome assemblies and full-length 16S rRNA genes, respectively: *P. aspalathi* (GCF_900116445.1); *P. caffeinilytica* (GCF_003368325.1, NR_152088.1); *P. fungorum* (GCF_000685055.1, NR 025058.1); *P. bryophila* (GCF_003269035.1); *P. aromaticivorans* (GCF_002278075.1, NR_163658.1); *P. phytofirmans* (GCF_000020125.1, NR_102845.1); *P. sediminicola* (GCF_900104005.1, NR_044383.1); *P. xenovorans* (GCF_000013645.1, NR_074325.2); *P. rhynchosiae* (GCF_002879865.1, NR_116248.1); *P. dilworthii* (GCF_000472525.1, NR_125580.1); *P. phenazinium* (GCF_900100735.1); *P. megapolitana* (GCF_900113825.1; NR_042594.1); *P. phymatum* (GCF_000020045.1) and *Caballeronia glathei* (GCF_000698595.1). If an accession is not specified, the 16S rRNA gene was recovered from the genome assembly indicated.

absent in *P. madseniana* RP11^T, indicating a maximum potential 16S rRNA gene dissimilarity of 0.2 % (1465/1468 nt; see Supplementary Methods). Like other members of its clade, strain RL16 encoded a high number of aromatic degradation genes (n=98) and secreted oxidases, but slightly fewer than *P. madseniana* RP11^T (n=105), including one fewer laccase. Strain RL16 did not encode a paralog of the 3-hydroxybenzoate 4-monooxygenase gene (*pobA*), which was deemed a characteristic of *P. madseniana* [29]. The functional gene content of strain RL16 also differed from *P. madseniana* RP11^T in several ways, including the absence of a large number of oligo and dipeptide ABC transporters, several amino acid biosynthesis and scavenging pathways, and xylose and ribose sugar utilization (Table S7). Neither strain encoded canonical plant growth-promoting genes for the synthesis of nitrogenases, nodulation factors, gibberellin or IAA. All *Paraburkholderia* genomes encoded the AAC deaminase operon and indoleacetamide hydrolase (*iaaH*), but only *P. phymatum* and *P. phenanzium* encoded the accompanying tryptophan 2-monooxygenase (*iaaM*) essential for indole-3-acetamide mediated synthesis of IAA.

PHYSIOLOGY AND CHEMOTAXONOMY

Chemotaxonomic characterizations were performed for strains $1N^T$, $5N^T$ and RL16 along with their three closest relatives: *P. madseniana* $RP11^T$, *P. aspalathi* $LMG\ 27731^T$ and *P. caffeinilytica* $CF1^T$. Determinations of enzyme activity and metabolic activity were performed using plate-based, colorimetric assay API ZYM strips (bioMérieux) and Biolog GEN III plates (Biolog), respectively, according to the manufacturers' instructions. The composition of cellular fatty acids was determined for cells grown on nutrient agar for 2 days at 22 °C and methylated according to [57] using an Agilent 6850 gas chromatograph configured by Microbial ID Inc. (MIDI) with the Sherlock Microbial Identification System (version 6.1) and the RTSBA6 database. Major respiratory quinones were determined by analysis of acetone extracts on an Agilent 6545 LC/Q-TOF MS using a modification of the methods described by [58] (details in Supplementary Methods). Antibiotic resistance was assessed on nutrient agar plates at 22 °C by measuring the zone of inhibition around filter paper discs containing ampicillin (10 µg), cephalexin (30 µg), chloramphenicol (30 µg), ciprofloxacin (5 µg), clindamycin (2 µg), erythromycin (15 µg), gentamycin (120 µg), kanamycin (30 µg), nalidixic acid (30 µg), penicillin (10 U), rifampicin (5 µg), spectinomycin (100 µg), streptomycin (10 µg), sulfamethoxazole (20 µg)+trimethoprim (4 µg), or tetracycline (30 µg) with a diameter of 10 mm or less considered resistant. Oxidase activity was tested using Oxistrips (MilliporeSigma). Catalase activity was assessed based on the production of bubbles after mixing a drop of 3% H_2O_2 (v/v; Wards Scientific) with a loop full of active culture. Gram staining was performed according to the method of Smibert and Krieg [59].

Strains $1N^T$ and $5N^T$ were phenotypically different from each other and their closest relatives according to the Biolog assay (Table 2). Strain $1N^T$ was uniquely able to metabolize acetic acid, while strain $5N^T$ was distinctly able to use D-glucose-6-phosphate, D-fructose-6-phosphate and D-galacturonic acid (Table 2). Strain $5N^T$ was singularly susceptible to several antibiotics (rifamycin, lincomycin and vancomycin) and unable to metabolize oligo- and polysaccharides used by close relatives (sucrose, raffinose and pectin). Strain $5N^T$ differed from $1N^T$ in its capacity to grow on ferulic acid and guaiacol, though it did not grow on phthalic acid despite encoding a phthalate 4,5-dioxygenase with high homology to $1N^T$ (88 and 89% identity for α- and β-subunits, respectively). Strains $1N^T$, RL16 and *P. madseniana* $RP11^T$ were capable of growth on phthalic acid and were the only other strains to encode a phthalate 4,5-dioxygenase (Table 1). Strain RL16 exhibited an identical phenotypic profile as *P. madseniana* $RP11^T$ except for the inability to metabolize several carbohydrates and organic acids (Table 2).

Strain $1N^T$ exhibited lipase (C14) and strain $5N^T$ alkaline phosphatase activity that were absent in close relatives, but, otherwise, shared the characteristic enzyme activity of related species, including acid phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase and naphthol-AS-BI-phosphohydrolase activities (Table S8). The major cellular fatty acids profiles of strains $1N^T$ and $5N^T$ were comparable to *P. madseniana*, which were substantially higher in the proportions of $C_{16:0}$ and $C_{17:0}$ cyclo and lower in 2OH-$C_{16:0}$ than *P. aspalathi* and *P. caffeinilytica* (Table 3). The fatty acid profile of strain $1N^T$ differed from $5N^T$ the proportion of summed features 3 ($C_{16:1}$ ω6*c*/ω7*c*) and 8 ($C_{18:1}$ ω7*c* and/or $C_{18:1}$ ω6*c*). The fatty acid profile of strain RL16 differed from *P. madseniana* $RP11^T$ in the proportions of summed feature 3 and 8 (full fatty acid methyl ester data in Table S9). The only respiratory quinone observed in strains $1N^T$, $5N^T$ and RL16, like all other relatives, was ubiquinone Q-8. Cells of $1N^T$, $5N^T$ and RL16 were resistant to ampicillin, penicillin and clindamycin (Table S9). Strain $5N^T$ was uniquely susceptible to vancomycin and lincomycin in the Biolog assay.

Salinity, pH and temperature growth optima were determined in dilute (0.1×) TSB medium based on measurements of optical density (OD at λ=600 nm). All assays were performed in duplicate in 20 ml test tubes bearing 10 ml liquid media shaken at a slant at 180 r.p.m. on an orbital shaker and monitored over a period of 72 h, except where specified otherwise. Cultures were inoculated with 20 µl of actively growing culture normalized to an OD_{600} of 0.5. The pH optimum was determined over pH range 3, 4, 5, 6, 6.5, 7, 8 and 8.5 using buffer systems described in the Supplementary Methods. The temperature optimum was assessed at 4, 25, 30 and 37 °C. Salinity tolerance was tested at 0.5, 1, 1.5, 2 and 3% (w/v NaCl). The capacity for growth on benzoic acid and phenolics (guaiacol, vanillin, syringic acid, ferulic acid, phthalic acid, salicylic acid and 4-coumaric acid) was determined in mineral salts media after a 1 week incubation according to [29], except that growth substrates were filter sterilized. The nutrient-dependent regulation of growth morphology of strains $1N^T$, $5N^T$, RL16 and *P. madseniana* $RP11^T$ was determined by assaying growth (OD_{600}) across a gradient of TSB (0.05, 0.1, 0.3, 0.6, 1 and 5×) with salinity maintained at 0.1% NaCl.

Changes in population-level cell size distributions were determined across the TSB gradient by phase-contrast microscopy (Olympus CX41). We imaged five fields of view per biological replicate (n=3) at ×200 magnification with an OMAX digital microscope camera (U3CMOS18000KPA) at 250 ms exposure. Cell size measurement was automated using the image analysis software ImageJ (version 1.52a) [60] with the MicrobeJ plugin (version 5.131) [61] (details in Supplementary Methods; images in Supplementary Data package). Scanning electron micrographs were taken of cells sampled at late log-phase growth on 1× and 1 : 10× TSB media. A 1 ml sample of cells were fixed with glutaraldehyde, post-fixed in osmium tetroxide and vacuum filtered onto Whatman #5 filter paper (protocol in Supplementary Methods). Critical point drying was performed in a Bal-tec 030 (Bal-tec) and samples were sputter coated with iridium in a Desk V thin film deposition sample preparation system (Denton Vacuum). SEM imaging was performed with a field emission scanning electron microscope (Zeiss GeminiSEM 500) at the Cornell Center for Materials Research.

Table 2. A summary of Biolog profiles and substrate use that differentiate strains 1N^T, 5N^T and RL16-012-BIC-B from the closest type strains of *Paraburkholderia*

Strains: 1, 1N^T; 2, 5N^T; 3, *Paraburkholderia madseniana* RP11^T; 4, *Paraburkholderia madseniana* RL16-012-BIC; 5, *Paraburkholderia aspalathi* LMG 27731^T; 6, *Paraburkholderia caffeinilytica* CF1^T. All data are from this study and the full assay results are provided in Table S9. A 'w' denotes 'weak growth' on a phenolic compound, indicating marginal growth was observed after 2 weeks.

Biolog profile/substrate	1	2	3	4	5	6
Acetic acid	+	–	–	–	–	–
Glycyl-L-proline	–	+	+	+	+	–
Aztreonam	+	+	–	–	–	–
α-Keto-butyric acid	+	+	–	–	+	+
Citric acid	–	–	+	+	+	+
D-Glucose-6-PO4	–	+	–	–	–	+
D-Fructose-6-PO4	–	+	–	–	–	+
D-Galacturonic acid	–	+	–	–	+	–
Sucrose	+	–	+	+	–	–
D-Raffinose	+	–	+	+	–	–
Pectin	+	–	+	+	–	–
4-Hydroxy-phenylacetic acid	+	–	+	+	+	+
D-Aspartic acid	+	–	+	+	+	+
L-Aspartic acid	+	–	+	+	+	+
Rifamycin SV*	+	–	+	+	+	+
Lincomycin*	+	–	+	+	+	+
Vancomycin*	+	–	+	+	+	+
L-Galacturonic acid lactone	–	+	+	–	+	–
D-Glucuronic acid	–	+	+	–	+	–
D-Malic acid	–	+	+	–	+	–
L-Rhamnose	–	+	+	–	+	+
L-Serine	+	–	+	–	+	–
Formic acid	+	–	+	–	+	+
α-Hydroxy-butyric acid	+	+	+	–	+	–
Growth on phenolic acids						
4-Hydroxybenzoic acid	+	+	+	+	+	+
Salicylic acid	+	+	+	+	+	+
Benzoic acid	+	+	+	+	+	+
4-Coumaric acid	+	+	–	–	+	+
Phthalic acid	+	–	+	+	–	–
guaiacol	–	+	–	–	–	–
Ferulic acid	–	+	–	–	w	+
Vanillin	–	–	–	–	–	+
Syringic acid	–	–	–	–	–	w

*Tolerance to antibiotic.

Table 3. Cellular fatty acid composition of strains 1N^T, 5N^T and RL16-012-BIC-B and closely related type strains from the genus *Paraburkholderia*

Strains: 1, 1N^T; 2, 5N^T; 3, *Paraburkholderia madseniana* RP11^T; 4, *Paraburkholderia madseniana* RL16-012-BIC; 5, *Paraburkholderia aspalathi* LMG 27731^T; 6, *Paraburkholderia caffeinilytica* CF1^T. Values are percentages of total fatty acids. Fatty acids that make up <1% of the total are not shown or are denoted as trace 'TR'. Shading indicates a property which distinguished strains. All data are from this study the full assay results are provided in Table S9.

Fatty acid	1	2	3	4	5	6
C12: 0	2.07	2.42	2.46	2.38	2.01	1.84
C14: 0	1.26	1.54	0.85	1.07	1.11	0.88
$C_{16:0}$	23.8	29.2	26.0	31.5	18.6	18.8
$C_{16:1}$ 2-OH	1.64	1.08	1.16	0.9	2.52	0.82
$C_{16:0}$ 2-OH	TR	TR	TR	0	2.19	0.42
$C_{16:0}$ 3-OH	3.86	4.37	4.37	4.77	4.34	4.37
$C_{17:0}$ cyclo	12.6	22.0	18.0	27.3	9.76	1.42
$C_{18:0}$	0.57	0.8	1.0	1.08	0.58	0.95
$C_{18:1}$ 2-OH	0.25	0.38	0.27	0	0.83	0.4
$C_{19:0}$ cyclo ω8*c*	3.23	4.18	4.2	5.24	3.5	0.43
Summed feature 2	4.43	4.86	5.14	5.28	4.79	5.01
Summed feature 3	22.6	11.3	16.6	9.04	14.4	25.8
Summed feature 8	20.7	15.2	18.1	8.75	33.4	38.0
Total	97	97.3	98.1	97.3	98.1	99.1

Cells of strains 1N^T and 5N^T are Gram-negative, rod-shaped, non-sporulating, oxidase and catalase positive. Cells of strain 1N^T (0.65×1.85 µm) and 5N^T (0.6×1.85 µm) grow as individual rods, as diplobacillus (Fig. 2) or in chains (three to 10+ cells). Cell morphology was found to vary depending on nutrient concentration, pH and salinity (Table S10). Cells of strain of strain 5N^T became narrower, elongated (0.57×7.0 µm) and filamentous (>10 cells per chain) at higher TSB concentrations, pH and salinity (Fig. 3). Cells of strain RL16 grow as slightly narrower rods (0.6×1.7 µm) than *P. madseniana* RP11^T (0.65×1.7 µm), and both grow as diplobacillus and occasionally in short to long chains (two to 10 cells). Cells of all strains are predominantly motile by flagella when growing in dilute TSB and non-motile in more concentrated nutrient conditions. Cells of 1N^T and 5N^T attach to surfaces at cell poles, evident in their adherence to other cells (Fig. S1), and the glass and cellulose-based mounts used in microscopy (Fig. 3).

The optimum growth conditions for strains 1N^T and 5N^T occurred at 25–30 °C, pH 6.5–7.0 and with <0.5% NaCl. Strains 1N^T and 5N^T exhibited a narrower tolerance range to pH (4.0–8.0; ≤1% NaCl) compared to RL16 and *P. madseniana* RP11^T (4.0–8.5; ≤1% NaCl). The optimal growth of strains RL16 and *P. madseniana* RP11^T occurred at 25–30 °C, pH 6.5 and with <0.5% NaCl. All strains could grow at 4 °C, but not at 37 °C. Above pH 7.0 and 0.5% NaCl, cells of strain 5N^T occurred predominantly in elongated and filamentous forms. All strains grew to the highest cell density at 0.1× TSB at mid log phase (Fig. 4a). *P. madseniana* RP11^T achieved the highest cell density of all strains in virtually all media concentrations. In more concentrated TSB (>0.3×), cells of strain 5N^T existed predominantly in elongated and filamentous forms, corresponding with decreased cell density (Fig. 4). The frequency and length of filaments increased with TSB concentration for strain 5N^T and, to a lesser extent 1N^T, but did not differ greatly over time (Fig. S2a). Cells of strain RL16 and *P. madseniana* RP11^T rarely formed chains longer than six to 10 cells and the occurrence of chained forms was invariant to TSB concentration (Fig. S2a).

DISCUSSION

Our polyphasic approach establishes that strains 1N^T and 5N^T constitute novel species in the genus *Paraburkholderia*. The strains met all phylogenetic and chemotaxonomic criteria for new species and exhibited major physiological and functional differences to related type strains. The names, *Paraburkholderia solitsugae* sp. nov. and *Paraburkholderia elongata* sp. nov. are proposed.

The chain formation observed in strains 1N^T, 5N^T, RL16 and *P. madseniana* RP11^T is not a common morphological characteristic of *Paraburkholderia*. Cells of *Paraburkholderia* are primarily described as 'rod-like' and found to occur singly and, for several species, as diplobacilli [21, 62–64], including relatives of the strains characterized here [29, 65]. Chain formation has only been reported for *P. madseniana* [29], though an earlier description of *Paraburkholderia* species noted the occurrence of 'irregular clusters' of cells [66] which was later included in the genus description [5]. Chain formation may be specific to the clade of *Paraburkholderia* examined here or may be more widespread but overlooked due its occurrence under specific growth conditions. Cells of *Caballeronia*, the neighboring genus to *Paraburkholderia*, also grow in pairs and in short chains [67], indicating chain formation in *Burkholderiaceae* may be paraphyletic and more prevalent than currently understood. In any case, our results demonstrate that chain formation does not distinguish *Caballeronia* from *Paraburkholderia* as suggested by the genus descriptions [5].

The regulation of cell motility, chain formation and elongation in response to TSB concentration represents a newly described characteristic of *Paraburkholderia*. In dilute media, all strains exhibited flagellar motility, with cells becoming predominantly non-motile at higher nutrient concentrations. Cells of 1N^T and 5N^T formed long chains and were increasingly adherent to surfaces and other cells at higher concentrations. For 5N^T, chain formation was accompanied by cell elongation. In contrast, the cell morphology of RL16 and *P. madseniana* RP11^T was invariant across the growth conditions tested. The greater sensitivity of 1N^T and 5N^T to nutrient concentrations in the range of a standard, nutrient-rich laboratory

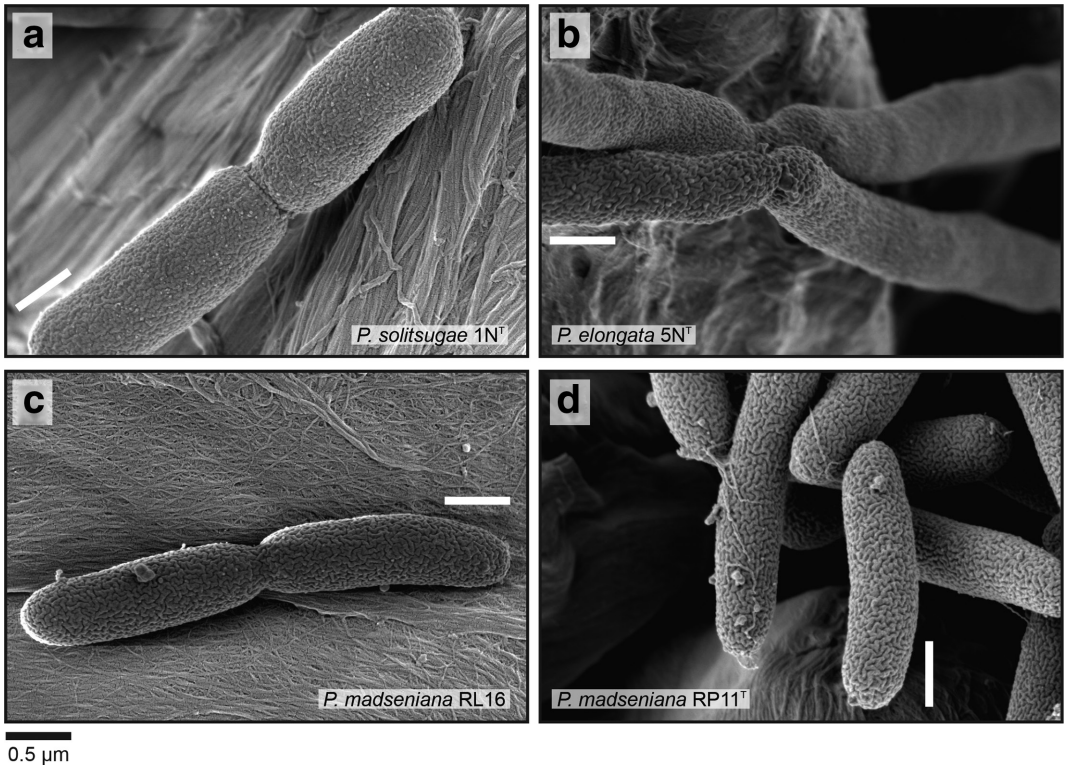

Fig. 2. Cell sizes of strains 1N^T (a), 5N^T (b), RL16-012-BIC-B (c) and *P. madseniana* RP11^T (d) revealed by SEM imaging. All images were taken at late-log phase of growth on 1× TSB and at the same magnification (each panel is 4 µm wide). Cells are mounted on filter paper made from cellulose fibres.

medium may reflect their isolation from a medium designed to approximate the soluble organic matter content of a forest soil [28]. The SESOM medium contained less organic carbon (0.18 g l^{-1} C) than the media used to isolate their closest relatives *P. madseniana* RP11^T (0.32 g l^{-1} C) [29], RL16 (~1.4 g l^{-1} C) [30] and *P. aspalathi* (~4 g l^{-1} C) [65].

Strain 5N^T exhibited the most striking changes in cell morphology in response to TSB concentration. In dilute media (≤0.1× TSB), cells resembled the morphology and size of 1N^T, RL16 and *P. madseniana* RP11^T. However, within a relatively narrow concentration range (0.1–0.3× TSB), cells of strain 5N^T elongated to over double their initial size and formed long filamentous chains. Above this tipping point, the frequency of cell chains and chain length was correlated with TSB concentration and corresponded with reduced cell density, indicating lower biomass yield. The underlying mechanism(s) regulating elongation and chain formation were not determined in the present study, but similar morphological changes were nutrient dependent in *Halomonas elongata*, a fellow member of the *Gammaproteobacteria* [68]. *H. elongata* occur as single and paired cells with polar flagella during log phase growth and form elongated, 'flexuous filaments of varying length' at stationary phase [68]. It remains to be determined whether the chained and elongated morphotype of strain 5N^T constitutes an adaptive trait or reflects the pathological condition of cells under stress. We can conclude that salinity and pH are not the sole regulators of filamentation in strain 5N^T, having been controlled in the TSB gradient, indicating phosphate, glucose and/or components of the peptone/soytone digests also regulate the morphotypic differentiation.

The characterization of RL16 was undertaken to expand the diversity of a clade of root- and soil-derived *Paraburkholderia*. Strain RL16 was isolated from soil from shallow roots [30] while its closest relative, *P. madseniana* RP11^T, was isolated from the underlying O-horizon of forest soil in a study of decomposition [29]. Our analyses determined RL16 is a strain of *P. madseniana* which differed in several aspects from the original species description. *P. madseniana* RL16 had a slightly smaller genome and encoded fewer dioxygenases than the closest type strain, *P. aspalathi* LMG 27731^T. Strain RL16 also lacked paralogs of the 3-hydroxybenzoate 4-monooxygenase gene (*pobA*), proposed to differentiate *P. madseniana* from several of its closest relatives [29]. Growth on phthalic acid was a distinct feature of both strains of *P. madseniana* compared to *P. aspalathi*, but not compared to the newly described strain 1N^T. Overall, strain RL16 encoded fewer amino acid biosynthesis and scavenging pathways than RP11^T and, in all cases, lacked the ability to metabolize compounds utilized by RP11^T. RL16 also exhibited a slower

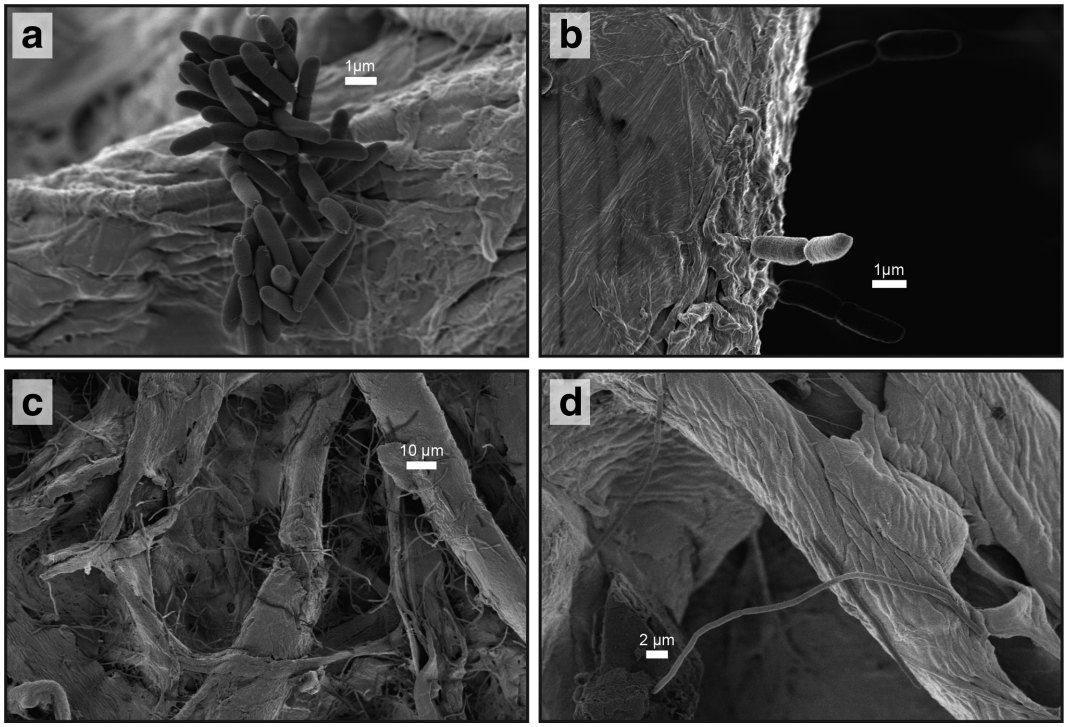

Fig. 3. Differences in cell morphology between strains 1N^T (AB) and 5N^T (CD) under high nutrient conditions where strain 1N^T exist predominantly as individual cells or diplobacillus and where 5N^T cells exist in elongated, chained, filamentous forms. All SEM images were taken at late-log phase of growth on 1× TSB. The capacity of cells to adhere to the cellulose fibres of the filter paper mount was evident in both 1N^T (b) and 5N^T (d).

growth rate and achieved lower cell densities across the TSB gradient. Neither strain encoded genes characteristic of plant-growth promoting activity observed in *P. phytofirmans* and *P. mimosarum*.

DESCRIPTION OF *PARABURKHOLDERIA SOLITSUGAE* SP. NOV.

Paraburkholderia solitsugae (so.li.tsu'gae. L. neut. n. *solum* soil; N.L. fem. n. *Tsuga* scientific name of hemlock; N.L. gen. n. *solitsugae* of/from soil of a hemlock forest).

Cells are aerobic, Gram-negative, motile, non-sporulating rods (0.65 µm wide by 1.85 µm long) that grow primarily as motile, in dilute media (<0.1× TSB), or non-motile, in more concentrated media (>0.3× TSB) bacillus or diplobacilli. Cells also grow in chains (three to 10 cells) at higher nutrient/solute concentration. Optimal growth occurs on 0.1× TSB at 25 °C (range: >4–30 °C), pH 6.5–7.0 (4.0–8.5) and salinity <0.5% NaCl (0–1.0%). Colonies are round, convex, translucent white in colour and mostly regular in shape with entire margins. Cells are resistant to ampicillin (10 µg), penicillin (10 U), and clindamycin (2 µg). Positive reactions are observed for acid phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), leucine arylamidase and naphthol-AS-BI-phosphohydrolase. Tests were negative for utilization of alkaline phosphatase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase. Tests were positive for utilization of 4-hydroxy-phenylacetic acid, acetic acid, bromo-succinic acid, D-arabitol, D-aspartic acid, D-fructose, D-galactose, D-gluconic acid, D-lactic acid methyl ester, D-mannitol, D-mannose, raffinose, D-saccharic acid, D-sorbitol, formic acid, glycerol, L-alanine, L-arginine, L-aspartic acid, L-fucose, L-glutamic acid, L-histidine, lincomycin, L-lactic acid, L-malic acid, L-pyroglutamic acid, L-serine, methyl pyruvate, mucic acid, myo-inositol, *N*-acetyl-D-glucosamine, pectin, quinic acid, sucrose, Tween-40, α-D-glucose, α-hydroxy-butyric acid, α-keto-butyric acid, β-hydroxy-D,L-butyric acid, γ-amino-butyric acid, 4-hydroxybenzoic acid, salicylic acid, benzoic acid, 4-coumaric acid and phthalic acid. Tests were negative for utilization of 3-methyl glucose, acetoacetic acid, citric acid, cellobiose, dextrin, D-fructose-6-PO_4, D-fucose, D-galacturonic acid, D-glucose-6-PO_4, D-glucuronic acid, D-malic acid, maltose, melibiose, D-salicin, D-serine, D-serine, trehalose, turanose, fusidic acid, gelatin, gentiobiose, glucuronamide, glycyl-L-proline, inosine, L-galacturonic acid lactone, L-rhamnose, *N*-acetyl neuraminic acid, *N*-acetyl-β-D-mannosamine, *N*-acetyl-D-galactosamine, propionic acid, lactose, α-keto-glutaric acid, methyl β-D-glucoside, guaiacol, ferulic acid, vanillin and syringic acid. The most

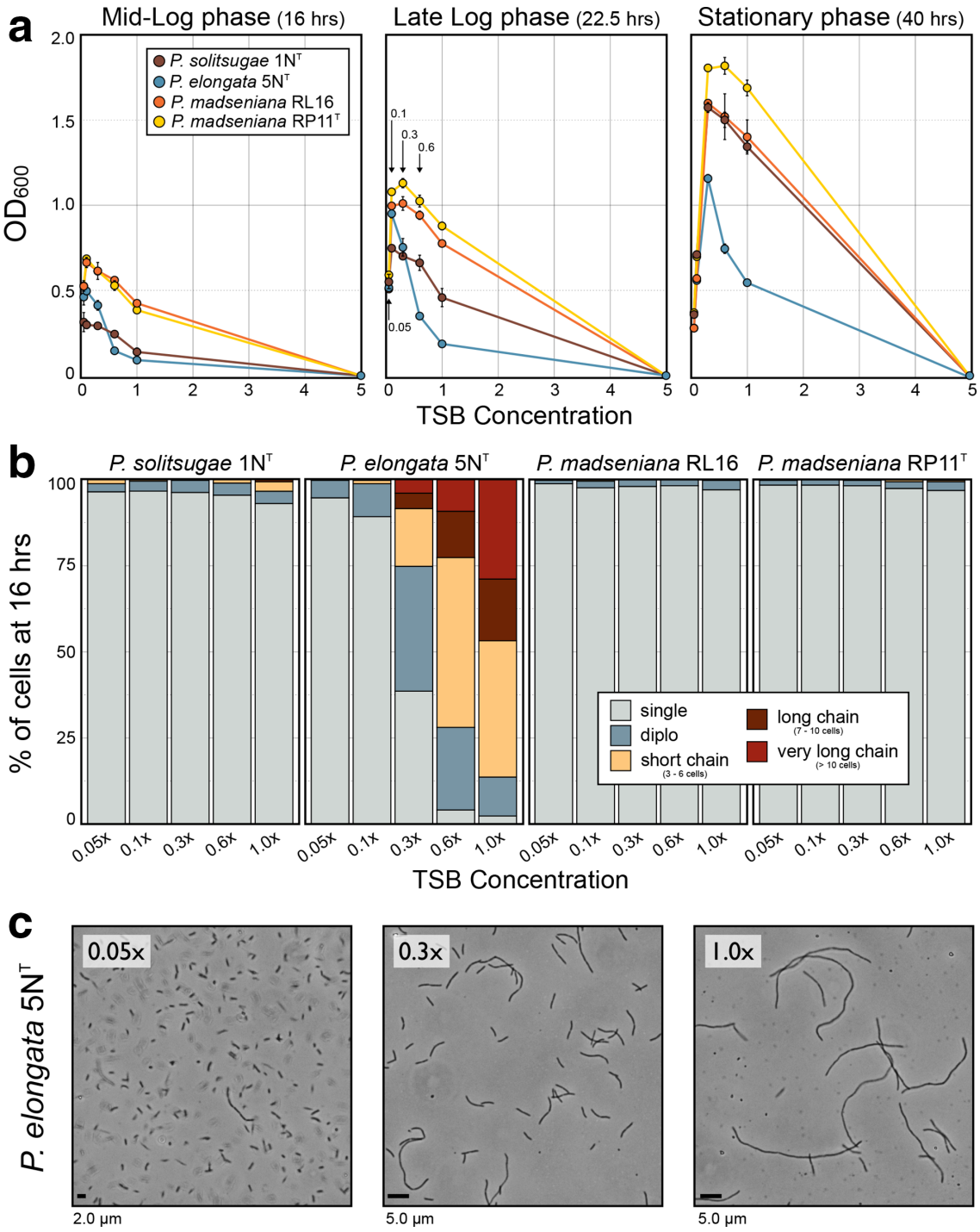

Fig. 4. Growth characteristics of strain 5N^T differed from closest relatives across a gradient in TSB concentration. In (a), the maximum cell density (OD_{600}) of strain 5N^T occurred at lower TSB concentrations (0.05–0.3×) than other strains. In (b), cells of strain 5N^T occurred predominantly in elongated and chained forms at TSB concentration ≥0.3×, with chain length increasing at higher concentrations. In (c), cells of 5N^T were motile in dilute media (0.05–0.1×), evident in blurring of cells in phase-contrast images at ×200 magnification), with cells shifting to predominantly chained form at 0.3× TSB. Cells had begun to decrease in size by late-log phase in 0.05 and 0.1×, indicating starvation (Fig. S2a).

abundant cellular fatty acids (ordered by abundance) are $C_{16:0,}$ summed feature 3, summed feature 8, $C_{17:0}$ cyclo, summed feature 2, $C_{16:0}$ 3-OH, $C_{19:0}$ cyclo ω8*c*, $C_{12:0}$, $C_{16:1}$ 2-OH, $C_{14:0}$, $C_{18:0}$, $C_{18:1}$ 2-OH and $C_{16:0}$ 2-OH. The major respiratory quinone is ubiquinone 8.

The type strain, 1N^T (=DSM 110721^T=LMG 31704^T), was isolated using a solubilized soil organic matter enrichment from the upper 5 cm of the B horizon of a moderately well-drained Dystrudept soil (pH 4.3–4.5) from a hemlock stand at the Arnot research forest. The DNA G+C content of the type strain is 60.6 mol%. The unassembled and assembled genome sequencing data (WOEZ00000000) and 16S rRNA gene (MN723156) were assigned to the NCBI BioProject: PRJNA590275.

DESCRIPTION OF *PARABURKHOLDERIA ELONGATA* SP. NOV.

Paraburkholderia elongata (e.lon.ga′ta. L. fem. part. adj. *elongata* elongated, stretched, pertaining to cell elongation and predisposition to form filamentous cell chains).

Cells are aerobic, Gram-negative, motile, non-sporulating rods (0.6 µm by 1.85 µm long) that grow primarily as motile bacillus or diplobacilli at dilute nutrient concentrations (<0.3× TSB) and as elongated rods (0.57×7.0 µm), forming short (three to six cells) or longer chains (>10 cells) at higher nutrient/solute concentrations. Optimal growth occurred on 0.1× TSB at 25 °C (range, 4–30 °C), pH 6.5–7.0 (4.0 to <8.0) and salinity <0.5% (0–0.5%). Colonies are round, convex, translucent white in colour and mostly regular in shape with entire margins. Cells are resistant to ampicillin (10 µg), penicillin (10 U) and clindamycin (2 µg). Positive reactions are observed for acid phosphatase, alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase and naphthol-AS-BI-phosphohydrolase. Tests are negative for utilization of lipase (C14), valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase. Tests were positive for utilization of 4-hydroxy-phenylacetic acid, bromo-succinic acid, D-arabitol, D-fructose, D-fructose-6-PO_4, D-galactose, D-galacturonic acid, D-gluconic acid, D-glucose-6-PO_4, D-glucuronic acid, D-malic acid, D-mannitol, D-mannose, D-saccharic acid, D-sorbitol, glycerol, glycyl-L-proline, L-alanine, L-arginine, L-fucose, L-galacturonic acid lactone, L-glutamic acid, L-histidine, L-lactic acid, L-malic acid, L-pyroglutamic acid, L-rhamnose, methyl pyruvate, mucic acid, myo-inositol, *N*-acetyl-D-galactosamine, *N*-acetyl-D-glucosamine, quinic acid, Tween 40, α-D-glucose, α-hydroxy-butyric acid, α-keto-butyric acid, β-hydroxy-D,L-butyric acid, γ-amino-butyric acid, 4-hydroxybenzoic acid, salicylic acid, benzoic acid, 4-coumaric acid, guaiacol and ferulic acid. Test results are negative for utilization of 3-methyl glucose, acetic acid, acetoacetic acid, citric acid, D-aspartic acid, cellobiose, dextrin, D-fucose, D-lactic acid methyl ester, maltose, melibiose, raffinose, D-salicin, D-serine, D-serine, trehalose, turanose, formic acid, fusidic acid, gelatin, gentiobiose, glucuronamide, inosine, L-aspartic acid, L-serine, *N*-acetyl neuraminic acid, *N*-acetyl-β-D-mannosamine, pectin, propionic acid, stachyose, sucrose, lactose, α-keto-glutaric acid, methyl β-D-glucoside, phthalic acid, vanillin and syringic acid. The most abundant cellular fatty acids (ordered by abundance) are $C_{16:0,}$ $C_{17:0}$ cyclo, summed feature 8, summed feature 3, summed feature 2, $C_{16:0}$ 3-OH, $C_{19:0}$ cyclo ω8*c*, $C_{12:0}$, $C_{14:0}$, $C_{16:1}$ 2-OH, $C_{18:0}$, $C_{18:1}$ 2-OH and $C_{16:0}$ 2-OH. The major respiratory quinone is ubiquinone 8.

The type strain, 5N^T (=DSM 110722^T=LMG 31705^T), was isolated using a solubilized soil organic matter enrichment from the upper 5 cm of the B horizon of a moderately well-drained Dystrudept soil (pH 4.3–4.5) from a hemlock stand at the Arnot research forest. The DNA G+C content of the type strain is 61.3 mol%. The unassembled and assembled genome sequencing data (WOEY00000000) and 16S rRNA gene (MN723157) were assigned to the NCBI BioProject: PRJNA590275.

EMENDED DESCRIPTION OF *PARABURKHOLDERIA MADSENIANA* WILHELM *ET AL.* 2020

Description is as given in Wilhelm *et al.* [55] except that cells are resistant to ampicillin (10 µg), penicillin (10 U) and clindamycin (2 µg), and response is variable in Biolog GenIII wells containing α-hydroxy-butyric acid, D-glucuronic acid, D-malic acid, formic acid, L-rhamnose and L-serine.

The *Paraburkholderia madseniana* strain RL16-012-BIC-B (=DSM 110723=LMG 31706) was isolated from soil adjacent to roots of understory *Digitalis* in a mixed deciduous woodland in a city park adjacent to Maple Place Towers in Burnaby, BC. The DNA G+C content of the strain is 61.5 mol%. The unassembled and assembled genome sequencing data (WVHR00000000) and 16S rRNA gene (MK373450) were assigned to the NCBI BioProject: PRJNA590275.

Funding information
This research was supported by the U.S. Department of Energy, Office of Biological and Environmental Research Genomic Science Program under award number DE-SC0016364 and by the AFRI Education and Workforce Development Program, grant no. 2019-67011-29513, from the U.S. Department of Agriculture, National Institute of Food and Agriculture. Partial funding was provided by the Cornell University Program in Cross-Scale Biogeochemistry and Climate, which is supported by NSF-IGERT and the Atkinson Center for Sustainable Future. Additional financial support for K.T.C. was provided by the College of Agriculture and Life Sciences at Cornell University in support of graduate research.

Acknowledgements
We would like to acknowledge the contributions of Jake Haeckl, Claire Fergusson and Dr. Roger Linington from Simon Fraser University for isolating and sharing *P. madseniana* RL16-012-BIC-B. We would like to acknowledge John Grazul and Malcolm Thomas at the Cornell Center for Materials Research for their outstanding expertise in sample preparation and SEM imaging. We thank Phurchhoki Sherpa for performing DNA extractions.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Sawana A, Adeolu M, Gupta RS. 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 env. *Front Genet* 2014;5:1–22.
2. Estrada-de Los Santos P, Palmer M, Beukes C, Steenkamp ET, Briscoe L *et al*. Whole genome analyses suggests that *Burkholderia sensu lato* contains two additional novel genera implications for the evolution of diazotrophy and nodulation in the *Burkholderiaceae*. *Genes* 2018;9:389.
3. Lin Q, Lv Y, Gao Z, Qiu LH. *Pararobbsia silviterrae* gen. nov., sp. nov., isolated from forest soil and reclassification of *Burkholderia alpina* as *Pararobbsia alpina*. *Int J Syst Evol Microbiol* 2019.
4. Beukes CW, Palmer M, Manyaka P, Chan WY, Avontuur JR *et al*. Genome data provides high support for generic boundaries in *Burkholderia sensu lato*. *Front Microbiol* 2017;8:1–12.
5. Dobritsa AP, Samadpour M. 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*2016;66:2836–2846.
6. Vandamme P, Goris J, Chen WM, De Vos P, Willems A. *Burkholderia tuberum* sp. nov. and *Burkholderia phymatum* sp. nov., nodulate the roots of tropical legumes. *Syst Appl Microbiol* 2002;25:507–512.
7. Chen WM, James EK, Coenye T, Chou JH, Barrios E *et al*. *Burkholderia mimosarum* sp. nov., isolated from root nodules of *Mimosa* spp. from Taiwan and South America. *Int J Syst Evol Microbiol* 2006;56:1847–1851.
8. Bournaud C, Moulin L, Cnockaert M, de Faria S, Prin Y *et al*. *Paraburkholderia piptadeniae* sp. nov. and *Paraburkholderia ribeironis* sp. nov., two root-nodulating symbiotic species of *Piptadenia gonoacantha* in Brazil. *Int J Syst Evol Microbiol* 2017;67:432–440.
9. Elliott GN, Chen WM, Chou JH, Wang HC, Sheu SY *et al*. *Burkholderia phymatum* is a highly effective nitrogen-fixing symbiont of *Mimosa* spp. and fixes nitrogen *ex planta*. *New Phytol* 2007;173:168–180.
10. Chen WM, de Faria SM, James EK, Elliott GN, Lin KY *et al*. *Burkholderia nodosa* sp. nov., isolated from root nodules of the woody Brazilian legumes *Mimosa bimucronata* and *Mimosa scabrella*. *Int J Syst Evol Microbiol* 2007;57:1055–1059.
11. Cunha C de O, Zuleta LFG, de Almeida LGP, Ciapina LP, Borges WL *et al*. Complete genome sequence of *Burkholderia phenoliruptrix* BR3459a (CLA1), a heat-tolerant, nitrogen-fixing symbiont of *Mimosa flocculosa*. *J Bacteriol* 2012;194:6675–6676.
12. Sheu SY, Chou JH, Bontemps C, Elliott GN, Gross E *et al*. *Burkholderia symbiotica* sp. nov., isolated from root nodules of *Mimosa* spp. native to north-east Brazil. *Int J Syst Evol Microbiol* 2012;62:2272–2278.
13. Sheu SY, Chou JH, Bontemps C, Elliott GN, Gross E *et al*. *Burkholderia diazotrophica* sp. nov., isolated from root nodules of *Mimosa* spp. *Int J Syst Evol Microbiol* 2013;63:435–441.
14. Martínez-Aguilar L, Salazar-Salazar C, Méndez RD, Caballero-Mellado J, Hirsch AM *et al*. *Burkholderia caballeronis* sp. nov., a nitrogen fixing species isolated from tomato (*Lycopersicon esculentum*) with the ability to effectively nodulate *Phaseolus vulgaris*. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol* 2013;104:1063–1071.
15. De MSE, Cnockaert M, Ardley JK, Maker G, Yates R *et al*. *Burkholderia sprentiae* sp. nov., isolated from *Lebeckia ambigua* root nodules. *Int J Syst Evol Microbiol* 2013;63:3950–3957.
16. De Meyer SE, Cnockaert M, Ardley JK, Van Wyk BE, Vandamme PA *et al*. *Burkholderia dilworthii* sp. nov., isolated from *Lebeckia ambigua* root nodules. *Int J Syst Evol Microbiol* 2014;64:1090–1095.
17. Caballero-Mellado J, Martínez-Aguilar L, Paredes-Valdez G, Estrada-de los Santos P. *Burkholderia unamae* sp. nov., an N_2-fixing rhizospheric and endophytic species. *Int J Syst Evol Microbiol* 2004;54:1165–1172.
18. Reis VM, Estrada-de los Santos P, Tenorio-Salgado S, Vogel J, Stoffels M *et al*. *Burkholderia tropica* sp. nov., a novel nitrogen-fixing, plant-associated bacterium. *Int J Syst Evol Microbiol* 2004;54:2155–2162.
19. Perin L, Martínez-Aguilar L, Paredes-Valdez G, Baldani JI, Estrada-de los Santos P *et al*. *Burkholderia silvatlantica* sp. nov., a diazotrophic bacterium associated with sugar cane and maize. *Int J Syst Evol Microbiol* 2006;56:1931–1937.
20. da Silva PRA, Simões-Araújo JL, Vidal MS, Cruz LM, de SEM *et al*. Draft genome sequence of *Paraburkholderia tropica* Ppe8 strain, a sugarcane endophytic diazotrophic bacterium. *Brazilian J Microbiol* 2018;49:210–211.
21. Chan YK. Utilization of simple phenolics for dinitrogen fixation by soil diazotrophic bacteria. *Plant Soil* 1986;90:141–150.
22. De Meyer SE, Cnockaert M, Moulin L, Howieson JG, Vandamme P *et al*. 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* 2018;68:2607–2614.
23. Guo JK, Ding YZ, Feng RW, Wang RG, YM X *et al*. *Burkholderia metalliresistens* sp. nov., a multiple metal-resistant and phosphate-solubilising species isolated from heavy metal-polluted soil in Southeast China. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol* 2015;107:1591–1598.
24. Gao Z-H, Ruan S-L, Huang Y-X, Lv Y-Y, Qiu L-H. *Paraburkholderia phosphatilytica* sp. nov., a phosphate-solubilizing bacterium isolated from forest soil. *Int J Syst Evol Microbiol* 2019;69:196–202.
25. Wilhelm RC, Singh R, Eltis LD, Mohn WW. Bacterial contributions to delignification and lignocellulose degradation in forest soils with metagenomic and quantitative stable isotope probing. *ISME J* 2019;13:413–429.
26. Wilhelm RC, DeRito CM, Shapleigh JP, Buckley DH, Madsen EL. Phenolic acid-degrading *Paraburkholderia* Prime Decomposition in Forest Soil. *MBio*.
27. Zwetsloot MJ, Muñoz Ucros J, Wickings K, Wilhelm RC, Sparks JP *et al*. Prevalent root-derived phenolics drive shifts in microbial community composition and prime decomposition in forest soil. *Soil Biol Biochem* 2020;145.
28. Cyle KT, Klein AR, Aristilde L, Martinez CE. Ecophysiological study of *Paraburkholderia* sp. 1N under soil solution conditions: Dynamic substrate preferences and characterization of carbon use efficiency. *Appl Environ Microbiol*.
29. Wilhelm RC, Murphy SJL, Feriancek NM, Karasz DC, Derito CM *et al*. *Paraburkholderia madseniana* sp. nov., a phenolic acid-degrading bacterium isolated from acidic forest soil. *Int J Syst Evol Microbiol* 2020.
30. Haeckl FPJ, Baldim JL, Iskakova D, Kurita KL, Soares MG *et al*. A selective genome-guided method for environmental *Burkholderia* isolation. *J Ind Microbiol Biotechnol* 2019;46:345–362.
31. Fahey TJ, Yavitt JB, Sherman RE, Groffman PM, Fisk MC *et al*. Transport of carbon and nitrogen between litter and soil organic matter in a northern hardwood forest. *Ecosystems* 2011;14:326–340.
32. Goodale CL, Fredriksen G, Weiss MS, McCalley CK, Sparks JP *et al*. Soil processes drive seasonal variation in retention of 15N tracers in a deciduous forest catchment. *Ecology* 2015;96:2653–2668.
33. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl Environ Microbiol* 2000;66:5488–5491.
34. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
35. Gordon A, G.J. H. Fastx-toolkit. FASTQ/A short-reads preprocessing tools. http://hannonlab.cshl.edu/fastx_toolkit 2010;5.
36. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.

37. Kolmogorov M, Raney B, Paten B, Pham S. Ragout - a reference-assisted assembly tool for bacterial genomes. *Bioinformatics* 2014;30:302–309.
38. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL *et al*. KBase: the United States department of energy systems biology Knowledgebase. *Nat Biotechnol* 2018;36:566-569.
39. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;5:e9490.
40. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;35:1547–1549.
41. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015;31:3210–3212.
42. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES *et al*. Integrative genome Viewer. *Nat Biotechnol* 2011;29:24–26.
43. Meier-Kolthoff JP, Klenk HP, Göker M. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–356.
44. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW *et al*. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;11:119.
45. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ *et al*. The seed and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 2014;42:206–214.
46. Eddy SR. Profile hidden Markov models. *Bioinformatics* 1998;14:755–763.
47. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O *et al*. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 2019;37:420–423.
48. Sun Y, Cheng Z, Glick BR. The presence of a 1-aminocycloprop ane-1-carboxylate (ACC) deaminase deletion mutation alters the physiology of the endophytic plant growth-promoting bacterium *Burkholderia phytofirmans* PsJN. *FEMS Microbiol Lett* 2009;296:131–136.
49. Checcucci A, Azzarello E, Bazzicalupo M, De Carlo A, Emiliani G *et al*. Role and regulation of ACC deaminase gene in *Sinorhizobium meliloti*: Is it a symbiotic, rhizospheric or endophytic gene? *Front Genet* 2017;8:6.
50. Nagel R, Bieber JE, Schmidt-Dannert MG, Nett RS, Peters RJ. A third class: functional gibberellin biosynthetic operon in beta-Proteobacteria. *Front Microbiol* 2018;9:1–8.
51. Weilharter A, Mitter B, Shin M V, Chain PSG, Nowak J *et al*. Complete genome sequence of the plant growth-promoting endophyte *Burkholderia phytofirmans* strain PsJN. *J Bacteriol* 2011;193:3383–3384.
52. Zúñiga A, Poupin MJ, Donoso R, Ledger T, Guiliani N *et al*. Quorum sensing and indole-3-acetic acid degradation play a role in colonization and plant growth promotion of *Arabidopsis thaliana* by *Burkholderia phytofirmans* PsJN. *Mol Plant-Microbe Interact* 2013;26:546–553.
53. Liu WH, Chen FF, Wang CE, HH F, Fang XQ *et al*. Indole-3-Acetic acid in *Burkholderia pyrrocinia* JK-SH007: enzymatic identification of the indole-3-acetamide synthesis pathway. *Front Microbiol* 2019;10:1–12.
54. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P *et al*. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;57:81–91.
55. Jain C, Rodriguez-r LM, Phillippy A, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:5114.
56. Aizawa T, Vijarnsorn P, Nakajima M, Sunairi M. *Burkholderia bannensis* sp. nov., an acid neutralizing bacterium isolated from torpedo grass (*Panicum repens*) growing in highly acidic swamps. *Int J Syst Evol Microbiol* 2011;61:1645–1650.
57. Sasser M. Identification of bacteria by gas chromatography of cellular fatty acids. technical note 101. *Microbial ID, Inc.,Newark, Del* 2001:1–6.
58. Kaiser P, Geyer R, Surmann P, Fuhrmann H. LC-MS method for screening unknown microbial carotenoids and isoprenoid quinones. *J Microbiol Methods* 2012;88:28–34.
59. Smibert RM, Krieg NR. Phenotypic characterization. In: Gerhardt P, Murray R, WA W, Krieg N (editors). *Methods for General and Molecular Bacteriology*. Washington, DC: American Society for Microbiology; 1994. pp. 607–654.
60. Schneider CA, Rasband WS, Eliceiri KW. NIH image to ImageJ: 25 years of image analysis. *Nat Methods* 2012;9:671–675.
61. Ducret A, Quardokus EM, Brun Y V. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nat Microbiol* 2016;1:1–7.
62. Lim JH, Baek SH, Lee ST. *Burkholderia sediminicola* sp. nov., isolated from freshwater sediment. *Int J Syst Evol Microbiol* 2008;58:565–569.
63. Vandamme P, De Brandt E, Houf K, Salles JF, van Elsas JD *et al*. *Burkholderia humi* sp. nov., *Burkholderia choica* sp. nov., *Burkholderia telluris* sp. nov., *Burkholderia terrestris* sp. nov. and *Burkholderia udeis* sp. nov.: *Burkholderia glathei*-like bacteria from soil and rhizosph. *Int J Syst Evol Microbiol*2013;63:4707–4718.
64. Peeters C, Meier-Kolthoff JP, Verheyde B, De Brandt E, Cooper VS *et al*. 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* 2016;7:1–19.
65. Mavengere NR, Ellis AG, Le Roux JJ. *Burkholderia aspalathi* sp. nov., isolated from root nodules of the South African legume *Aspalathus abietina* Thunb. *Int J Syst Evol Microbiol* 2014;64:1906–1912.
66. Partida-Martinez LP, Groth I, Schmitt I, Richter W, Roth M *et al*. *Burkholderia rhizoxinica* sp. nov. and *Burkholderia endofungorum* sp. nov., bacterial endosymbionts of the plant-pathogenic fungus *Rhizopus microsporous*. *Int J Syst Evol Microbiol* 2007;57:2583–2590.
67. Zolg W, Ottow J. *Pseudomonas glathei* sp. nov., a new nitrogen-scavenging rod isolated from acid lateritic relicts in Germany. *Zeitschrift fur Allgemaine Mikrobiol* 1975;15:287–299.
68. Vreeland RH, Litchfield CD, Martin SEL, Elliot E. *Halomonas elongata*, a new genus and species of extremely salt-tolerant bacteria. *Int J Syst Bacteriol* 1980;30:485–495.