



Rhizobium aouanii sp. nov., efficient nodulating rhizobia isolated from Acacia saligna roots in Tunisia

Jihed Hsouna, Takwa Gritli, Houda Ilahi, Jia-Cheng Han, Walid Ellouze, Xiao Xia Zhang, Maroua Mansouri, Praveen Rahi, Mustapha Missbah El Idrissi, Mouad Lamrabet, et al.

► To cite this version:

Jihed Hsouna, Takwa Gritli, Houda Ilahi, Jia-Cheng Han, Walid Ellouze, et al.. Rhizobium aouanii sp. nov., efficient nodulating rhizobia isolated from Acacia saligna roots in Tunisia. International Journal of Systematic and Evolutionary Microbiology, 2024, 74 (9), pp.006515. <10.1099/ijsem.0.006515>. <hal-05262835>

HAL Id: hal-05262835

<https://hal.inrae.fr/hal-05262835v1>

Submitted on 16 Sep 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Rhizobium aouanii sp. nov., efficient nodulating rhizobia isolated from *Acacia saligna* roots in Tunisia

Jihed Hsouna^{1,2}, Takwa Gritli¹, Houda Ilahi¹, Jia-Cheng Han³, Walid Ellouze⁴, Xiao Xia Zhang³, Maroua Mansouri¹, Praveen Rahi⁵, Mustapha Missbah El Idrissi⁶, Mouad Lamrabet⁶, Pierre Emmanuel Courty⁷, Daniel Wipf⁷, Abdelkader Bekki⁸, James T. Tambong^{9,*} and Bacem Mnasri^{1,*}

Abstract

Three bacterial strains, 1AS14I^T, 1AS12I and 6AS6, isolated from root nodules of *Acacia saligna*, were characterized using a polyphasic approach. Phylogenetic analysis based on *rrs* sequences placed all three strains within the *Rhizobium leguminosarum* complex. Further phylogeny, based on 1756 bp sequences of four concatenated housekeeping genes (*recA*, *atpD*, *glnII* and *gyrB*), revealed their distinction from known rhizobia species of the *R. leguminosarum* complex (Rlc), forming a distinct clade. The closest related species, identified as *Rhizobium laguerreae*, with a sequence identity of 96.4% based on concatenated *recA-atpD-glnII-gyrB* sequences. The type strain, 1AS14I^T, showed average nucleotide identity (ANI) values of 94.9, 94.3 and 94.1% and DNA–DNA hybridization values of 56.1, 57.4 and 60.0% with the type strains of closest known species: *R. laguerreae*, *Rhizobium acaciae* and '*Rhizobium indicum*', respectively. Phylogenomic analyses using 81 up-to-date bacteria core genes and the Type (Strain) Genome Server pipeline further supported the uniqueness of strains 1AS14I^T, 1AS12I and 6AS6. The relatedness of the novel strains to NCBI unclassified *Rhizobium* sp. (396 genomes) and metagenome-derived genomes showed ANI values from 76.7 to 94.8% with a species-level cut-off of 96%, suggesting that strains 1AS14I, 1AS12I and 6AS6 are a distinct lineage. Additionally, differentiation of strains 1AS14I^T, 1AS12I and 6AS6 from their closest phylogenetic neighbours was achieved using phenotypic, physiological and fatty acid content analyses. Based on the genomic, phenotypic and biochemical data, we propose the establishment of a novel rhizobial species, *Rhizobium aouanii* sp. nov., with strain 1AS14I^T designated as the type strain (=DSM 113914^T=LMG 33206^T). This study contributes to the understanding of microbial diversity in nitrogen-fixing symbioses, specifically within *Acacia saligna* ecosystems in Tunisia.

INTRODUCTION

Rhizobium leguminosarum is recognized as the primary nodulating species of the tribe Vicieae [1–3]. Saidi et al. [4] showed that *R. leguminosarum* includes at least two distinct species namely *R. leguminosarum* USDA2370^T and *R. laguerreae* FB206^T. Young et al. [5] analysed 439 whole-genome sequences of bacterial strains previously identified in NCBI database as *R. leguminosarum* and concluded that the strains are too diverse to be classified as a single species and named the group the *R. leguminosarum* complex (Rlc). In addition, 18 genospecies (gsp) were defined plus seven single unique strains that have no close relatives within the Rlc

Author affiliations: ¹Laboratory of Legumes and Sustainable Agrosystems, Centre of Biotechnology of Borj-Cédria, BP 901 Hammam-lif 2050, Tunisia; ²University of Carthage, Faculty of Sciences of Bizerte, Tunis, Tunisia; ³Agricultural Cultural Collection of China, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100080, PR China; ⁴Agriculture and Agri-Food Canada, 4902 Victoria Avenue North, Vineland Station, Ontario, L0R 2E0, Canada; ⁵Institut Pasteur, Université Paris Cité, Biological Resource Center of Institut Pasteur (CRBIP), Paris, France; ⁶Faculty of Sciences, Centre de Biotechnologies Végétale et Microbienne, Biodiversité et Environnement, Mohammed V University in Rabat, Rabat, Morocco; ⁷Agroécologie, Institut Agro Dijon, CNRS, Univ. Bourgogne, INRAE, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France; ⁸Biotechnology of Rhizobia and Plant Breeding Laboratory, Department of Biotechnology, Faculty of Sciences, University of Oran1, Sénia, Algeria; ⁹Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Ontario, K1A 0C6, Canada.

***Correspondence:** James T. Tambong, james.tambong@agr.gc.ca; Bacem Mnasri, mnbacemm@yahoo.com

Keywords: *Acacia saligna*; nodules; *Rhizobium aouanii*; *Rhizobium leguminosarum* complex; rhizobia.

Abbreviations: AM, agar medium; ANI, average nucleotide identity; dDDH, digital DNA–DNA hybridization; gsp, genospecies; ML, maximum-likelihood; YMA, yeast mannitol agar.

The Genbank accession numbers for the *rrs* (OM392036, PP495259, PP495260), *atpD* (OM429377, OM429378, OM429380), *recA* (OM429400, OM429399, OM429403), *glnII* (OM429436, OM429437, OM429448), and *gyrB* (OM429451, OM429452, OM429461) gene sequences; and the draft genome accession numbers are JBAMYD0000000000, JBAMYC0000000000 and JBAMYB0000000000 for strains 1AS14I^T, 1AS12I, and 6AS6, respectively. The bioproject accession number for the draft genomes is PRJNA1070255.

Five supplementary figures and one supplementary table are available with the online version of this article.

006515 © 2024 Crown Copyright



This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

[5]. Ten of the genospecies include the type strains of validly published species. Genospecies E includes *R. leguminosarum* USDA 2370^T, gspR is *R. laguerreae* [4], gspG is *Rhizobium sophorae*, gspC is *Rhizobium ruizarguesonis* [6], gspI is '*Rhizobium indicum*' [7], gspS is *Rhizobium changzhienae* [8], gspA is *Rhizobium brockwellii*, gspB is *Rhizobium johnstonii*, gspD is *Rhizobium beringeri* [9] and gspJ is *Rhizobium acaciae* [10]. This classification suggests the presence of several other clades within the Rlc that equally deserved formal species designation due to their distinct genetic makeup.

Hsouna et al. [11] reported that *Acacia saligna* plants grown in Tunisian soils were nodulated by fast-growing organisms, suggesting two putative new species within the genus *Rhizobium*. The authors published a new species (*R. acaciae*) within the genus *Rhizobium*, represented by strains 1AS11^T, 1AS12, and 1AS13 [10]. The other potential new lineage, comprising the three strains 1AS14I^T, 1AS12I and 6AS6, remained undescribed. Analysis of 1734 bp concatenated sequences of *recA*, *atpD*, *glnII* and *gyrB* genes placed these three strains in a distinct clade within the *R. leguminosarum* complex (Rlc), with *R. laguerreae* being the closest described species [11]. Interestingly, based on the divergence of *nodA/nodC* sequences and host range nodulation patterns, a criterion previously used for symbiovar delineation [12], strains 1AS14I^T, 1AS12I and 6AS6 together with strains of *R. acaciae* (1AS11^T, 1AS12, and 1AS13) were classified as a new symbiovar named *saligna* [11]. The representative strain 1AS14I^T was found to nodulate *A. saligna*, *A. salicina*, *Leucaena leucocephala*, but not *Glycine max*, *Phaseolus vulgaris* or *Retama raetam*, highlighting its unique symbiotic properties [11]. Despite this, the exact taxonomic position of these three strains remained unclear. This study presents comprehensive phenotypic, biochemical, genomic and phylogenomic data for strains 1AS14I^T, 1AS12I and 6AS6, underscoring their distinctiveness and suggesting their classification as a new taxon for which the name *Rhizobium aouanii* sp. nov. is proposed.

METHODS

Isolation and ecology

Three novel *Rhizobium* strains were isolated from the root nodules of *A. saligna* plants grown in Tunisia, following the method described by Vincent [13]. Strains 1AS14I^T and 1AS12I were obtained from Borj Cedria (36° 42' 35.5" N 10° 25' 42" E), located in northern Tunisia, while strain 6AS6 was isolated from Monastir (35° 40' 55.6" N 10° 19' 59.9" E), situated in the central region of Tunisia. Single colonies of the bacterial cultures were obtained by repetitively streaking on yeast mannitol agar (YMA), as previously described by Hsouna et al. [11]. The type strain 1AS14I^T was deposited in the BCCM/LMG Bacteria Collection, University of Ghent, Belgium (LMG 33206^T) and in the DSMZ-German Collection of Microorganisms and Cell Cultures, Leibniz Institute (DSM 113914^T).

rrs, *atpD*, *recA*, *glnII* and *gyrB* phylogeny

DNA extraction from strains 1AS14I^T, 1AS12I and 6AS6 was performed following the method described by Terefework et al. [14]. The 16S rRNA gene sequence (1245 bp) and partial sequences of the *atpD*, *recA*, *glnII* and *gyrB* genes were amplified as previously described by Hsouna et al. [10]. PCR-amplified products were purified from agarose gels using the GeneJET Gel Extraction Kit (ThermoScientific) following the manufacturer's instructions. Purified DNA fragments were sequenced at the Génome Québec Innovation Centre, Montreal, Canada. Sequences of the 16S rRNA, *atpD*, *recA*, *glnII* and *gyrB* housekeeping genes were used for phylogenetic analyses. Sequence similarities in the DNA databases were performed using the BLAST program (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Reference sequences of 16S rRNA and the housekeeping genes were obtained from Genbank (NCBI). Sequences were aligned using the ClustalW2 software (www.ebi.ac.uk/Tools/clustalw2/). Phylogenetic trees were reconstructed using the MEGA 10 software package by the maximum-likelihood, neighbour-joining and minimum-evolution statistical methods [15], using the Kimura two-parameter substitution model [16]. Bootstrap values were based on 1000 replications and values >50% are shown on the phylogenetic trees.

Genomic relatedness and phylogenomics

Draft whole-genome sequences were generated using Illumina NovaSeq 6000 (for strains 1AS14I^T and 1AS12I) or MiSeq (for strain 6AS6) technology (Génome Québec, Montreal, Canada). The quality of the raw reads [2×150 bp (NovaSeq) or 2×250 (MiSeq)] was checked using the FastQC algorithm [17]. Whole-genome sequence reads were assembled using Unicycler [18] and contigs <300 bp long were discarded, as implemented in BV-BRC pipeline [19]. The quality and completeness of the genome were evaluated using CheckM [20]. Annotations and protein function assignments were implemented in BV-BRC [19] or Prokka [21]. Average nucleotide identity (ANI [22]) and digital DNA–DNA hybridization (dDDH [23]) values were computed through pairwise comparisons between 1AS14I^T and type strains of closest related *Rhizobium* species. Phylogenomic analyses were done using Type (Strain) Genome Server (TyGS) [24]. The up-to-date bacterial core gene set pipeline of 81 single-copy genes (UBCG2 [25]) was used to validate the clustering patterns of the genome sequences and strains of *Rhizobium aouanii* sp. nov. identified on the TyGS-derived evolutionary tree. The UBCG-based tree was generated using RAXML-NG [26] with 1000 non-parametric bootstrap replicates. In addition, the type strain of *R. aouanii* sp. nov., strain 1AS14I^T, was compared to genome sequences retrieved from GenBank as unclassified *Rhizobium* sp. (396 genomes) or uncultured metagenome-derived *Rhizobium* sp. (55 genomes) using ANI, dDDH and phylogenomics as described above.

In addition, we implemented the GTDB-Tk pipeline version 2.4.0 ([27] <https://github.com/Ecogenomics/GTDBTk>) to compare genome sequences of *R. aouanii* sp. nov. against the most recent GTDB database (release220) [28] consisting of over 12000 bacterial and archaeal genomes from isolates and metagenomes derived from different environments. The `classify_wf` function of GTDB-Tk was used, which implements a workflow consisting of ANI screening with `skani` [29], `Prodigal` [30] with HMM models [31], and `HMMER` [32] to identify 120 bacterial and 53 archaeal marker genes, and `pplacer` [33] to find the maximum-likelihood placement of each genome in the GTDB-TK reference tree. `FastTree` [34] was used to infer the approximately maximum-likelihood trees for large alignments.

Physiology and chemotaxonomy

Colony morphology was examined on YMA medium, and the phenotypic and physiological features of the three strains were determined and compared with their most related type strains: *R. acaciae* 1AS11^T, '*R. indicum*' JKLM 12A2^T, and *R. laguerreae* FB206^T. Characteristics tested included carbon sources utilization, tolerance to NaCl, pH, resistance to antibiotics, enzymatic activities, and fatty methyl ester content. Various phenotypic tests, including 70 carbon source utilization and 18 chemical sensitivity assays, were carried out using Biolog GEN III MicroPlates following the manufacturer's instructions. Enzymatic activities were assessed using the API ZYM systems (07584D and 25200, bioMérieux). Antibiotic resistance testing was checked on Yeast Extract Mannitol (YEM) plates supplemented with five concentrations of each antibiotic: streptomycin, kanamycin, neomycin, and erythromycin. Tolerance to NaCl, pH, and temperature was monitored in liquid YEM medium as previously reported by Hsouna et al. [10]. The optimum pH for growth was assessed on YMA medium plates adjusted from pH 4 to 10 using 1 M HCl (pH 4.0 and 5.0) or 1 M NaOH (pH 7 to 10 prior to autoclaving). Fatty acid methyl esters analysis was performed using the Sherlock Microbial Identification System version 6.1 (MIDI) and a TSBA6 database.

RESULTS AND DISCUSSION

rrs, *atpD*, *recA*, *glnII* and *gyrB* phylogeny

The phylogenetic analysis based on partial *rrs* gene sequences (1215 nt) revealed that all three new strains clustered together, showing 100% similarity to type strains of *R. laguerreae*, *R. changzhiense*, *R. anhuiense*, *R. leguminosarum*, *R. ruizarguesonis*, *R. acidisoli*, *R. hidalgonense*, '*R. indicum*', *R. sophorae*, *R. redzepovicii*, *R. acaciae*, *R. brockwellii*, *R. johnstonii* and *R. beringeri*. These species are classified within the Rlc, with the exceptions of *R. anhuiense*, *R. acidisoli*, *R. hidalgonense*, and *R. redzepovicii*. The maximum-likelihood phylogenetic tree in Fig. 1 shows the clustering of the three strains of the proposed *R. aouanii* sp. nov. with the type strains of the Rlc. Similar tree topologies were obtained using the neighbour-joining (Fig. S1, available in the online Supplementary Material) and minimum-evolution (Fig. S2) algorithms. Given the limited taxonomic resolution at the species level obtained from the 16S rRNA gene sequences [35, 36], multilocus sequence analysis of housekeeping genes was used to elucidate the *Rhizobium* species more precisely. Phylogenetic analysis based on concatenated *atpD-recA-glnII-gyrB* sequences (1734 bp) using the maximum-likelihood statistical method showed that the three strains of the proposed *R. aouanii* sp. nov formed a distinct cluster separate from the type strains of previously described *Rhizobium* species of the Rlc group (Fig. 2).

Genomic relatedness and phylogenomics

Table 1 shows the basic statistics of the sequenced draft genomes. Strains 1AS14I^T and 1AS12I had slightly over 30 million raw NovaSeq reads, while strain 6AS6 had about 1.35 million raw MiSeq reads (Table 1). After quality checks, the reads used in the assembly process resulted in average coverage depths of 1376×, 1318× and 85×, with contig counts of 68, 67 and 58, for strains 1AS14I^T, 1AS12I and 6AS6, respectively. The completeness of the genomes was 100%, with coarse and fine consistencies of 99.4 or 99.9 and 96.7%, respectively, and a contamination rate ranging from 0.2 to 0.5%. The N_{50} values were 288843 (1AS14I^T or 1AS12I) and 314702 (6AS6) bases, with an L_{50} of 9 for each strain (Table 1). The total length of the draft genomes of 1AS14I^T, 1AS12I and 6AS6 were 7110981, 7111185 and 7107636 nucleotides, respectively, with a similar G+C content of 60.93 mol% (Table 1). A total of 4446 (1AS12I) or 4451 (1AS12I or 6AS6) protein-encoding genes with functional assignment were identified with 44 tRNA, three rRNA and one tmRNA each (Table 1). The newly generated draft whole-genome sequences were used to confirm the uniqueness of these strains relative to the closest related *Rhizobium* species.

Genome-based dDDH and ANI values were computed and compared to eight validly published close *Rhizobium* species to validate the taxonomic position of the strains. The three strains (1AS14I^T, 1AS12I and 6AS6) exhibited pairwise dDDH and ANI values of 99.7 and 100%, respectively. These values are significantly above the species-level cut-offs of 70 and 96%, respectively, confirming their classification within the same *Rhizobium* species. Strain 1AS14I^T exhibited low dDDH and ANI values, falling below species delineation cut-off of 70 and 96%, respectively, with known *Rhizobium* species (Table 2). The closest *Rhizobium* species identified were '*R. indicum*' JKLM 12A2^T, *R. acaciae* 1AS11^T and *R. laguerreae* FB206^T (Table 2). For example, '*R. indicum*' JKLM 12A2^T showed the highest dDDH of 60.0% (cut-off of 70%) and ANI of 94.95% (cut-off of 96.0%) with strain 1AS14I^T followed by *R. acaciae* 1AS11^T with values of 57.4 and 94.08% (Table 2). These findings confirmed the results of the *recA-atpD-gyrB-glnII* phylogenetic analysis, suggesting that strains 1AS14I^T, 1AS12I and 6AS6 represent a novel species within the genus *Rhizobium*.

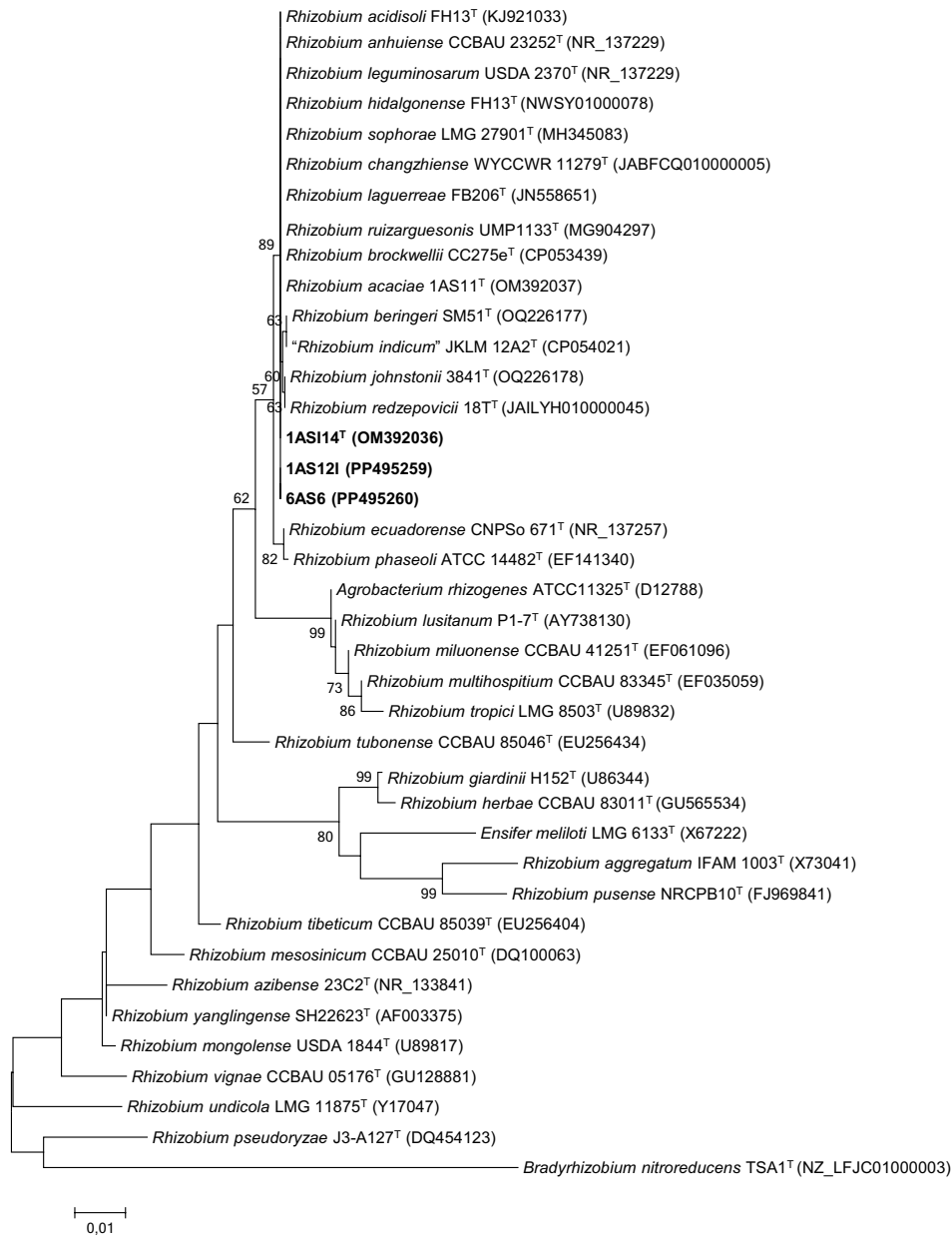


Fig. 1. Maximum-likelihood-based phylogenetic tree of *rrs* sequences (1215 nucleotides) showing the *Rhizobium aouanii* strains clustering within the *Rhizobium leguminosarum* complex. The three novel strains are in bold. Bootstrap values ≥ 50 are indicated at the node (1000 replicates). After each species name, the strain code followed by the NCBI accession number are provided. The scale indicates the number of substitutions per site.

The dDDH and ANI results were further supported by phylogenomics tree inferred using theTyGS [24] based on Genome BLAST Distance Phylogeny (Fig. S4). In addition, the uniqueness of the strains 1AS14^T, 1AS12I and 6AS6 was validated by the 81 UBCG pipeline [25] (Fig. 3). Both TyGS and UBCG generated phylogenomic trees showing that strains 1AS14^T, 1AS12I and 6AS6 clustered distinctly from known species of the genus *Rhizobium* (Figs 3 and S3). These genome data and analyses strongly indicate that these strains represent a potential new species in the genus *Rhizobium*.

Additionally, the genome data obtained in this study were utilized to investigate the strains' relatedness to the proposed 18 genospecies groupings (gsp) of Young *et al.* [5]. A maximum-likelihood phylogenomics tree generated using the UBCG tool showed strains 1AS14^T, 1AS12I and 6AS6 clustering distinctly from the representative genomes of the 18 genospecies proposed by Young *et al.* [5] (Fig. 4). This tree suggests that the *R. aouanii* sp. nov. strains constitute a unique genospecies. Furthermore, Young *et al.* [5] identified seven unique strains (WSM1689, CCBAU10279, CC278f, Norway, WYCCWR10014, Tri-43 and Vaf12) that could not be associated to any of the defined 18 genospecies. A comparison of the genomes of these seven unique strains to

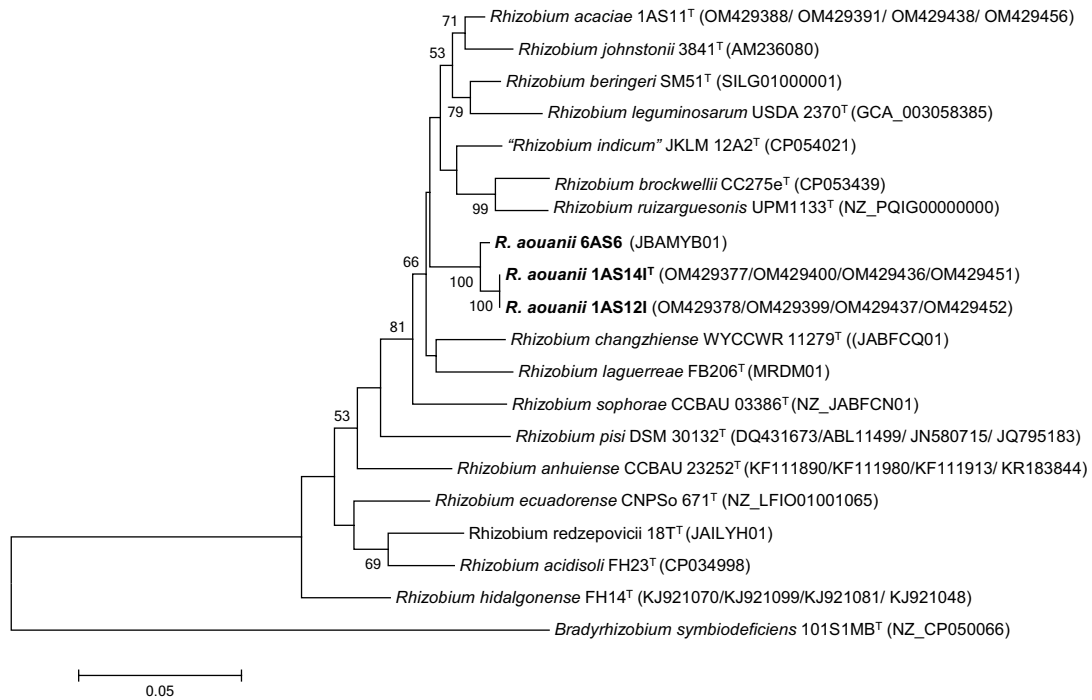


Fig. 2. Maximum-likelihood-based evolutionary tree inferred from the concatenated *atpD-recA-glnII-gyrB* sequences (1734 nt). The three novel strains are in bold. Bootstrap values ≥ 50 are indicated for each node (1000 replicates). The scale indicates the number of substitutions per site.

R. aouanii sp. nov. strains revealed them to be highly divergent. For example, significantly low dDDH values ranging 48.20–60.5% were recorded between the seven strains and the type strain of *R. aouanii* sp. nov., 1AS141^T. These values which are significantly below the species delineation threshold of 70% suggest these strains represent distinct potential species that significantly differ from members of *R. aouanii* sp. nov.

Table 1. Whole-genome sequence statistics of the three strains of *Rhizobium aouanii* sp. nov. sequenced using Illumina NovaSeq 6000 or MiSeq technology

Description	1AS141 ^T	1AS12I	6AS6 ⁺
GenBank accession	JBAMYD01	JBAMYC01	JBAMYB01
Number of short paired end reads (151 or 250 bp length)	33112583	31707554	1349589
Average coverage depth	1376.8×	1318.7×	85.65×
Contig count	68	67	58
Largest contig (bp)	709429	709429	709429
Total length (bp)	7110981	7111185	7107636
Contigs N ₅₀ (bp)	288843	288843	314702
Contigs L ₅₀	9	9	8
Protein-encoding genes with functional assignment	4446	4451	4451
Total protein-coding sequences (CDSs)	7289	7294	7300
ncRNAs	4	4	4
rRNA (5S, 16S, 23S)	1, 1, 1	1, 1, 1	1, 1, 1
Number of tRNA	44	44	44
Number of tmRNA	1	1	1
G+C content (mol%)	60.93	60.93	60.93

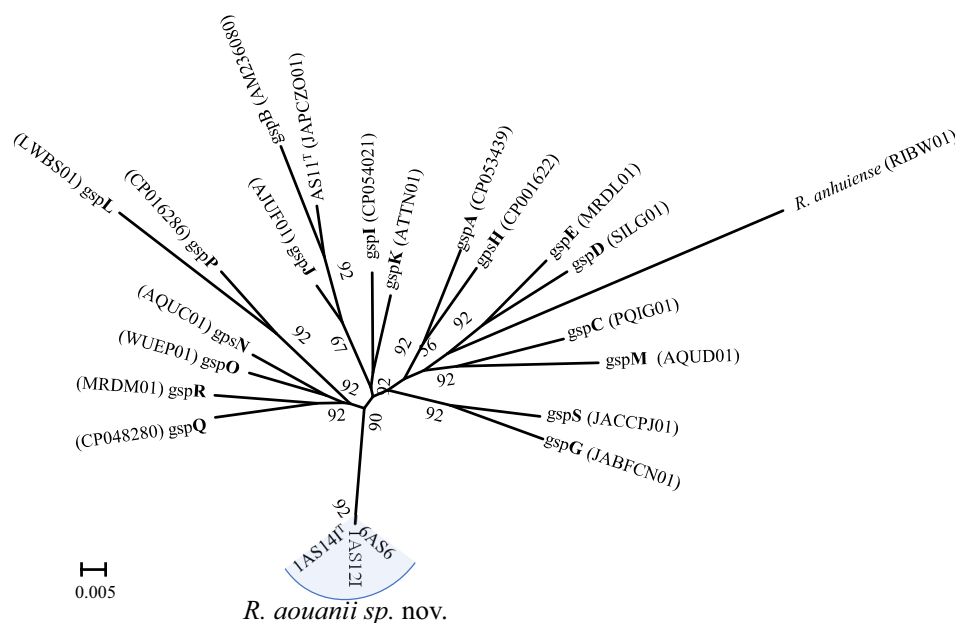


Fig. 4. Up-to-date bacterial core genes (UBCG) maximum-likelihood-based phylogenetic tree of strains of *Rhizobium aouanii* sp. nov. and representatives of the 18 genospecies proposed (letters in bold by Young et al. [5] suggests a new genospecies. The tree was inferred using RAXML-ng with the GTR +CAT model, and percentage bootstrap values are given at branching points.

close genera *Ectorhizobium*, *Agrobacterium*, *Neorhizobium* and *Allorhizobium*. Strains of *R. aouanii* sp. nov. clustered in group III together with 70 other *Rhizobium* strains and metagenome-derived genomes with *R. laguerreae* (=gspR) and a GTDB-defined taxon, *s_R. leguminosarum_T* (=gspQ), as the closest phylogenetic relatives (Fig. 5). This confirmed that strains 1AS14I, 6AS6 and 1AS12I constitute a novel lineage.

The *R. aouanii* sp. nov. strains were also evaluated against NCBI GenBank unclassified *Rhizobium* sp. (396 genomes) and the results show ANI values that are below the species-level cut-off and ranged from 76.7 to 94.8%, suggesting that strains 1AS14I^T, 1AS12I and 6AS6 belong to a distinct lineage. We also investigated the relatedness between *R. aouanii* sp. nov. and metagenome-derived genomes. Of the 55 uncultured metagenome-derived *Rhizobium* sp. sequences, 15 were removed from the analysis due to potential contamination or had a genome size that is less than 4.0 Mb. The remaining 40 metagenome-derived genome sequences had ANI or dDDH values of 77.8–81.9% and 19.9–23.1% with strain 1AS14I^T, which are below the species delineation threshold of 96% and 70%, respectively (Table S1). This affirmed the uniqueness of strains of the proposed *R. aouanii* sp. nov. The phylogenomic tree (Fig. S5) inferred using the 40 metagenome-derived genome sequences grouped the genomes into six distinct clusters. Cluster 1 regrouped the strains of *R. aouanii* sp. nov. with five metagenome-derived sequences: DAMCIK01 (urban metagenome, USA), CALLIZ01 (metagenome), JBBNAU01 (forest soil metagenome), DALZSY01 (urban metagenome) and SSFH01 (wastewater metagenome, USA) but the topology and branch lengths indicate significant genetic divergence. This is corroborated by low ANI and dDDH values (Table S1). Most of the metagenome-derived genomes analysed were obtained either in the USA or Europe with no entries from African countries including Tunisia. Also, the source materials, e.g. wastewater metagenome, of these MAG genomes are profoundly different from that of the strains of *R. aouanii* sp. nov. isolated from root nodules of *A. saligna* in Tunisia. Any similarities or differences between these taxa, however, may be overemphasized/overestimated given that metagenome-derived genome sequences are known to have quality and other issues including differences in genome sizes. For example, one metagenome-derived genome sequence deposited in GenBank as *Rhizobium* sp. has a genome size of 2.4 Mb instead of the required ca 7.1 Mb. A small genome may lack key genes required for bacterial classification. In fact, only 10 of these 40 metagenome-derived genomes had a copy of the 16S rRNA gene, a key genus-level marker. The metagenome-derived genome era has started but to effectively harness its usefulness in bacterial diversity and taxonomy, new sequencing methods to obtain very long reads coupled with smart bioinformatic tools to sort and assemble only fragments of the same molecule are key requirements.

Physiology and chemotaxonomy

Colonies of strains 1AS14I^T, 1AS12I and 6AS6 were typical rhizobial-like translucent, cream-white, and convex. These strains tolerated up to 0.8% NaCl, although none of them were able to grow at a concentration higher than 1%. All strains grew at pH levels from 5 to 10 and at temperatures ranging from 16 to 37°C, with optimal growth at pH 7 and 28°C. Carbon substrate



Fig. 5. Subtree of GTDB-TK Bac120 protein set showing relatedness of strains of *Rhizobium aouanii* sp. nov. (in red) and the 70 whole-genome sequences of members of the *Rhizobium* group III generated using the classify_wf command and the most recent database (release220) of the Genome Taxonomy Database (GTDB-TK) pipeline version 2.4.0.

utilization patterns (Table 3) shared among all three strains of *R. aouanii* sp. nov. differed from at least one of their closest related type strains. Particularly, these strains could assimilate methyl β -D-glucoside, D-fucose, L-rhamnose, D-arabitol, myo-inositol, glycerol, L-serine, β -hydroxy-D,L butyric acid, acetoacetic acid and acetic acid. However, they were unable to assimilate trehalose, gentiobiose, sucrose, turanose, stachyose, raffinose, N-acetyl-D-glucosamine, N-acetyl- β -D-mannosamine, α -D-glucose, D-mannose, D-glucose-6-PO₄, D-fructose-6-PO₄, L-pyroglutamic acid, D-galacturonic acid, D-gluconic acid, D-glucuronic acid, methyl pyruvate, L-malic acid and bromo-succinic acid. Other characteristics shown by the strains of the novel species that differed from at least one of these closest type strains were their positive enzymatic reactions for esterase, trypsin, acid phosphomonoesterase, β -galactosidase, α -galactosidase and N-acetyl- β -glucosaminase; and negative reaction for β -glucosidase. *Rhizobium aouanii* sp. nov. strains were resistant to lincomycin, 10 μ g ml⁻¹ kanamycin, 20 μ gml⁻¹ streptomycin and 20 μ gml⁻¹ neomycin but sensitive to 20 μ gml⁻¹ erythromycin. Analysis of fatty acid profiles revealed distinctive compositions, with notable differences in the amounts of certain fatty acids compared to related strains. The major fatty acids of the type strain 1AS141^T were summed feature 8 (C_{18:1} ω 7c; 48.68%), C_{18:1} ω 7c 11-methyl (12.9%) and C_{19:0} cyclo ω 8c (12.4%) (Table 4). The amount of C_{18:1} ω 7c 11-methyl was higher than that in '*R. indicum*' JKLM 12A2^T and *R. laguerreae* FB206^T and the amount of C_{19:0} cyclo ω 8c was higher than that in *R. acaciae* 1AS11^T, '*R. indicum*' JKLM 12A2^T and *Rhizobium laguerreae* FB206^T. We also detected trace amounts of C_{15:0} 3OH (1.24%), C_{17:1} ω 8c (1.06%), C_{17:1} ω 6c (0.15%), C_{17:0} 3OH (1.24%), C_{17:1} anteiso ω 9c (0.25%) and C_{18:1} ω 9c (0.79%) that were absent in JKLM 12A2^T and FB206^T.

Table 3. Phenotypic characteristics on Biolog GENIII MicroPlates and API ZYM for strains 1AS141^T, 1AS12I and 6AS6 of *Rhizobium aouanii* sp. nov. and related strains (1) '*Rhizobium indicum*' JKLM 12A2^T, (2) *Rhizobium laguerreae* strain FB206^T; (3) *Rhizobium acaciae* 1AS11^T

Characteristic	<i>Rhizobium aouanii</i> sp.nov					
Biolog substrate	1AS141 ^T	1AS12I	6AS6	1	2	3
Trehalose	–	–	–	w	–	+
Gentiobiose	–	–	–	–	–	+
Sucrose	–	–	–	+	–	–
Turanose	–	–	–	w	–	+
Stachyose	–	–	–	–	–	+
Raffinose	–	–	–	–	–	+
Methyl β-D-glucoside	+	+	+	–	+	+
N-Acetyl-D-glucosamine	–	–	–	–	+	+
N-Acetyl-β-D-mannosamine	–	–	–	–	+	+
α-D-Glucose	–	–	–	–	–	+
D-Mannose	–	–	–	–	–	+
D-Fucose	+	+	+	–	+	+
L-Rhamnose	+	+	+	–	–	+
D-Arabitol	+	+	+	–	–	+
myo-Inositol	+	+	+	–	–	+
Glycerol	+	+	+	–	+	+
D-Glucose-6-PO ₄	–	–	–	–	+	–
D-Fructose-6-PO ₄	–	–	–	+	+	–
L-Pyroglutamic acid	–	–	–	–	+	+
L-Serine	+	+	+	–	–	–
D-Galacturonic acid	–	–	–	+	+	–
D-Gluconic acid	–	–	–	–	+	+
D-Glucuronic acid	–	–	–	+	+	–
Methyl pyruvate	–	–	–	–	–	+
L-Malic acid	–	+	+	–	+	–
Bromo-succinic acid	–	+	+	–	–	–
β-Hydroxy-D,L Butyric Acid	+	+	+	–	+	–
Acetoacetic acid	+	+	+	–	–	–
Acetic acid	+	+	+	w	–	–
Chemical sensitivity						
pH 6	+	+	+	–	–	+
pH 5	+	+	+	–	–	+
NaCl 1%	–	–	–	–	+	–
Tetrazolium violet	+	+	+	–	+	+
Tetrazolium blue	+	+	+	–	–	+
Potassium tellurite	+	+	+	–	–	+
Antibiotic resistance						
Lincomycin	+	+	+	–	–	–

Continued

Table 3. Continued

Characteristic	<i>Rhizobium aouanii</i> sp.nov					
	1AS14I ^T	1AS12I	6AS6	1	2	3
Kanamycin (10 µg ml ⁻¹)	+	+	+		–	+
Erythromycin (20 µg ml ⁻¹)	–	–	–		w	w
Streptomycin (20 µg ml ⁻¹)	+	+	+		–	–
Neomycin (20 µg ml ⁻¹)	+	+	+		–	–
API ZYM						
Esterase	+	+	+	–	–	ND
Trypsin	+	+	+	w	w	ND
Acid phosphomonoesterase	+	+	+	–	–	ND
β-Galactosidase	+	+	+	w	w	ND
α-Galactosidase	+	+	+	+	–	ND
β-Glucosidase	–	–	–	w	+	ND
N-Acetyl-β-glucosaminase	+	+	+	–	–	ND

+, Positive; –, negative; w, weakly positive; ND, not determined.

Based on the genotypic and phenotypic data presented in this study, we propose that strains 1AS14I^T, 1AS12I and 6AS6 represent a novel species, named *Rhizobium aouanii* sp. nov.

DESCRIPTION OF *RHIZOBIUM AOUANII* SP. NOV.

Rhizobium aouanii (a.ou.a'ni.i. N.L. gen. masc. n. *aouanii*, of Aouani, to honour Dr. Mohamed Elarbi Aouani, a Tunisian microbiologist who made significant contributions to research on rhizobia).

Cells of *Rhizobium aouanii* are Gram-stain-negative, aerobic, and non-spore-forming. When cultured on YMA medium, colonies appear translucent, cream-white and convex after 3 days of growth at 28 °C. The strains tolerated up to 0.8% NaCl but did not grow at concentrations >1%; and grew at pH levels from 5 to 10 and at temperatures ranging from 16 to 37 °C, with optimal growth at pH 7 and 28 °C. Strain 1AS14I^T demonstrated the ability to utilize various carbon sources, including methyl β-D-glucoside, D-fucose, L-rhamnose, D-arabitol, *myo*-inositol, glycerol, L-serine, β-hydroxy-D,L-butyric acid, acetoacetic acid and acetic acid; but is unable to use trehalose, gentiobiose, sucrose, turanose, stachyose, raffinose, N-acetyl-D-glucosamine, N-acetyl-β-D-mannosamine, α-D-glucose, D-mannose, D-glucose-6-PO₄, D-fructose-6-PO₄, L-pyrogutamic acid, D-galacturonic acid, D-gluconic acid, D-glucuronic acid, methyl pyruvate, L-malic acid, and bromo-succinic acid. Furthermore, strain 1AS14I^T is resistant to lincomycin, kanamycin (10 µg ml⁻¹), streptomycin (20 µg ml⁻¹) and neomycin (20 µgml⁻¹), but sensitive to erythromycin (20 µg ml⁻¹). It is also capable of forming effective nitrogen-fixing nodules on *Acacia saligna*. The predominant fatty acid of the type strain 1AS14I^T is summed feature 8 (C_{18:1} ω7c).

Table 4. Fatty acid composition (%) of strain 1AS14I^T and related strains

Strains: 1, 1AS14I^T; 2, *Rhizobium acaciae* 1AS11^T; 3, '*Rhizobium indicum*' JKLM 12A2^T; 4, *Rhizobium laguerreae* FB206^T. –, Not detected.

Fatty acid	1	2	3	4
C _{15:0} 3OH	1.24	0.27	–	–
C _{17:0}	4.18	0.78	–	–
C _{17:1} ω8c	1.06	0.26	–	–
C _{17:0} 3OH	1.24	0.16	–	–
C _{18:1} ω7c 11-methyl	12.9	13.24	1.8	1.9
C _{19:0} cyclo ω8c	12.4	3.0	3.9	6.4
Summed feature 8 (C _{18:1} ω7c)	48.68	57.81	68	69.3

The type strain of *Rhizobium aouanii* sp. nov. 1AS141^T (=DSM 113914^T = LMG 33206^T), was isolated from root nodules of *Acacia saligna* used as trapping host in Borj cedria soil located in northeastern Tunisia. The genome size of strain 1AS141^T is approximately 7.11 Mb with a G+C content of 60.93 mol%. The Genbank accession numbers for the 16S rRNA, *atpD*, *recA*, *glnII*, *gyrB*, *nodA* and *nodC* gene sequences are OM392036, OM429377, OM429400, OM429436, OM429451, OM429502 and OM429498, respectively. The GenBank draft genome accession number of strain 1AS141^T is JBAMYD000000000.

Funding information

This study was funded by the PHC Maghreb project (34MAG21) and the Tunisian Ministry of Higher Education and Scientific Research. The contribution of J.T.T. was funded by Agriculture and Agri-Food Canada through projects J-002272 and J-002295.

Acknowledgements

All contributions made by members of the Centre of Biotechnology of Borj Cedria are sincerely appreciated. The authors would like to extend their gratitude to Mira Mnasri and Chahd Mnasri for their invaluable contributions to nodules and grain sampling. We are thankful to A. Oren for assisting with the naming of the new species.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

- Gritli T, Ellouze W, Chihaoui S-A, Barhoumi F, Mhamdi R, et al. Genotypic and symbiotic diversity of native rhizobia nodulating red pea (*Lathyrus cicera* L.) in Tunisia. *Syst Appl Microbiol* 2020;43:126049.
- Ilahi H, Hsouna J, Ellouze W, Gritli T, Chihaoui S-A, et al. Phylogenetic study of rhizobia nodulating pea (*Pisum sativum*) isolated from different geographic locations in Tunisia. *Syst Appl Microbiol* 2021;44:126221.
- Tian CF, Young JPW, Wang ET, Tamimi SM, Chen WX. Population mixing of *Rhizobium leguminosarum* bv. viciae nodulating *Vicia faba*: the role of recombination and lateral gene transfer. *FEMS Microbiol Ecol* 2010;73:563–576.
- Saïdi S, Ramírez-Bahena M-H, Santillana N, Zúñiga D, Álvarez-Martínez E, et al. *Rhizobium laguerreae* sp. nov. nodulates *Vicia faba* on several continents. *Int J Syst Evol Microbiol* 2014;64:242–247.
- Young JPW, Moeskjær S, Afonin A, Rahi P, Maluk M, et al. Defining the *Rhizobium leguminosarum* species complex. *Genes* 2021;12:111.
- Jorriin B, Palacios JM, Peix Á, Imperial J. *Rhizobium ruizarguesonis* sp. nov., isolated from nodules of *Pisum sativum* L. *Syst Appl Microbiol* 2020;43:126090.
- Rahi P, Giram P, Chaudhari D, diCenzo GC, Kiran S, et al. *Rhizobium indicum* sp. nov., isolated from root nodules of pea (*Pisum sativum*) cultivated in the Indian trans-Himalayas. *Syst Appl Microbiol* 2020;43:126127.
- Zhang J, Peng S, Andrews M, Liu C, Shang Y, et al. *Rhizobium changzhienense* sp. nov., isolated from effective nodules of *Vicia sativa* L. in North China. *Int J Syst Evol Microbiol* 2019;71.
- Young JPW, Jorriin B, Moeskjær S, James EK. *Rhizobium brockwellii* sp. nov., *Rhizobium johnstonii* sp. nov. and *Rhizobium beringeri* sp. nov., three genospecies within the *Rhizobium leguminosarum* species complex. *Int J Syst Evol Microbiol* 2023;73.
- Hsouna J, Ilahi H, Han J-C, Gritli T, Ellouze W, et al. *Rhizobium acaciae* sp. nov., a new nitrogen-fixing symbiotype isolated from root nodules of *Acacia saligna* in Tunisia. *Int J Syst Evol Microbiol* 2023;73.
- Hsouna J, Gritli T, Ilahi H, Ellouze W, Mansouri M, et al. Genotypic and symbiotic diversity studies of rhizobia nodulating *Acacia saligna* in Tunisia reveal two novel symbiotypes within the *Rhizobium leguminosarum* complex and *Bradyrhizobium*. *Syst Appl Microbiol* 2022;45:126343.
- Rogel MA, Ormeño-Orrillo E, Martínez Romero E. Symbiotype in rhizobia reflect bacterial adaptation to legumes. *Syst Appl Microbiol* 2011;34:96–104.
- Vincent J. *A Manual for the Practical Study of the Root-Nodule Bacteria*, 15th edn. Oxford and Edinburgh: Blackwell Scientific Publications; 1970.
- Terefework Z, Nick G, Suomalainen S, Paulin L, Lindstrom K. Phylogeny of *Rhizobium galegae* with respect to other rhizobia and agrobacteria. *Int J Syst Bacteriol* 1998;48:349–356.
- Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–1874.
- Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980;16:111–120.
- Andrews S. A quality control tool for high throughput sequence data; 2010. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 2017;13:e1005595.
- Olson RD, Assaf R, Brettin T, Conrad N, Cucinell C, et al. Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. *Nucleic Acids Res* 2023;51:D678–D689.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
- Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.
- 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.
- Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinform* 2013;14:60.
- 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.
- Kim J, Na SI, Kim D, Chun J. UBCG2: Up-to-date bacterial core genes and pipeline for phylogenomic analysis. *J Microbiol* 2021;59:609–615.
- Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 2019;35:4453–4455.
- Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 2022;38:5315–5316.
- Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P-A, et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 2022;50:D785–D794.

29. Shaw J, Yu YW. Fast and robust metagenomic sequence comparison through sparse chaining with skani. *Nat Methods* 2023;20:1661–1665.
30. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform* 2010;11:119.
31. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol* 2011;7:e1002195.
32. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* 2016;17:132.
33. Matsen FA, Kodner RB, Armbrust EV. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinform* 2010;11:538.
34. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;5:e9490.
35. González AJ, Cleenwerck I, De Vos P, Fernández-Sanz AM. *Pseudomonas asturiensis* sp. nov., isolated from soybean and weeds. *Syst Appl Microbiol* 2013;36:320–324.
36. Mulet M, Lalucat J, García-Valdés E. DNA sequence-based analysis of the *Pseudomonas* species. *Environ Microbiol* 2010;12:1513–1530.

The Microbiology Society is a membership charity and not-for-profit publisher.

Your submissions to our titles support the community – ensuring that we continue to provide events, grants and professional development for microbiologists at all career stages.

Find out more and submit your article at microbiologyresearch.org