Pseudomonas fortuita sp. nov., isolated from the endosphere of a wild yam

Aurelien Carlier, Marine Beaumel, Sandra Moreau, Tessa Acar, Thibault Géry Sana, Margo Cnockaert, Peter Vandamme

INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

RESEARCH ARTICLE

Carlier *et al.*, *Int. J. Syst. Evol. Microbiol.* 2024;74:006395
DOI 10.1099/ijsem.0.006395

MICROBIOLOGY SOCIETY

OPEN ACCESS

Pseudomonas fortuita sp. nov., isolated from the endosphere of a wild yam

Aurelien Carlier[1,2,*], Marine Beaumel[1], Sandra Moreau[1], Tessa Acar[1,2], Thibault Géry Sana[1,†,‡], Margo Cnockaert[2] and Peter Vandamme[2]

Abstract

A Gram-negative, strictly aerobic bacterial strain was isolated from asymptomatic leaf tissue of a wild yam plant. Optimal growth was observed at 28°C and pH 7, and catalase and oxidase activities were detected. Polyphasic taxonomic and comparative genomics revealed that strain LMG 33091^T represents a novel species of *Pseudomonas*. The nearest phylogenetic neighbours of strain LMG 33091^T were *Pseudomonas putida* NBRC 14164^T (with 99.79% 16S rRNA sequence identity), *Pseudomonas alkylphenolica* KL28^T (99.28%) and *Pseudomonas asplenii* (99.07%) ATCC 23835^T. MALDI-TOF MS analysis yielded distinct profiles for strain LMG 33091^T and the nearest phylogenetic neighbours. Average nucleotide identity analyses between the whole genome sequence of strain LMG 33091^T and of the type strains of its nearest-neighbour taxa yielded values below the species delineation threshold and thus confirmed that the strain represented a novel *Pseudomonas* species, for which we propose the name *Pseudomonas fortuita* sp. nov., with strain LMG 33091^T (=GMI12077^T= CFBP 9143^T) as the type strain.

DATA AVAILABILITY

Lists of all public sequence accessions numbers used in the study, as well as phylogenetic trees in Newick format are available in the recherche.data.gouv.fr data repository at https://doi.org/10.57745/BDH4CS. All nucleotide sequence data generated in the study are available in NCBI Genbank under BioProject ID PRJNA909601.

INTRODUCTION

The genus *Pseudomonas* was first described by Migula in 1894 [1] and contains to date 322 validly named species, excluding subspecies and synonymous species (https://lpsn.dsmz.de/genus/pseudomonas, accessed November 2023). *Pseudomonas* are motile, non-spore-forming, Gram-negative rods belonging to the *Gammaproteobacteria*. *Pseudomonas* species have been isolated from a broad range of environments, and are commonly found in various ecological niches such as soil, water, plants and animals [2]. Some *Pseudomonas* species cause disease in animals or plants (e.g. *Pseudomonas aeruginosa* or *Pseudomonas syringae*, respectively), while others harbour useful biocontrol strains or even serve as model organisms for the degradation of xenobiotics or as a chassis for synthetic biology [3].

The taxonomy of the genus *Pseudomonas* has been revised several times to include new data and interpretations based on sequence analysis and phylogeny. More recently, a large effort focused on a systematic sequencing of genomes of *Pseudomonas* type strains, which resulted in the proposition of 43 new *Pseudomonas* species [4]. Sixteen of these novel species had been previously identified as members of the *Pseudomonas putida* group, which Girard and colleagues proposed to partition into 15 subgroups [4]. Several

Author affiliations: [1]LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France; [2]Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, K. L. Ledeganckstraat 35, 9000 Ghent, Belgium.
***Correspondence:** Aurelien Carlier, aurelien.carlier@inrae.fr
Keywords: endosphere; *Pseudomonas*; *putida* group.
Abbreviations: ANI, average nucleotide identity; CTAB, cetyltrimethylammonium bromide; dDDH, digital DNA-DNA hybridization; MALDI-TOF MS, matrix assisted laser desorption ionization - time of flight mass spectrometry; NB, nutrient broth; PBS, phosphate buffered saline; TEM, transmission electron microscopy; TSA, trypticase soy agar.
†Present address: 'Structural Biology of Biofilms' Group, European Institute of Chemistry and Biology (IECB), Pessac, France
‡Present address: Université de Bordeaux, CNRS, Bordeaux INP, CBMN, UMR 5248, F-33600, Pessac, France.
The Genbank/EMBL/DDJB accessions of the whole genome sequence and 16S rRNA gene sequence of strain LMG 33091^T (=GMI12077^T=CFPB 9143^T) are CP114035 and OQ231605, respectively.
Seven supplementary figures and four supplementary tables are available with the online version of this article.

of these subgroups were defined based on genomes obtained from public databases, but could not be formally described or named for lack of publicly available strains.

In the course of our studies on leaf symbiosis of the wild yam *Dioscorea sansibarensis* [5, 6], we isolated a strain of *Pseudomonas* from leaf tips of the host plant that normally harbour the symbiotic bacteria *Orrella dioscoreae*. We provide in the present study the genome sequence and polyphasic taxonomic data for this isolate and assign it as a representative member of a *Pseudomonas putida* subgroup (*Pseudomonas* sp. 12 [4]), which lacked formal description.

ISOLATION AND HABITAT

To investigate if the leaf symbiosis between *D. sansibarensis* and its naturally occurring bacterial symbiont *O. dioscoreae* is facultative or obligate, explants of *D. sansibarensis* were treated with antibiotics *in vitro*, and whole plants were regenerated in sterile conditions as described previously [5, 7]. After 4 weeks of growth *in vitro*, aposymbiotic plants were moved to soil and left to grow under non-sterile conditions in a climate chamber at 25°C with 16 h/8 h day/night. To isolate endophytic bacteria, leaf tips were surface-sterilized for 5 min in a solution of 1% sodium hypochlorite, then in a solution of 70% ethanol, rinsed three times in sterile distilled water and macerated in a sterile solution of 0.4% NaCl. Cell suspensions were serially diluted in sterile PBS, and spread on trypticase soy agar (TSA; Oxoid) at 28°C. Although the plant did not present symptoms of infection, macerates from a surface-sterilized leaf gland yielded copious growth of a single colony type on TSA after 48 h of growth, which was fluorescent under UV light. Strain LMG 33091^T was isolated from a single colony and sub-cultivated on TSA three times before preservation in 15% glycerol at −80°C.

16S rRNA GENE AND PHYLOGENETIC ANALYSES

DNA extracts were prepared by heating 1 µl loop of cell material at 95°C for 15 min in 20 µl lysis buffer containing 0.25% (w/v) SDS and 0.05 M NaOH. Following lysis, 180 µl distilled water was added to the lysate [8]. The 16S rRNA gene sequence was amplified using the forward primer pA 5′-AGAGTTTGATCCTGGCTCAG-3 and the reverse primer pH 5′-AAGGAGGTGATCCAGC-CGCA-3′ [9]. PCR products were generated by using the Promega GoTaq2 PCR kit and purified using the Wizard SV PCR clean-up system (Promega) after which the purified products were sequenced by a commercial company (Eurofins, Germany) using the primer BKL1 5′-GTATTACCGCGGCTGCTGGCA-3′ [10] and the primer *Gamma 5′-CTCCTACGGGAGGCAGCAGT-3′ [11]. Sequence assembly was performed using the Geneious Prime software version 2021.2 and yielded a partial 16S rRNA gene sequence of 1006 bp. Sequence comparison to public databases using the EzBioCloud identification service placed the strain within the *P. putida* group [12]. Subsequently, the genome sequence of strain LMG 33091^T was determined (see below) and a complete 16S rRNA gene sequence was extracted from the whole genome sequence. Full-length 16S rRNA gene sequences of strain LMG 33091^T and of the type strains of all *Pseudomonas* species with valid names as downloaded from the EzBioCloud web server were aligned using MAFFT version 7.475 [13]. The alignment was inspected in MEGA7 [14] and poorly aligned regions were manually removed. Maximum-likelihood phylogenetic analysis was performed using IQ-TREE version 2.2.0.3 [15] with automatic model selection (TPM3u+F+I+G4 chosen based on the Bayesian information criterion) and 1000 SH-aLRT and 1000 ultrafast bootstrap replicates (Fig. 1). The strains with the most similar 16S rRNA gene sequences towards strain LMG 33091^T were the type strains of *P. putida* (with 99.79% 16S rRNA sequence identity), *Pseudomonas alkylphenolica* (99.28%), *Pseudomonas asplenii* (99.07%), *Pseudomonas reidholzensis* (99.00%), *Pseudomonas sichuanensis* (99.14%), *Pseudomonas huaxiensis* (99.14%) and *Pseudomonas donghuensis* (99.00%). All other *Pseudomonas* type strains exhibited 16S rRNA sequence identity values below 99% (data not shown).

GENOME SEQUENCING AND WHOLE GENOME PHYLOGENY

We sought to obtain the genome sequence of strain LMG 33091^T to gain further insights into its taxonomy and physiology. Genomic DNA from overnight cultures of strain LMG 33091^T was extracted using CTAB and phenol-chloroform followed by isopropanol precipitation [16]. Contaminating RNA was digested using RNAse A (Qiagen) at 2 mg ml^{-1}, 5 µl per 100 µl extract and incubated for 1 h at 37°C. DNA was further purified using Sera-Mag Select magnetic beads (Cytiva) according to manufacturer recommendations. DNA quality was checked using 1% agarose gel electrophoresis and quantification was done using a Qubit 3 instrument and kit (Thermo Fisher Scientific). Sequencing libraries were prepared using the Oxford Nanopore Technologies Native Barcoding kit SQK NBD 112.4 and sequenced on an Oxford Nanopore MinION instrument using an FLO-MIN112 flow cell. Base calling was done with Guppy version 6.2.1 in GPU mode on a Windows 10 PC equipped with an NVIDIA GeForce RTX 3080 graphics card and 10 Gb of on-board memory. Reads were filtered with FiltLong with default settings (https://github.com/rrwick/Filtlong) and assembled on a Linux server running Debian 5.16.11–1 (64 cores and 504 Gb of RAM) with Canu version 2.2 [17] with the following options: useGrid, False; corOutCoverage, 100; genomeSize, 6.3 m; enableOEA, false. Contigs were circularized using Circlator version 1.5.5 [18]. Assembly polishing using ONT reads was done using Medaka version 1.7.1. (https://github.com/nanoporetech/medaka). In addition, paired-end 2×150 bp libraries were generated from the same sample using standard protocols at Novogene Europe (Cambridge, UK) and sequenced on an Illumina platform. Illumina sequencing

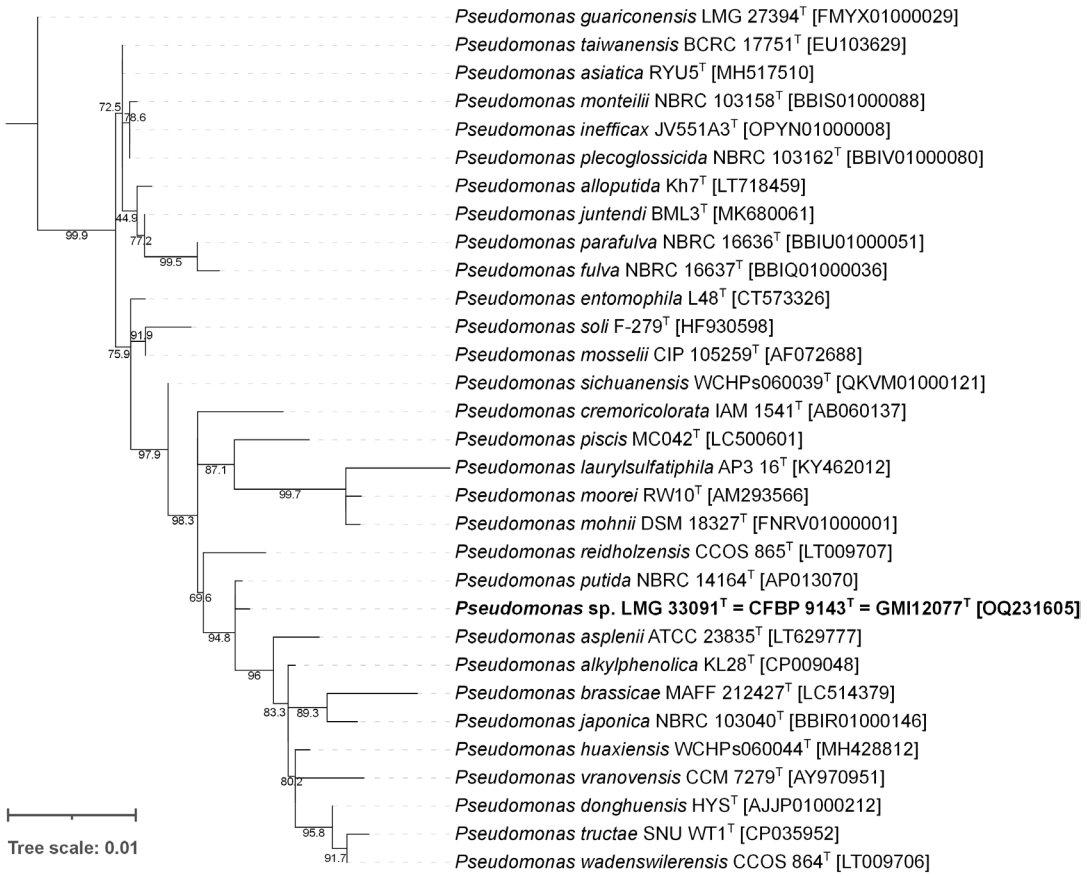

Fig. 1. Detail of a phylogenetic tree based on partial 16S rRNA gene sequences showing the relationships of *Pseudomonas* sp. LMG 33091^T among type strains of *Pseudomonas* species with valid names. For clarity, only a subset of the original phylogenetic tree containing 354 branches is represented. NCBI GenBank sequence accession numbers are given between brackets. Bootstrap percentages are shown next to the branch points. Bar, 0.01 changes per nucleotide position. The best tree topology was inferred using 1000 bootstrapping replications. The tree was rooted using the 16S rRNA gene sequence of *Cellvibrio japonicus* Ueda107 (accession AF452103.1, not shown) [25]. The tree was viewed and edited in iTol [43]. The complete tree can be downloaded in Newick format from the Data INRAE repository (https://doi.org/10.57745/BDH4CS).

reads were trimmed to remove sequencing adapters, deduplicated, and low-quality sequences were filtered using Fastp version 0.22 with default settings [19]. Sequences from contigs assembled from long-read data alone were corrected with Illumina data with Pilon version 1.24 [20]. The genome of strain LMG 33091^T was assembled into a single circular contig of approximately 6.31 Mb, with an average coverage of 26× (Nanopore data) and 299× (Illumina paired-end data). The contig sequence was annotated with the NCBI Prokaryotic Genome Annotation Pipeline version 6.3 [21] and deposited in NCBI Genbank under accession CP114035. Genome statistics are provided in Table S1 (available in the online version of this article). To determine the precise taxonomic status of strain LMG 33091, 521 *Pseudomonas* genome sequences representing the full set of *Pseudomonas* species for which genome data was publicly available as of July 2023 were downloaded using the NCBI datasets command line tool version 14.6.4 [22]. Average nucleotide identity (ANI) values between strain LMG 33091^T and each of the *Pseudomonas* representative genomes were computed using FastANI version 1.33 [23]. ANI computed between strain LMG 33091^T and each of the *Pseudomonas* representative genomes yielded values below 95%, with the highest value given between strain LMG 33091^T and *P. putida* NBRC 14164^T (93.83%) (Table S2). Nucleotide sequences of 92 gene markers were extracted with UBCG [24], and used to generate a maximum-likelihood phylogeny generated using RAxML version 8.2.12 with the GTR +CAT model [25]. This phylogenetic analysis placed strain LMG 33091^T as a distinct lineage within the *P. putida* group, with *P. putida* NBRC 14164^T, *Pseudomonas juntendi* BML3^T, *Pseudomonas monteilii* NBRC 103158^T and *Pseudomonas kurunegalensis* RW1P2^T as the nearest neighbours with valid names (Fig. 2).

PHYSIOLOGY AND CHEMOTAXONOMIC CHARACTERIZATION

Strains LMG 33091^T, *P. juntendi* LMG 33152^T, *P. putida* LMG 2257^T, *P. monteilii* LMG 21609^T and *P. kurunegalensis* LMG 32023^T were cultured on TSA at 28°C for 48 h and cell extracts were used to generate MALDI-TOF mass spectra as described previously

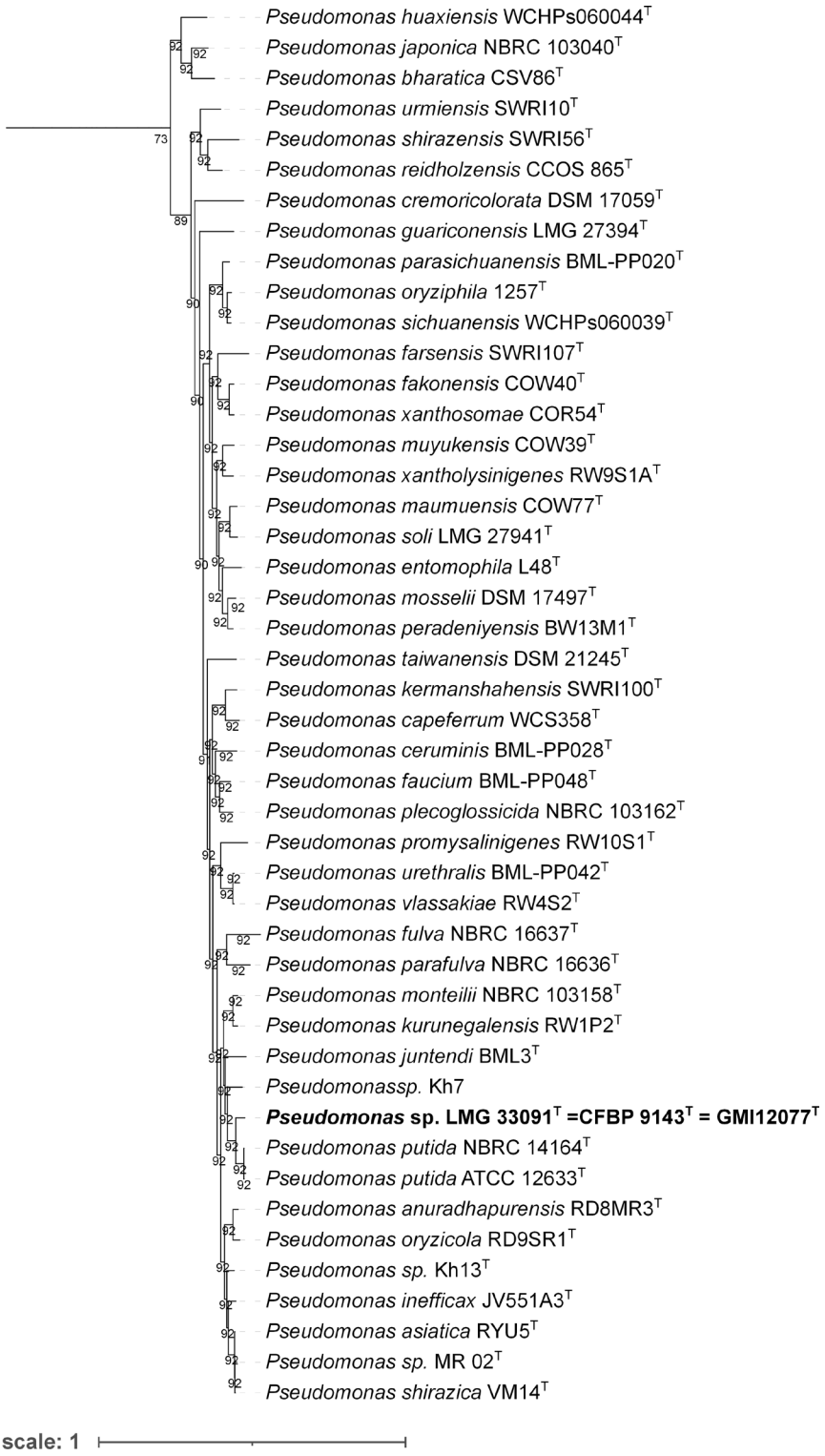

Fig. 2. Phylogenomic tree based on 92 concatenated sequence alignments of gene markers partial showing the relationships of *Pseudomonas* sp. LMG 33091^T among type or representative strains of *Pseudomonas* species with valid names. For clarity, only a detail of the original phylogenetic tree containing 541 branches is represented. Branch support values indicate the number of individual marker gene trees which support the topology, with branch support values of 92 indicating full agreement. Bar, 1 change per nucleotide position. The complete tree in Newick format and the list of NCBI sequence accession numbers of the assemblies used to generate the alignments can be downloaded from the Data INRAE repository (https://doi.org/10.57745/BDH4CS). The tree was viewed and edited in iTol [43].

[26]. All cell extracts were spotted in duplicate and the profiles were visualized with MBT Compass Explorer 4.1 software (Bruker Daltonik) [27]. Each strain displayed clearly distinct MALDI-TOF MS profiles from the other type strains, further confirming the distinct taxonomic status of strain LMG 33091^T (Fig. 3). We further investigated the physiological and metabolic characteristics of *Pseudomonas* sp. LMG 33091^T, *P. juntendi* LMG 33152^T, *P. putida* LMG 2257^T, *P. monteilii* LMG 21609^T and *P. kurunegalensis* LMG 32023^T. These included growth on TSA, TSA supplemented with horse blood, McConkey agar (Oxoid), *Pseudomonas* cetrimide agar (Oxoid) and Drigalski medium (Bio-Rad) under aerobic conditions. Anaerobic nitrate respiration was tested on TSA medium supplemented with or without 10 mM KNO_3 and incubated at 28°C in an anaerobic chamber. Growth at pH 4, 5, 6, 7, 8 and 9 was tested in nutrient broth (NB; Oxoid) using appropriate biological buffers (acetate for pH 4.0–5.0, citrate–Na_2HPO_4 for pH 5.0–6.0, phosphate buffer for pH 6.0–7.0 and Tris–hydrochloride for pH 8.0–9.0); also growth at 4, 15, 20, 28, 37, 40 and 45°C was determined in NB medium. Growth on McConkey agar and activities of catalase, oxidase, DNase, as well as the hydrolysis of starch, casein and Tween 20 and 80, as well as the glucose oxidation–fermentation test were determined using standard methods [28]. Salt tolerance was determined in NB using concentrations ranging from 0 to 10% w/v NaCl in 1% increments. Cell morphology was examined using a light microscope (BX41, Olympus) after growth on TSA at 28°C under aerobic conditions for 48 h. Motility was observed in young cultures by examining wet mounts in broth by phase-contrast microscopy. All five strains tested yielded the same test results (see species description below for details). We further determined cell morphology using transmission electron microscopy (TEM). Briefly, bacterial cells of a 24h-old agar culture of strain LMG 33091^T were picked with a sterile plastic loop and gently suspended in sterile PBS. Samples were centrifuged at 123 g for 20 min, washed in PBS and fixed for 20 min in 50 µl 4% paraformaldehyde in PBS at pH 7.3. To prevent clumping of cells, 0.1% Triton-X 100 (Sigma) was added to the fixative. The cells were then adsorbed onto a Formvar-coated copper single slot grid for 10 min and rinsed twice for 10 s in PBS and once for 10 s in distilled water. Cells were negatively stained with a 2% (w/v) uranyl acetate for 10 s. The excess fluid was removed with a filter paper and the grid was then air-dried. TEM analysis was performed using a Hitachi HT-7700 transmission electron microscope at 80 kV, equipped with an AMT XR-41 CCD camera. Cells appeared as small bacilli (about 1 µm wide and 2–3 µm long) with rounded ends with single polar flagella (Fig. 4).

GENOME FEATURES AND METABOLISM

To refine our comparative genome analysis, we downloaded all publicly available genomes belonging to strains of the *Pseudomonas* putida group in NCBI GenBank, for a total of 638 genomes, including that of strain LMG 33091^T (accessed June 2023, list of accession numbers available from the Data INRAE repository at https://doi.org/10.57745/BDH4CS). ANI values were first computed as above with FastANI, and pairwise comparisons yielding FastANI values>94% were further refined with PyANI version 0.01.12, using the more sensitive NCBI BLASTn+algorithm to generate the alignments [29]. Digital DNA–DNA hybridization (dDDH) values were also computed using the TYGS web service [30]. Twenty-two genome sequences, representing 20 strains, yielded pairwise ANI values >95% and dDDH values >70%, indicating that strain LMG 33091^T is a representative of a group of strains with uncertain taxonomic status (Tables S3 and S4, respectively). Genome sequences of the 20 strains displayed similar

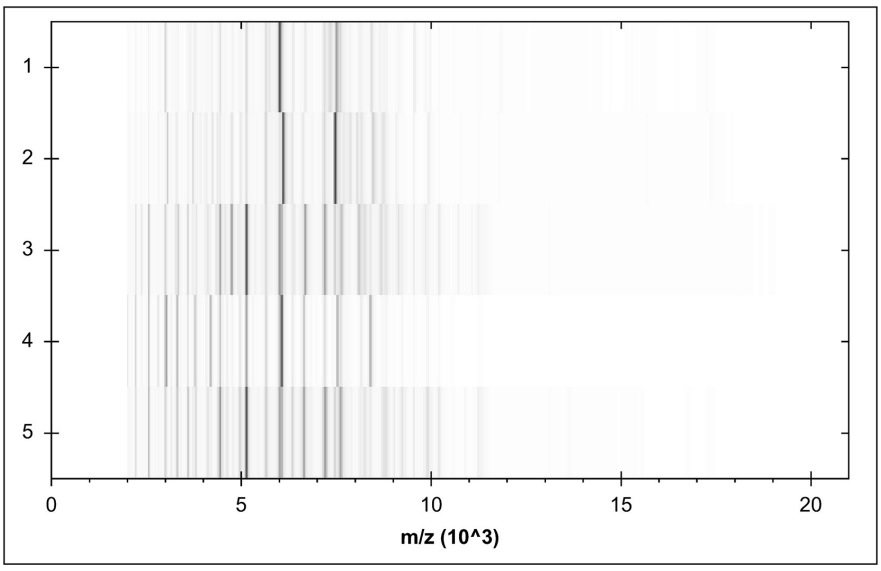

Fig. 3. MALDI-TOF mass spectra of *Pseudomonas* sp. LMG 33091^T (1), *Pseudomonas juntendi* LMG 33152^T (2), *Pseudomonas putida* LMG 2257^T (3), *Pseudomonas monteilii* LMG 21609^T (4) and *Pseudomonas kurunegalensis* LMG 32023^T (5).

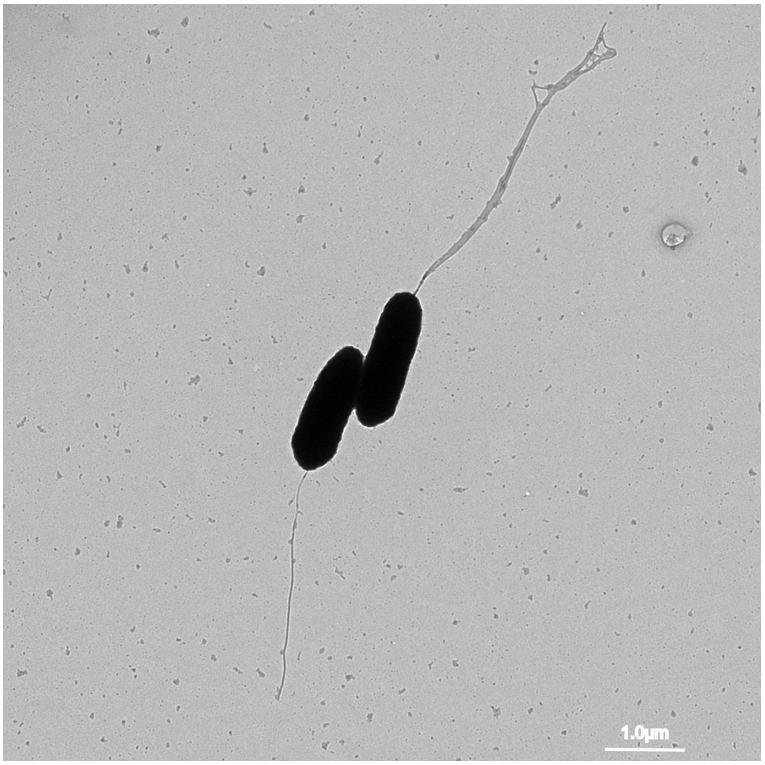

Fig. 4. Transmission electron microscopy image of cells of strain LMG 33091^T.

features, included genome sizes of 6.45 Mb on average (ranging from 5.6 to 7.47 Mb), and an average percentage G+C content of 61.77 mol% (ranging from 61.2 to 62.11 mol%) (Table S1). These strains were isolated from diverse environments, including clinical and environmental samples, but rarely plants (Table S1). This highlights the status of the species represented by strain LMG 33091^T as an opportunistic commensal or pathogen, including as an endophyte of *D. sansibarensis*.

We also analysed putative secretion systems and flagellar genes with Macsyfinder version 2.1.2 with the '--models TXSScan all' parameter [31]. The genome of strain LMG 33091^T was predicted to encode a functional type I, type II and type VI secretion system, flagellar motility, as well as a Type IV tight adherence (tad) pilus involved in biofilm formation in other organisms [32, 33]. No genes linked to Type III secretion systems could be detected. These features were shared with all other high-quality genomes of the same species cluster (Table S1). The core genome of the intra-species dataset (17 strains with high-quality genome sequence data; Table S1) was computed using Orthofinder version 2.5.4 with default settings [34]. The core genome of the putative species included 4195 gene families, for an average number of 5911 genes per genome and a pan-genome size of 7626. The relatively high proportion of core genes suggested that gene inventories are homogenous within the group of strains, further supporting the hypothesis that they belong to a single species. To gain insight into the metabolic pathways shared between strains, we annotated the predicted core genes of strain LMG 33091^T using EggNOG-mapper v2.1.11 [35] and the eggNOG 5 database [36]. We retrieved KEGG functional identifiers (KO numbers) from the EggNOG-mapper output and mapped them to pathways in the KEGG Orthology database using the web service at www.genome.jp/kegg/ko.html [37]. The core genome of the 17 strains contained the full complement of genes required for oxidative phosphorylation (cytochrome c oxidase, NADH dehydrogenase and F-type ATPase), TCA cycle, gluconeogenesis but not glycolysis (with a missing phosphofructokinase enzyme), and the non-oxidative branch of the pentose-phosphate pathway (Fig. S4). A complete benzoate to succinyl-CoA degradation pathway is conserved (Fig. S5), a trait shared with other species of the *P. putida* group [38]. Furthermore, the genome of strain LMG 33091^T encoded several putative secreted enzymes as predicted using the dbCAN3 web server with DIAMOND as the search engine with default e-value cut-off=1e-102 [39]. Enzymes linked to complex carbohydrate metabolism that were also part of the core genome of the 17 strains include: a putative β-1,3 glucanase (Cazy family GH148, locus tag OZ911_15290); an endo-β-1,4-glucuronan lyase (PL5_1, OZ911_23975); a β-agarase (GH50, OZ911_22545); and a β-glucosidase (GH3, OZ911_23140). Assimilatory nitrate and nitrite reduction were predicted for all 17 strains, but dissimilatory nitrate reduction or nitrogen fixation pathways were absent (Fig. S6). Putative biosynthetic genes coding for the siderophore pyoverdin were predicted by antiSMASH [40] and also figure among the core genome (Fig. S7). Three putative antibiotic efflux pumps (locus tags OZ911_13520, OZ911_15405 and

OZ911_23235) were predicted using the online CARD database [41] to provide some level of resistance to fluoroquinolone and tetracycline antibiotics. Orthologs of these three genes were present in all 17 genomes of this novel taxon.

Finally, two putative proteins were conserved in all genomes of the novel taxon, but figured in less than 1% of representative genomes of the genus *Pseudomonas*. Protein WAP63962.1 is small (53 AA), predicted membrane protein containing a DUF1328 domain of unknown function. Protein WAP61180.1 is also a small protein (60AA) containing a conserved KGG motif, homologous to YciG of *E. coli* and *Salmonella* sp. KGG-domain containing proteins are induced by stress in enterobacteria, and may play a role in biofilm formation and resistance to antibiotics in various γ-proteobacteria [42].

DESCRIPTION OF *PSEUDOMONAS FORTUITA* SP. NOV.

Pseudomonas fortuita (for.tu.i'ta., L. fem. adj. *fortuita* accidental, unpremeditated; referring to the species fortuitous isolation).

P. fortuita cells are Gram-negative small bacilli (about 1 µm wide and 2–3 µm long) with rounded ends that occur as single units. Cells collected from 24 h agar cultures displayed single polar flagella (Fig. 4) although motility was not observed. After 48 h incubation on TSA at 28°C, colonies were beige, round, convex and shiny with smooth margins, and 2–3 mm wide. Grows on TSA supplemented with horse blood, McConkey agar, cetrimide agar and Drigalski agar under aerobic conditions. No haemolysis on horse blood agar. Yellow discolouration of McConkey agar. No growth on TSA or TSA supplemented with 10 mM KNO_3 in anaerobic conditions. Growth on DNase, gelatin, skimmed milk, starch and Tween 80 agar bases, but hydrolytic activities were not observed. No growth on Tween 20 agar base. Oxidase and catalase activity were detected. Growth was observed at 28–40°C and pH 6.0–9.0. Tolerance to NaCl was up to 7% (w/v) with optimum growth below 5% (w/v) NaCl. No reduction of nitrate to nitrite, and no production of gas from nitrate.

The type strain is LMG 33091^T (=GMI12077^T=CFBP 9143^T) and was isolated from the endosphere of the yam *Dioscorea sansibarensis* in 2021. Its G+C content is 62.02 mol% (calculated based on genome sequence). The 16S rRNA gene and whole genome sequences are publicly available under accessions OQ231605 and CP114035.2, respectively.

Funding information
A.C. acknowledges support from the French National Research Agency under grant agreement ANR-19-TERC-0004-01 and from the French Laboratory of Excellence project "TULIP" (ANR-10-LABX-41; ANR-11-IDEX-0002-02). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements
We thank Bruno Payré at the 'Centre de Microscopie Electronique Appliquée à la Biologie' (CMEAB), Université Paul Sabatier, Toulouse for his help in the preparation and observation of microscopy samples.

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

References

1. Migula N. *Arbeiten aus dem Bakteriologischen Institut der Technischen Hochschule zu Karlsruhe*. 1894.
2. Hesse C, Schulz F, Bull CT, Shaffer BT, Yan Q, *et al*. Genome-based evolutionary history of *Pseudomonas* spp. *Environ Microbiol* 2018;20:2142–2159.
3. Nikel PI, Martínez-García E, de Lorenzo V. Biotechnological domestication of pseudomonads using synthetic biology. *Nat Rev Microbiol* 2014;12:368–379.
4. Girard L, Lood C, Höfte M, Vandamme P, Rokni-Zadeh H, *et al*. The ever-expanding *Pseudomonas* genus: description of 43 new species and partition of the *Pseudomonas putida* group. *Microorganisms* 2021;9:1766.
5. Acar T, Moreau S, Coen O, De Meyer F, Leroux O, *et al*. Motility-independent vertical transmission of bacteria in leaf symbiosis. *mBio* 2022;13:e0103322.
6. De Meyer F, Danneels B, Acar T, Rasolomampianina R, Rajaonah MT, *et al*. Adaptations and evolution of a heritable leaf nodule symbiosis between *Dioscorea sansibarensis* and *Orrella dioscoreae*. *ISME J* 2019;13:1831–1844.
7. Acar T, Moreau S, Jardinaud M-F, Houdinet G, Maviane-Macia F, ***et al***. The association between *Dioscorea sansibarensis* and *Orrella dioscoreae* as a model for hereditary leaf symbiosis. 2023.
8. Niemann S, Pühler A, Tichy HV, Simon R, Selbitschka W. Evaluation of the resolving power of three different DNA fingerprinting methods to discriminate among isolates of a natural *Rhizobium meliloti* population. *J Appl Microbiol* 1997;82:477–484.
9. Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC. Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res* 1989;17:7843–7853.
10. Cleenwerck I, Camu N, Engelbeen K, De Winter T, Vandemeulebroecke K, *et al*. *Acetobacter ghanensis* sp. nov., a novel acetic acid bacterium isolated from traditional heap fermentations of Ghanaian cocoa beans. *Int J Syst Evol Microbiol* 2007;57:1647–1652.
11. Coenye T, Falsen E, Vancanneyt M, Hoste B, Govan JRW, *et al*. Classification of *Alcaligenes faecalis*-like isolates from the environment and human clinical samples as *Ralstonia gilardii* sp. nov. *Int J Syst Evol Microbiol* 1999;49:405–413.
12. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, *et al*. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.
13. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;30:772–780.
14. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–1874.

15. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, *et al*. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;37:1530–1534.
16. Wilson K. Preparation of genomic DNA from bacteria. *Curr Protocols Mol Biol* 2001;56.
17. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, *et al*. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res* 2017;27:722–736.
18. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, *et al*. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 2015;16:294.
19. Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* 2023;2:e107.
20. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, *et al*. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014;9:e112963.
21. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, *et al*. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016;44:6614–6624.
22. Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, *et al*. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 2021;49:D10–D17.
23. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:5114.
24. Na S-I, Kim YO, Yoon S-H, Ha S-M, Baek I, *et al*. UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J Microbiol* 2018;56:280–285.
25. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–1313.
26. Dumolin C, Aerts M, Verheyde B, Schellaert S, Vandamme T, *et al*. Introducing SPEDE: high-throughput dereplication and accurate determination of microbial diversity from matrix-assisted laser desorption-ionization time of flight mass spectrometry data. *mSystems* 2019;4:e00437-19.
27. Strohalm M, Kavan D, Novák P, Volný M, Havlícek V. mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Anal Chem* 2010;82:4648–4651.
28. Smibert JF, Krieg NR. Phenotypic characterization. In: Gerhardt P, Murray RGE, Wood WA and Krieg NR (eds). *Methods for General and Molecular Bacteriology*. Washington, DC: American Society for Microbiology; 1994. pp. 607–654.
29. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, *et al*. BLAST+: architecture and applications. *BMC Bioinform* 2009;10:421.
30. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 2019;10:2182.
31. Néron B, Denise R, Coluzzi C, Touchon M, Rocha EPC. (2023) MacSyFinder v2: improved modelling and search engine to identify molecular systems in genomes. *Peer Commun J*;3.
32. Pu M, Rowe-Magnus DA. A Tad pilus promotes the establishment and resistance of *Vibrio vulnificus* biofilms to mechanical clearance. *NPJ Biofilms Microbiom* 2018;4:10.
33. Planet PJ, Kachlany SC, Fine DH, DeSalle R, Figurski DH. The widespread colonization island of *Actinobacillus actinomycetemcomitans*. *Nat Genet* 2003;34:193–198.
34. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019;20:238.
35. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 2021;38:5825–5829.
36. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, *et al*. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 2019;47:D309–D314.
37. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 2016;44:D457–D462.
38. Sudarsan S, Blank LM, Dietrich A, Vielhauer O, Takors R, *et al*. Dynamics of benzoate metabolism in *Pseudomonas putida* KT2440. *Metab Eng Commun* 2016;3:97–110.
39. Zheng J, Ge Q, Yan Y, Zhang X, Huang L, *et al*. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res* 2023;51:W115–W121.
40. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, *et al*. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res* 2021;49:W29–W35.
41. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, *et al*. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2023;51:D690–D699.
42. Kaleta MF, Petrova OE, Zampaloni C, Garcia-Alcalde F, Parker M, *et al*. A previously uncharacterized gene, PA2146, contributes to biofilm formation and drug tolerance across the γ-proteobacteria. *NPJ Biofilms Microbiomes* 2022;8:54.
43. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2019;47:W256–W259.