

Scleromatobacter humisilvae gen. nov., sp. nov., a novel bacterium isolated from oak forest soil

Sophie Mieszkin, Blandine Trouche, Julien Ancousture, Youssef Raouf, Stéphane Uroz, Karine Alain

► To cite this version:

Sophie Mieszkin, Blandine Trouche, Julien Ancousture, Youssef Raouf, Stéphane Uroz, et al.. Scleromatobacter humisilvae gen. nov., sp. nov., a novel bacterium isolated from oak forest soil. International Journal of Systematic and Evolutionary Microbiology, 2023, 73 (3), pp.005793. 10.1099/ijsem.0.005793. hal-04236566

HAL Id: hal-04236566 https://hal.science/hal-04236566

Submitted on 11 Oct 2023 $\,$

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.

1 Scleromatobacter humisilvae gen. nov., sp. nov., a novel bacterium isolated 2 from an oak forest soil 3 Sophie Mieszkin^{1,2*,} Blandine Trouche¹, Julien Ancousture¹, Youssef Raouf¹, 4 Stéphane Uroz² and Karine Alain¹ 5 6 7 Author affiliations : 8 ¹Univ Brest, CNRS, Ifremer, UMR6197 Biologie et Ecologie des Ecosystèmes marins Profonds, F-29280 Plouzané, France 9 ²Centre INRAE-Grand Est-Nancy, Université de Lorraine, INRAE, UMR IAM, 54280 10 11 Champenoux, F-54000 Nancy, France 12 13 * Correspondence: Sophie.Mieszkin@univ-brest.fr 14 15 Keywords: Scleromatobacter humisilvae; Betaproteobacteria; Burkholderiales; 16 Comamonadaceae; oak forest soil 17 **Repositories**: 18 19 The 16S rRNA gene sequence and the assembled genome sequences of strain BS-T2-15^T have been deposited in GenBank under the accession numbers OM630150 20 21 and JAJLJH00000000, respectively. 22 23 Abbreviations: 24 AAI, Average Amino acids Identity; AF, Alignment Fraction; ANI, Average Nucleotide Identity; CAPSO, N-cyclohexyl-3-aminopropanesulfonic acid; COGs, cluster of 25

26 orthologous groups; dDDH, Digital DNA-DNA hybridization; DMSO, dimethyl diphosphatidylglycerol; DSMZ, Deutsche Sammlung von 27 sulfoxide; DPG, 28 Mikroorganismen und Zellkulturen; HEPES, 4-(2-hydroxyethyl)-1piperazineethanesulfonic Homopiperazine-1,4-bis(2-29 acid: HOMOPIPES, 30 ethanesulfonic acid); HPIC, Hight Pressure Ion Chromatography; ICSP, International Committee on Systematics of Prokaryotes; KEGG, Kyoto Encyclopedia of Genes and 31 Genomes; MES, 2-(N-Morpholino) EthaneSulfonic acid; MMN, Mannitol-Mobility-32 33 Nitrate medium; OGRI, overall genome related indices; PE, agar 34 Phosphatidylethanolamine; PG, Phosphatidylglycerol; PHA, Polyhydroxyalkanoates; PIPES, Piperazine-N,N'-bis(2-ethanesulfonic acid); POCP, Percentage Of Conserved 35 Proteins; TEM, Transmission Electron Microscopy; TSA, Tryptic Soy Agar; TSB, 36 37 Tryptic Soy Broth; UBOCC, UBO Culture Collection.

38 Abstract

A novel bacterial strain, designated BS-T2-15^T, isolated from forest soil in close 39 proximity to decaying oak wood, was characterized using a polyphasic taxonomic 40 approach. Phylogenetic analyses based on 16S rRNA gene sequences as well as 41 42 phylogenomic analyses based on coding sequences of 340 concatenated core proteins indicated that strain BS-T2-15^T forms a distinct and robust lineage in the 43 44 Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella branch of the 45 order Burkholderiales. The amino acid identity (AAI) and the percentage of conserved proteins (POCP) between genome of strain BS-T2-15^T and genomes of 46 closely related type strains ranged from 54.27 to 66.57% and from 43.17 to 49.27%, 47 respectively, providing genomic evidence that strain BS-T2-15^T represents a new 48 genus. Its cells stain Gram-negative, are aerobic, motile by a polar flagellum, rod-49 50 shaped and form incrusted white to ivory colonies. Optimal growth is observed at 20-22°C, pH 6 and 0% NaCl. The predominant fatty acids of strain BS-T2-15^T are $C_{16:1}$ 51 52 ω 7c, C_{16:0} and C_{14:0} 2-OH. Its polar lipid profile consists of a mixture of 53 phosphatidylethanolamine, diphosphatidylglycerol and phosphatidylglycerol and its 54 main respiratory quinone is ubiquinone 8. The estimated size of its genome is 6.28 Mb with a DNA G+C content of 69.56 mol%. Therefore, on the basis of phenotypic 55 and genotypic properties, the new strain BS-T2-15^T represents a new genus and a 56 novel species for which the name Scleromatobacter humisilvae gen. nov., sp. nov., is 57 proposed. The type strain is BS-T2-15^T (DSM 113115^T = UBOCC-M-3373^T). 58

59 In forest ecosystems, wood decay is an important process that participates to the soil 60 fertility through the action of a complex guild of decomposers [1, 2]. Among them, fungi are considered as the main actors, while the role of bacteria is much less 61 documented [3]. To fill this gap, the taxonomic diversity and the metabolic and 62 functional potential of bacterial strains isolated from a soil-decaying-wood continuum 63 64 have been investigated [4]. Overall, this study demonstrated a community dominated Paraburkholderia 65 by representatives of the genera (Betaproteobacteria), Streptomyces, Kitasatospora, Arthrobacter and Streptacidiphilus (Actinobacteria), 66 Dyella (Gammaproteobacteria) and by uncharacterized bacterial strains [4]. At the 67 68 functional level, this study also revealed that soil bacterial communities have the potential to decompose organic matter and a stronger effectiveness than those 69 isolated from decaying oak wood. The taxonomic assignation and functional 70 71 screening allowed the identification of various strains of interest, among which strain BS-T2-15^T, one of the uncharacterized strains, was selected to be further 72 73 characterized in the present study to clarify its taxonomic position within the 74 Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella branch of the 75 order Burkholderiales, in the class Betaproteobacteria. To date, there is a taxonomic uncertainty regarding the classification of these genera at the family rank level. 76 77 Recently, Liu et al. [5] proposed to refer to this branch as the closest-to-Comamonadaceae (CTC) group, as it forms a phylogenetically coherent group close 78 to genera belonging to the Comamonadaceae family. They proposed to classify the 79 CTC group as a new family, 'Sphaerotilaceae' fam. nov., based on phylogenomic 80 81 analyses, genome relatedness indices and phenotypic data. However, this new 82 delineation is not yet recognized by the International Committee on Systematics of 83 Prokaryotes (ICSP). In addition, they also proposed several taxonomic revisions at

the genera and species levels. Among remarkable physiological features of the *Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella* branch, some species are known to be involved in the biodegradation of petroleum compounds or in the reduction of chlorate to chloride under anaerobic conditions [6-8].

In the present study, we performed a polyphasic taxonomic characterization of strain BS-T2-15^T and provided phenotypic, phylogenomic, and genomic evidences that it meets the criteria for delineating a new species of a new genus within the *Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella* branch. We propose to name this new species *Scleromatobacter humisilvae* gen. nov., sp. nov.

93

94 ISOLATION AND ECOLOGY

Strain BS-T2-15^T is from a collection of 308 bacterial strains established as part of a 95 study on the functional abilities of bacterial isolates from decaying wood and their 96 97 comparison with the properties of bacteria from underlying soil [4]. The experimental 98 set up consisted in oak discs that have been placed during 9 months on the soil 99 surface of the forest experimental site of Champenoux (north-eastern France; Lat: 48.718420° N; Long: 6.346580° E; Alt: 248 m; 0.5 ha of surface). Strain BS-T2-15^T 100 101 was isolated from the bulk soil in direct contact with decaying oak disc by serial 102 dilutions on 1/10 diluted Tryptic Soy Agar (TSA) medium (Difco's Tryptic Soy Broth (TSB) 3 $g.L^{-1}$ and agar 15 $g.L^{-1}$) containing cycloheximide (100 $g.L^{-1}$, final 103 104 concentration), with a pH adjusted to 5. The strain was then purified by 3 successive 105 spreads on 1/10 diluted TSA plates at pH 5 to obtain a pure culture. The strain BS-T2-15^T was then grown routinely on 1/10 TSA or 1/10 TSB media adjusted to pH 5 106 during 3 days at 20°C, under agitation (250 rpm). Its purity was routinely confirmed 107

by microscopic observations, and by sequencing of its 16S rRNA gene and genome. Stock cultures were stored at -80° C in 1/10 TSB medium supplemented with 5% (v/v) dimethylsulfoxide (DMSO). Main chemical characteristics of the bulk soil samples were pH 4.65 ± 0.10, 48.38 ± 7.93 g.kg⁻¹ total carbon, 3.26 ± 0.49 g.kg⁻¹ total nitrogen, 0.17 ± 0.01 g.kg⁻¹ phosphorus extracted according to the Duchaufour method and 83.80 ± 13.92 g.kg⁻¹ organic matter [4].

Strain BS-T2-15^T (DSM 113115^T = UBOCC-M-3373^T) is available in the public culture
collections Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ;
https://www.dsmz.de/collection) and UBO Culture Collection (UBOCC;
https://www.univ-brest.fr/ubocc).

118

119 MORPHOLOGY AND PHYSIOLOGY

Colony morphology of the isolate BS-T2-15^T was observed on 1/10 TSA adjusted to 120 pH 6. Cell morphology and motility were determined by light microscopy (Olympus 121 122 BX60 and CX40) and transmission electron microscopy (TEM; JEOL JEM 1400). 123 Motility was also observed by using the Mannitol-Motility-Nitrate agar medium (MMN, 124 composed (per liter) of 10 g tryptic hydrolysate of casein, 1 g potassium nitrate, 7.5 g mannitol, 40 mg phenol red and 3.5 g agar), which was also used to evidence 125 126 mannitol fermentation and nitrate reductase activity. Gram-staining was determined using standard procedures and confirmed with a KOH (3%) test. Catalase and 127 cytochrome oxidase activities were respectively evaluated using H₂O₂ and strips of 128 N, N, N', N'-tetramethyl-*p*-phenylenediamine dihydrochloride (Bio-Rad). The oxidase 129 test was repeated several times using independent cultures to confirm the result. 130 Colonies of BS-T2-15^T are circular and rough with a color between white and ivory. A 131 major phenotypic feature is that its colonies are incrusted in the agar medium. To our 132

133 knowledge, this was not yet reported in the literature for genera belonging to the Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella branch. Cells 134 135 are Gram-negative coccobacilli that divide by binary fission and occur mainly singly, but can also form aggregates (Fig. 1A, B). Cell size range from 0.59 to 0.93 µm wide 136 137 (mean 0.78 µm; *n*=45) and from 1.71 to 2.73 µm long (mean 2.15 µm; *n*=45). Strain 138 BS-T2-15^T is motile, as confirmed by growth observations on MMN agar medium and 139 the observation of a single polar flagellum per cell by TEM (Fig. 1C, D). Refringent 140 intracellular granules, which could be polyhydroxyalkanoates (PHA) storage 141 granules, were also observed (Fig. 1A). The strain is catalase positive and oxidase negative. The latter phenotypic feature clearly distinguishes strain BS-T2-15^T from 142 143 other genera of the Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-144 Ideonella branch which are all oxidase positive. However, it is important to note that, 145 cytochrome C oxidase complex was identified in its genome. This point is discussed 146 latter in the study.

147

Physiological characterization of the novel strain BS-T2-15^T was carried out 148 149 aerobically, in triplicate, on 1/10 TSA or TSB adjusted to pH 6 at 20°C and under 150 agitation (250 rpm). Determination of the temperature range for growth and salt 151 tolerance were respectively tested over the range 0-45°C, at 5°C intervals and 0-10% NaCl (w/v), at 0.5% intervals, both for 10 days on 1/10 TSA at pH 6. Growth was 152 153 observed from 4 to 30°C with optimal growth between 20-22°C. Concerning salt tolerance, the strain grew only at 0.5% NaCl and exhibited an optimal growth without 154 155 sodium chloride. The pH range for growth was tested from pH 2.0 to pH 12 (at 20°C), 156 with increments of 1 unit on 1/10 TSB medium for 8 days. Cells were routinely enumerated by direct cell counting using a modified Thoma chamber (Preciss, 157

France; surface: 0.0025 mm², depth: 10 µm). The following buffers (each at 20 mM, 158 Sigma-Aldrich, Darmstadt, Germany) were used to adjust the required pH: pH 4.0 159 and 5.0 with HOMOPIPES buffer, pH 6.0 with MES buffer, pH 7.0 with PIPES buffer, 160 pH 8.0 with HEPES buffer, and pH 9.0 and 10.0 with CAPSO buffer. For pH 2.0, 3.0, 161 11.0 and 12.0, no buffer was used. Growth of strain BS-T2-15^T was observed from 162 pH 2.0 to 12 with optimal growth at pH 6. The growth kinetics of BS-T2-15^T under 163 164 optimal conditions was then studied in triplicates, on 1/10 TSB adjusted to pH 6 at 165 22°C with no NaCl salt, under agitation (250 rpm) for 4 days. Cell growth was 166 monitored by direct cell counting, every 3 h using a modified Thoma chamber, in order to determine the growth rate and doubling time of the strain under optimal 167 culture conditions. The growth rate and the generation time of strain BS-T2-15^T are 168 respectively, 0.17 h⁻¹ and 3 h 58 min. 169

170

171 Utilization of carbon sources was investigated using the mineral basis of the 1/10 TSB medium containing (L⁻¹): 0.25 g KH₂PO₄ and adjusted at pH 6. Each substrate 172 173 (lactose, D-mannose, D-ribose, D-maltose, D-glucose, N-acetylglucosamine, butyric 174 acid, gluconate, citrate, benzoate, acetate, pyruvate and urea) was supplied at a final 175 concentration of 20 mM. Utilization of nitrogen sources was investigated using 1/10 176 TSB medium adjusted at pH 6 with nitrate and ammonium supplied at a final concentration of 20 mM, while nitrite was supplied at a final concentration of 5 mM. 177 178 Before inoculation, cells at the end of their exponential growth phase were harvested and washed three times with distilled water and then the bacterial suspension was 179 adjusted to obtain a Mac Farland index value of 1. A medium composed of mineral 180 181 basis of the 1/10 TSB without carbon source was used as negative control for each carbon utilization bioassay and a positive control was performed using 1/10 TSB 182

183 medium adjusted at pH 6. The hydrolysis of cellulose was determined using Carboxymethyl Cellulose (CM Cellulose) as described by Mieszkin et al. [4]. These 184 185 tests were completed with liquid or solid cultures allowing to determine: (i) mannitol 186 fermentation and presence of nitrate reductase (MMN agar medium), (ii) fermentative 187 pathways (mixed acids or butane-2,3-diol pathways; Clark and Lubs liquid medium), 188 (iii) glucose and lactose fermentation, gas and H₂S production (Kligler-Hajna agar 189 medium), and, (iv) citrate utilization (Simmons citrate agar medium). For these 190 assays, culture media, cell cultures and washes and preparation of bacterial 191 suspensions were performed as described above and by Mieszkin et al. [9]. Strain BS-T2-15^T is chemoorganoheterotrophic and grows by aerobic respiration. It 192 193 catabolizes *D*-glucose, lactose and pyruvate while a moderate growth was obtained 194 with *D*-ribose, *N*-acetylglucosamine, urea and gluconate. A weak growth was also 195 observed with all the nitrogen substrates tested (nitrate, nitrite and ammonium). In addition, BS-T2-15^T is capable of using nitrate or nitrite as terminal electron 196 197 acceptors showing that the nitrate reductase is functional (Table 1).

198

Volatile fatty acid production of strain BS-T2-15^T after growth in duplicate in 1/10 TSB was assessed by high-pressure ion chromatography (Dionex ICS-6000 HPIC) (Growth conditions: pH 6, 20°C, under agitation (250 rpm), for 3 days). HPIC analyses showed that the strain consumed all the pyruvate present in the medium during its growth (approximately 930 μ M), but did not produce acetate, lactate, pyruvate or oxalate under these culture conditions. This is consistent with previous results showing that it is capable to grow with pyruvate as sole carbon source.

Strain BS-T2-15^T was not capable to use chlorate as sole terminal electron acceptor (at 20 mM), when grown anaerobically on 5 g.L⁻¹ NaCl, with 1 g.L⁻¹ tryptone as carbon and energy source (with respect to negative and positive controls).

Sensitivity to antibiotics of strain BS-T2-15^T was tested by the disc diffusion method 209 210 after spreading cell suspensions (1 McFarland) on 1/10 TSA adjusted to pH 6 at 20°C for 7 days. Discs were impregnated with antibiotic solutions to obtain a final 211 212 quantity of antibiotic per disc of: kanamycin (40µg), rifampicin (30 µg), novobiocin (30 213 μ g), gentamicin (15 μ g), ampicillin (10 μ g), penicillin (10 U), oxacillin (5 μ g), 214 oxytetracycline (30 µg), chloramphenicol (30 µg), and streptomycin (10 µg). The 215 discs were then placed on the agar surface. The inhibition zones were read after 3 days of incubation at 20°C. Among the antibiotics tested, strain BS-T2-15^T was 216 217 sensitive to kanamycin, novobiocin, gentamicin and streptomycin.

218

219 CHEMOTAXONOMY

In order to analyze respiratory quinones, polar lipids, and fatty acids, cells of strain BS-T2-15^T were cultured in 1/10 TSB pH 6 at 20°C under agitation (250 rpm) and then harvested by centrifugation (800 *g*; 10 min) at the end of their exponential phase of growth. These analyses were carried out by the Identification Service of the DSMZ (Braunschweig, Germany) as described by Tindall [10, 11] and Kuykendall *et al.* [12]. The major respiratory quinone identified for strain BS-T2-15^T is ubiquinone 8 (Q-8).

The polar lipid profile of the novel isolate consisted in phosphatidylethanolamine (PE), diphosphatidylglycerol (DPG), and phosphatidylglycerol (PG). A common feature of the new strain and species belonging to the related genera *Leptothrix*, *Rubrivivax*, *Ideonella* and *Aquincola* is the high proportion of PE and PG, which is consistent with the *Comamonadaceae* family [13, 14]. However, a major

characteristic of strain $BS-T2-15^{T}$ is its high content of DPG, a polar lipid that is absent or only present in low proportions in the type strains of the closest genera (Table S1 et Fig. S1).

The predominant cellular fatty acids (>10% of the total fatty acids) of strain BS-T2-234 15^{T} are C_{16:1} ω 7c (40.45%), C_{16:0} (25.18%) and C_{14:0} 2-OH (10.49%) (Table 2). 235 236 Noticeably, the main saturated fatty acid detected (*i.e.*, C_{16:0}) appeared dominant and common among BS-T2-15^T and the four closely related genera (*Leptothrix*, 237 Rubrivivax, Ideonella and Aquincola), while the hydroxyl fatty acid (*i.e.* C_{14:0} 2-OH) 238 was only detected in strain BS-T2-15^T (Table 1). In addition, the presence of the 239 unsaturated fatty acid $C_{16:1} \omega 7c$ is another major feature of the new strain but could 240 also be represented by summed feature 3 (C16: 1 ω 7c o and/or C16: 1 ω 6c) in 241 Rubrivivax gelatinosus DSM 1709^{T} and Ideonella dechloratans CGUG 30977^{T} . 242

243

244 16S rRNA GENE PHYLOGENY

A full 16S rRNA gene sequence (1512 nt) of strain BS-T2-15^T was extracted from the 245 246 genome using the bacterial ribosomal predictor tseemann/barrnap on Galaxy v1.2.1 247 [19]. Pairwise 16S rRNA gene sequence similarity was calculated using the global 248 alignment algorithm implemented at the EzTaxon-e server (http://eztaxon-249 e.ezbiocloud.net/; [20]). Phylogenetic analyses were performed using the software Seaview version 4.7 for trees reconstruction by the neighbor-joining (NJ) and 250 251 maximum parsimony (MP) methods [21-23]. The evolutionary distances were respectively calculated using the Kimura two-parameters and the Dnapars modes. 252 253 The robustness of the inferred topology was assessed by bootstrap analyses based 254 on 1,000 replications. A third tree was build using the maximum likelihood (PhyML) 255 method subtree-pruning-regrafting algorithm with the

(http://phylogeny.lirmm.fr/phylo_cgi/simple_phylogeny.cgi) [24, 25]. The 16S rRNA 256 gene sequence analysis of strain BS-T2-15^T showed that it belongs to the class 257 Betaproteobacteria, the order Burkholderiales and the family Comamonadaceae. The 258 NJ and MP phylogenetic trees based on 16S rRNA gene sequences revealed that 259 strain BS-T2-15^T does not form a monophyletic clade with any representatives of 260 261 related taxa (Fig. S2 and S3). However, most of the bootstrap percentages of the trees had a very low value (<70%). Concerning the PhyML tree, a clade was 262 obtained with strain BS-T2-15^T and *Leptothrix mobilis* (Feox-1: X97071) but it was 263 distantly related to most of the genera of the Rubrivivax-Roseateles-Leptothrix-264 Azohydromonas-Aquincola-Ideonella branch (Fig. S4). 265

Comparative analysis of the 16S rRNA gene sequence of the novel strain with 266 relatives having validly published names showed that BS-T2-15^T is equidistant to 267 species belonging to the genera Ideonella (96.55-95.78%), Leptothrix (96.47-268 95.50%), Aquincola (96.47–96.13%), Rubrivivax (96.40–95.85%) and Aquabacterium 269 (96.13-95.53%). Its closest relative is *Ideonella azotifigens* $1a22^{T}$ (96.55%) followed 270 by Leptothrix mobilis $FeOx-1^{T}$ (96.47%), Aquincola tertiaricarbonis $L10^{T}$ (96.47%), 271 and *I. dechloratans* CCUG 30898^T (96.47%), *Rubrivivax gelatinosus* ATCC 17011^T 272 (96.40%) and *R. benzoatilyticus* $JA2^{T}$ (96.40%). 273

Overall, these results demonstrate that the 16S rRNA gene-based phylogeny lacks of resolution to evaluate the taxonomic position of strain BS-T2-15^T among the *Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella* branch, as previously underlined by Liu *et al.* [25]. Genome based phylogeny is therefore mandatory to gain a better resolution to assign with certainty the taxonomic position of the new strain BS-T2-15^T.

280

GENOME FEATURES

Genomic DNA of strain BS-T2-15^T was extracted with a standard PCI (Phenol: 282 Chloroform: Isoamyl Alcohol (25:24:1)) protocol [26]. Whole genome sequencing was 283 performed by the Novogene company (Cambridge, United Kingdom), using the 284 285 Illumina NovaSeq 6000 PE150 technology (2×150 bp paired-end reads). Reads 286 quality was then evaluated with Fast QC (v0.11.9; 287 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [27]. The whole genome 288 was assembled by using the Shovill pipeline with SPAdes and default parameters 289 [28]. Genome completeness and potential contamination were estimated with CheckM on the MicroScope Microbial Genome Annotation and Analysis Platform 290 291 (MaGe; https://mage.genoscope.cns.fr) [29]. The overall genome size was 6,276,266 292 bp (for 100% completion and 0% contamination) for 28 contigs and the GC content 293 was 69.56%. The N50 and L50 values were respectively 579,168 bp and 3 contigs. A 294 total of 5,712 coding DNA sequences (CDSs), 3 rRNA and 55 tRNA genes 295 (corresponding to the 20 essential amino acids) were detected using the MaGe 296 pipeline (Table 3 and Fig S5). Most of the CDSs (78.29%) could be assigned to at 297 least one cluster of orthologous groups (COGs). COGs categories related to 298 metabolic processes were dominant (38.30% of the CDSs) and the main processes 299 involved (>5% of the CDSs) were: (i) amino acid transport and metabolism (E; 9.04%); (ii) carbohydrate transport and metabolism (G; 6.45%); inorganic ion 300 301 transport and metabolism (P; 5.45%) and (iv) energy production and conversion (C; 302 5.38%). Then, 24.56% of CDSs were allocated to the COGs category related to 303 cellular process and signaling and the following processes were dominant: signal 304 transduction mechanisms (T; 7.64%) and cell wall, membrane and envelope

biogenesis (M; 5.26%). Finally, 14.91% of the CDSs were predicted to be involved in
the processes dedicated to the information storage and processing (Table S2).

307

In addition to phylogenetic trees based on the 16S rRNA sequences that do not provide robust phylogenetic reconstructions with this dataset, a phylogenomic tree for was constructed based on 340 core gene clusters (gene clusters appearing once in each genome).

312 This tree encompassed 21 reference genomes representative of the Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella 313 branch and phylogenetically related to strain BS-T2-15^T (Aguabacterium commune DSM 11901^T, 314 Aquincola rivuli KYPY4^T, Aquincola tertiaricarbonis DSM 18512^T, Azohydromonas 315 lata NBRC 102462^T, Ideonella azotifigens DSM 21438^T, Ideonella benzenivorans B7, 316 Ideonella dechloratans CGUG 30977^T, Ideonella livida TBM-1^T, Ideonella paludis 317 KCTC 32238^T, Ideonella sakaiensis 201-F6^T, Leptothrix cholodnii SP-6^T, Leptothrix 318 *mobilis* DSM 10617^T, *Roseateles aquatilis* CCUG 48205^T, *Roseateles depolymerans* 319 KCTC 42856^T, Rubrivivax albus ICH-3^T, Rubrivivax benzoatilyticus JA2^T, Rubrivivax 320 gelatinosus DSM 1709^T, Schlegelella brevitalea DSM 7029^T, Sphaerotilus hippei 321 DSM 566^T, Sphaerotilus natans ATCC 13338^T and Sphaerotilus natans subsp. 322 sulfidivorans D-507^T), the target genome BS-T2-15^T and two closely related 323 Metagenome Assembled Genomes (MAGs; from the Genome Taxonomy Database 324 325 (GTDB, https://gtdb.ecogenomic.org/); accession numbers GB_GCA_013289985.1 and GB_GCA_903845345.1) artificially classified as genus CAIMXF01. These 326 closely related MAGs were respectively recovered from a forest soil [30] and from 327 328 stratified freshwater lakes and ponds [31-33]. The phylogenomic tree was built as follows: (i) amino acid sequences were extracted, aligned and concatenated using 329

330 command line anvi-get-sequences-for-gene-clusters on Anvi'o v. 7.1 [34], then (ii) positions with over 0.5 gap frequency were masked, and finally (iii) a maximum 331 332 likelihood tree was computed using IQTREE (v2.0.3, [35, 36]) under the model WAG with 1,000 ultrafast bootstrap replicates (See SI for more details). The tree was 333 334 rooted using the genome of Alcaligenes faecalis ZD02 (GCA_000967305.2) as an outgroup. A second phylogenomic tree was built by placing BS-T2-15^T in the GTDB 335 336 tree using GTDB-tk (v2.1, [37]), then subsetting the tree to family Burkholderiaceae 337 and genus Escherichia, extracting the amino acid alignment, masking positions with 338 over 0.5 gap and computing a maximum likelihood (ML) tree with IQTREE under the 339 model WAG as described above. This second tree was built with 229 MAGs and 60 340 cultivated bacterial strains from the Rubrivivax-Roseateles-Leptothrix-341 Azohydromonas-Aquincola-Ideonella branch.

342 The phylogenomic tree built with the 21 reference genomes and the two MAGs closely related to strain BS-T2-15^T, was very robust with high ultrafast bootstrap 343 values, and showed a distinctly deeper branching of strain BS-T2-15^T with its two 344 345 closely related MAGs and a close proximity to the Ideonella clade (Fig. 2). A similar 346 result was also obtained with the phylogenomic tree built with the 229 MAGs (Fig. S6). Therefore, these two phylogenomic trees indicate that strain BS-T2-15^T indeed 347 348 represents a new genus of the Comamonadaceae family within the order 349 Rubrivivax-Roseateles-Leptothrix-Burkholderiales, and belongs the to 350 Azohydromonas-Aquincola-Ideonella branch.

351

In addition to the phylogenomic trees, Overall Genome Related Indices (OGRI) were obtained by pairwise calculations between genome of strain BS-T2-15^T, the two closely related MAGs of strain BS-T2-15^T and the 21 reference genomes. The

355 average nucleotide identity scores (ANI) were obtained using the FastANI calculator tool from the GTDB web server (https://gtdb.ecogenomic.org/tools/fastani) [39]. In 356 357 addition, FastANI calculator provided also the alignment fraction (AF) values between genomes. Digital DNA-DNA hybridization (dDDH) scores were determined by the 358 359 genome-to-genome distance calculator (GGDC 2.1), using formula 2 [40]. The 360 average amino acid identity (AAI) calculations between genomes were determined as 361 described by Kim et al. [41] using the EzAAI pipeline (http://leb.snu.ac.kr/ezaai). The 362 percentage of conserved proteins (POCP) was estimated as described by Qin et al. 363 [42].

Digital DNA-DNA hybridization and ANI values between the strain $BS-T2-15^{T}$ 364 365 genome and the genomes of type strains of the closest genera ranged from 20.00 to 366 20.50% and from 78.50 to 79.70%, respectively (Table 3). These values are both far 367 below the dDDH threshold level of 70% and the ANI value of 95-96% that are generally accepted for species delineation [43, 44]. Concerning the AAI and POCP 368 values between genome of strain BS-T2-15^T and genomes of the type strains of the 369 370 closest genera, they ranged, respectively, from 64.27 to 66.57% and from 43.17 to 371 49.27% (Table 3). AAI values are close to the 65% cutoff value generally accepted 372 for delineating a new genus, while POCP values are below the 50% threshold 373 recognized for demarcating a new genus [42, 45]. Finally, the AF values ranged between 0.40 and 0.48 between genomes of strain BS-T2-15^T and type strains of the 374 closest genera (Table 3), and were observed in conjunction with low ANI values [46], 375 which is not incompatible with the description of a new genus. In this case, POCP 376 377 and, to a lesser extent, ANI, are the most relevant indices to demonstrate that strain BS-T2-15^T represents a new genus. Similar results were obtained when considering 378 379 the 21 reference genomes (Fig. S7 and Tab. S3). In addition, all OGRI values

obtained between the BS-T2-15^T strain genome and the genomes of its two closely 380 related MAGs were higher compared to the values obtained against genomes of 381 382 cultured type strains. These OGRI values relative to the closest MAGs were respectively in the range of 30.10-40.40% and 87.40-88.90% for dDDH and ANI, and 383 384 in the range of 84.14-86.43%, 70.43-74.76% and 0.70-0.80 for AAI, POCP and AF indices (Tab S4-S8). These dDDH and ANI values are below the threshold accepted 385 386 for species delineation while the AAI, POCP and AF values are above the threshold 387 for a new genus delineation. Overall, these OGRI analyses clearly show that strain BS-T2-15^T and its two closely related MAGs (i) belong to the same genus, (ii) 388 389 represent a new genus of the Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella branch and, (iii) belong to three distinct genomic species. 390 391 Furthermore, values of the different OGRI indices calculated between the species of the genus *Ideonella* and strain BS-T2-15^T and its two closely related MAGs are much 392 lower than those obtained between strain BS-T2-15^T and its two closely related 393 394 MAGs (Tab S4-S8). Thus, OGRI indices, as well as phylogenomic analyses, lead all to the conclusion that strain BS-T2-15^T represents a new species of a new genus of 395 396 the Rubrivivax-Roseateles-Leptothrix-Azohydromonas-Aquincola-Ideonella branch.

397

For gene function prediction, genome annotations were performed with the MaGe platform using KEGG and BioCyc databases. The genome of the new strain BS-T2-15^T encodes metabolic pathways for organoheterotrophic growth. Notably, the genome encodes a complete pentose phosphate pathway, and a complete Entner-Doudoroff pathway. The tricarboxylic acid cycle (TCA) pathway for aerobic respiration is almost complete (one enzyme missing). The genome of the novel isolate encodes also the nitrate VIII reduction pathway (dissimilatory) for anaerobic

nitrate respiration, which has been shown to be functional experimentally. 405 Degradation pathways for organic compounds such as amino acids (alanine, 406 407 arginine, asparagine, serine, glutamine, glycine, histidine, cysteine, lysine, ornithine, taurine, threonine and tryptophan), carbohydrates (acetoin, chitin, glucose, glucose-408 409 1-phosphate, lactose, ribose and xylose), aromatic compounds (anthranilate, 410 gentisate, phenylacetate and protocatechuate), alcohols (ethanol and glycerol), 411 aldehyde (L-lactaldehyde), amines and polyamines (choline, 4-aminobutyrate, 412 ethanolamine and urea), and carboxylates (D-gluconate, glutaryl CoA, glycolate and 413 glyoxylate) are evidenced. The cytochrome C oxidase complex (EC number: 7.1.1.9; 414 locus tags: COMBST215_v1_20371; 20374; 20375; 20376; 80152; 80156; 80157) 415 was also identified despite that the oxidase activity not being demonstrated 416 experimentally. The cyclopropane fatty acid (CFA) biosynthesis via the cyclopropane-417 fatty-acyl-phospholipid synthase (EC number: 2.1.1.79; locus tag: COMBST215 v1 10589; 50189) and the arginine dependent acid resistance 418 419 (arginine decarboxylase; EC number: 4.1.1.19; locus tag: COMBST215_v1_10270) 420 pathways are present and may give the strain a competitive advantage in acidic environments such as the forest soil ecosystem. The new strain BS-T2-15^T has also 421 422 the genetic potential to detoxify arsenate using glutaredoxin (gene: arsC; EC number: 423 1.20.4.1; locus tag: COMBST215 v1 40055) and to degrade superoxide radicals (identification of superoxide dismutase (EC number: 1.15.1.1; locus tags: 424 425 COMBST215_v1_20566; 40196; 50105) and catalase enzymes (EC number: 1.11.1.21; locus tag: COMBST215_v1_11754 and EC: 1.11.1.6; locus tags: 426 COMBST215 v1 60157; 140122)). In comparison with other closely related strains, 427 the new strain BS-T2-15^T is the only one with the genetic potential to achieve the 428

429 biosynthesis of betaxanthin (a secondary metabolite) and to degrade alanine and430 anthranilate (Table S9).

In conclusion, from the clear genotypic distance with the closest genera, the physiological similarities and some phenotypical and chemotaxonomic differences, we comply with the phylo-phenetic concept that currently prevails for the description of a new genus. Thus, we assigned strain $BS-T2-15^{T}$ to a novel species, of a novel genus, for which the name *Scleromatobacter humisilvae* gen. nov., sp. nov. is proposed.

437

438 **DESCRIPTION OF SCLEROMATOBACTER GEN. NOV.**

439 Scleromatobacter sp. (scle.ro'mae. N.L. gen. n. scleromae, of scleroma, a hardened
440 part; derived from the colonies that are incrusted in solid medium when the strain
441 grows in Petri dishes).

442 Cells are Gram-negative, mesophilic, aerobic, chemoorganotrophic, straight rod-443 shaped (0.6 to 0.9 µm by 1.7 to 2.7 µm in size) and motile with presence of storage 444 granules. Oxidase-negative and catalase-positive. The predominant fatty acids are 445 $C_{16:1} \omega 7c$, $C_{16:0}$ and $C_{14:0}$ 2-OH. The polar lipid profile consists of a mixture of phosphatidylethanolamine, diphosphatidylglycerol and phosphatidylglycerol and the 446 main ubiquinone is Q-8. The estimated size of the genome is 6.28 Mb with a DNA 447 448 G+C content of 69.56 mol%. Phylogenetically, on the base of whole genome 449 comparisons, the genus belongs to the family Comamonadaceae, order 450 Burkholderiales. Overall, the genus belongs to the Rubrivivax-Roseateles-Leptothrix-451 Azohydromonas-Aquincola-Ideonella branch. The type species is Scleromatobacter humisilvae (BS-T2-15^T). 452

453

454 **DESCRIPTION OF SCLEROMATOBACTER HUMISILVAE GEN. NOV. SP. NOV.**

455 *Scleromatobacter humisilvae* (hu.mi.sil'vae. L. fem. n. *humus*, soil; L. fem. n. *silva*, 456 forest; N.L. gen. fem. n. *humisilvae*; from forest soil; from the forest soil where the 457 strain has been isolated).

458 Displays the following properties in addition to those given in the genus description: 459 Colonies are circular, rough and incrusted in the agar medium with a color between white and ivory. Optimal growth occurs at 20-22°C, pH 6 and 0% NaCl. Positive for 460 461 nitrate reduction. Does not respire chlorate. The following carbon sources are used: 462 L-arabinose, D-cellobiose, D-glucose, D-ribose, D-arabitol, adonitol, butyrate, 463 gluconate. pyruvate. L-leucine, L-proline, N-acetylglucosamine, p-hydroxy 464 phenylacetic acid, bromo succinic acid, L-pyroglutamique acid, tween 40 and 80, 465 uridine, putrescine and dextrine. On the contrary, unable to use: citrate, acetate, 466 aspartate, benzoate, formate, methyl-pyruvate, mono-methyl-succinate, cis-aconitic 467 acid, D-galactonic acid lactone, D-galacturonic acid, D-gluconic acid, D-glucosaminic 468 acid, D-glucuronic acid, α -, β -, γ -hydroxy butyric acid, p-hydroxyphenylacetic acid, 469 itaconic acid, α -keto butyric acid, α -keto glutaric acid, α -keto valeric acid, D,L-lactic 470 acid, malonic acid, propionic acid, quinic acid, D-saccharic acid, sebacic acid, 471 succinic acid, succinamic acid, L-aspartic acid, L-glutamic acid, glycyl-L-Aspartic 472 acid, glycyl-L-glutamic acid, L-pyroglutamic acid, y-amino butyric acid, urocanic acid, 473 phosphate, glucose-1-phosphate, glucose-2-phosphate, D-fructose, L-fucose, D-474 galactose, D-galactose, gentiobiose, CM-cellulose, D-psicose, D-raffinose, L-475 rhamnose, sucrose, D-maltose, D-trehalose, turanose, D-mannose, D-melibiose, α-476 D-lactose, lactulose, maltose, ß-methyl-D-glucoside, m-inositol, xylitol, D-mannitol, i-477 erythritol, D-sorbitol, glycerol, D,L-α-glycerol, 2,3-butanediol, N-acetyl-D-478 glucosamine. *N*-acetyl-D-galactosamine, 2-aminoethanol, phenylethylamine,

Glucuronamide, L-alaninamide, D-, L-alanine, L-phenylalanine, L-alanyl-glycine, Lasparagine, L-histidine, L-ornithine, hydroxy-L-proline, D-, L-serine, L-threonine, D,Lcarnitine, thymidine, α -cyclodextrin, inosine, glycogen, xylan, lignin and chitin. Produces the α -glucosidase but is negative for urease.

483

The type strain BS-T2-15^T (DSM 113115^T =UBOCC-M-3373^T) was isolated from bulk soil in direct contact with decaying oak placed for 9 months on the floor of the Champenoux forest experimental site (France). The 16S rRNA gene sequence and the assembled genome sequences of strain BS-T2-15^T have been deposited in GenBank under the accession numbers OM630150 and JAJLJH000000000, respectively.

490

491 Acknowledgements

We are grateful to Claire Geslin (UBO, LM2E) and Philippe Eliès (UBO, Plateforme 492 493 d'imagerie et de mesures en microscopie (PIMM)) for their technical assistance in 494 TEM imaging. We acknowledge Nadège Bienvenu (UBO, LM2E) for her help 495 concerning the deposit of the strain at the UBOCC. We thank Erwann Vince and 496 Xavier Philippon for their technical help concerning ionic chromatography. The 497 LABGeM (CEA/Genoscope & CNRS UMR8030), the France Génomique and French 498 Bioinformatics Institute national infrastructures (funded as part of Investissement 499 d'Avenir program managed by Agence Nationale pour la Recherche, contracts ANR-10-INBS-09 and ANR-11-INBS-0013) are acknowledged for support within the 500 501 MicroScope annotation platform. The UMR1136 is supported by the ANR through the 502 Labex Arbre (ANR-11-LABX-0002-01).

503

504 **Funding information**

505 This research was funded by the Sino-French IRP 1211 MicrobSea to K.A. and by

- 506 the UMR6197.
- 507
- 508 **Conflict of interest**
- 509 The authors declare that there are no conflicts of interest.
- 510

511 References

- Ulyshen MD. Wood decomposition as influenced by invertebrates. *Biol Rev* 2016;
 91:70–85.
- Tláskal V, Brabcová V, Větrovský T, Jomura M, López-Mondéjar R, et al.
 Complementary roles of wood-inhabiting fungi and bacteria facilitate deadwood decomposition. *Msystems* 2021; 6:e01078–20.
- 517 3. Johnston SR., Boddy L, Weightman AJ. Bacteria in decomposing wood and 518 their interactions with wood-decay fungi. *FEMS Microbiol. Ecol.* 2016; 92(11).
- Mieszkin S, Richet P, Bach C, Lambrot C, Augusto L, et al. Oak decaying
 wood harbors taxonomically and functionally different bacterial communities in
 sapwood and heartwood. Soil Biol Biochem 2021; 155:108160.
- 522 5. Liu Y, Du J, Pei T, Du H, Feng GD, et al. Genome-based taxonomic 523 classification of the closest-to-*Comamonadaceae* group supports a new family 524 *Sphaerotilaceae* fam. nov. and taxonomic revisions. *Syst Appl Microbiol* 2022; 525 45:126352.
- 6. Bedics A, Táncsics A, Tóth E, Banerjee S, Harkai P, et al. Microaerobic
 enrichment of benzene-degrading bacteria and description of *Ideonella benzenivorans* sp. nov., capable of degrading benzene, toluene and
 ethylbenzene under microaerobic conditions. *Antonie von Leeuwenhoek* 2022;
 115: 1113-1128.
- 531 7. Lechner U, Brodkorb D, Geyer R, Hause G, Härtig C et al. Aquincola
 532 tertiaricarbonis gen. nov., sp. nov., a tertiary butyl moiety-degrading bacterium. Int
 533 J Syst Evol Microbiol 2007; 57:1295–1303.
- 8. Malmqvist Å, Welander T, Moore E, Ternström A, Molin G et al. Ideonella dechloratans gen. nov., sp. nov., a new bacterium capable of growing anaerobically with chlorate as an electron acceptor. Syst Appl Microbiol 1994; 17:58–64.
- Mieszkin S, Pouder E, Uroz S, Simon-Colin C, Alain K. Acidisoma silvae sp. nov. and Acidisoma cellulosilytica sp. nov., two acidophilic bacteria isolated from decaying wood, hydrolyzing cellulose and producing poly-3-hydroxybutyrate. *Microorganisms* 2021; 9:2053.
- 542 10. **Tindall BJ.** A comparative study of the lipid composition of *Halobacterium* 543 *saccharovorum* from various sources. *Syst Appl Microbiol* 1990;13:128–130.
- 544 11. Tindall BJ. Lipid composition of *Halobacterium lacus* profundi. *FEMS Microbiol* 545 Lett 1990; 66: 199–202.

- 546 12. Kuykendall LD, Roy MA, O'Neill JJ, Devine TE. Fatty acids, antibiotic
 547 resistance, and deoxyribonucleic acid homology groups of *Bradyrhizobium* 548 *japonicum. Int J Syst Bact* 1988; 38:358–361.
- 549 13. Jeon CO., Park W, Ghiorse WC, Madsen EL. Polaromonas naphthalenivorans
 550 sp. nov., a naphthalene-degrading bacterium from naphthalene-contaminated
 551 sediment. Int J Syst Evol Microbiol 2004; 54:93–97.
- 14. Blümel S, Busse H J, Stolz A, Kämpfer P. Xenophilus azovorans gen. nov., sp.
 nov., a soil bacterium that is able to degrade azo dyes of the Orange II type. Int J
 Syst Evol Microbiol 2001; 51:1831–1837.
- 555 15. Chen WM, Chen LC, Sheu DS, Tsai JM, Sheu SY. Ideonella livida sp. nov.,
 556 isolated from a freshwater lake. Int J of Syst Evol Microbiol 2020;70:4942–4950.
- 16. **Sheu SY, Hsieh TY, Chen WM.** *Aquincola rivuli* sp. nov., isolated from a freshwater stream. *Int J Syst Evol Microbiol* 2019; 69:2226–2232.
- 17. Spring S, Kampfer P, Ludwig W, Schleifer KH. Polyphasic characterization of the genus *Leptothrix*: new descriptions of *Leptothrix mobilis* sp. nov. and *Leptothrix discophora* sp. nov. nom. rev. and emended description of *Leptothrix cholodnii* emend. Syst Appl Microbiol 1996; 19:634–643.
- 18. Willems A, Gillis M, De Ley J. Transfer of Rhodocyclus gelatinosus to
 Rubrivivax gelatinosus gen. nov., comb. nov., and phylogenetic relationships with
 Leptothrix, Sphaerotilus natans, Pseudomonas saccharophila, and Alcaligenes
 latus. Int J Syst Evol Microbiol 1991; 41:65–73.
- 567 19. **Seemann T, Booth T.** BARNAP: BAsic Rapid Ribosomal RNA Predictor 568 [Internet]. 2020; Berlin: GitHub; 2013. p.
- 569 20. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y. Introducing EzBioCloud: a
 570 taxonomically united database of 16S rRNA gene sequences and whole-genome
 571 assemblies. Int J Syst Evol Microbiol 2017; 67:1613.
- 572 21. Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical
 573 user interface for sequence alignment and phylogenetic tree building. *Mol Biol* 574 *Evol* 2010; 27:221–224
- 575 22. Saitou N. Nei, M. The neighbor-joining method: A new method for reconstructing
 576 phylogenetic trees. *Mol Biol Evol* 1987; 4:406–425.
- 577 23. **Kimura M.** The Neutral Theory of Molecular Evolution. Cambridge: Cambridge 578 University Press; 1983.
- 579 24. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, et al. Phylogeny. fr: robust
 580 phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 2008; 36: W465 581 W469.
- 582 25. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, *et al.* New
 583 algorithms and methods to estimate maximum-likelihood phylogenies: assessing
 584 the performance of PhyML 3.0. *Syst Biol* 2010; 59:307-321.
- 26. Charbonnier F, Forterre P, Erauso G, Prieur D. Purification of plasmids from
 thermophilic and hyperthermophilic archaebacteria. In Archaea: A Laboratory
 Manual; Robb, F.T., Place, A.R., DasSarma, S., Schreier, H.J., Fleischmann,
 E.M., Eds.; Cold Spring Harbor Laboratory Press: Woodbury, NY, USA, 1995; pp.
 87–90.
- 590 27. **Andrews, S.** FastQC: A quality control tool for high throughput sequence data. 591 2010.
- 592 28. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M. SPAdes: a new
 593 genome assembly algorithm and its applications to single-cell sequencing. J
 594 Comput Biol 2012; 19:455–477.

- 595 29. Vallenet D, Calteau A, Dubois M, Amours P, Bazin, A. *et al.* MicroScope: an
 596 integrated platform for the annotation and exploration of microbial gene functions
 597 through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids* 598 *Res* 2020: 48(D1):D579-D589.
- 30. Alteio LV, Schulz F, Seshadri R, Varghese N, Rodriguez-Reillo W, et al.
 Complementary metagenomic approaches improve reconstruction of microbial
 diversity in a forest soil. *msystems* 2020; 5:e00768-19.
- 31. Martin G, Rissanen AJ, Garcia SL, Mehrshad M, Buck M, et al. Candidatus
 Methylumidiphilus drives peaks in methanotrophic relative abundance in stratified
 lakes and ponds across Northern landscapes. *Front Microbiol* 2021; 12:669937.
- 32. Garcia SL, Mehrshad M, Buck M, Tsuji JM, Neufeld JD, et al. Freshwater
 Chlorobia exhibit metabolic specialization among cosmopolitan and endemic
 populations. msystems 2021; 6:e01196-20.
- 33. Rissanen AJ, Saarela T, Jäntti H, Buck M, Peura S, et al. Vertical stratification
 patterns of methanotrophs and their genetic controllers in water columns of
 oxygen-stratified boreal lakes. *FEMS Microbiol Ecol* 2021; 97:fiaa252.
- 611 34. **Eren AM, Kiefl E, Shaiber A, Veseli I, Miller S** *et al.* Community-led, integrated, 612 reproducible multi-omics with anvi'o. *Nat Microbiol* 2021; 6:3–6.
- 613 35. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2:
 614 improving the ultrafast bootstrap approximation. *Mol Biol Evol* 2018; 35:518–522.
- 36. 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.
- 618 37. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to
 619 classify genomes with the Genome Taxonomy Database. *Bioinform* 2020; 6:1925620 1927.
- 38. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH, et al. GTDB-Tk v2:
 memory friendly classification with the Genome Taxonomy Database. *bioRxiv* 2022.
- 39. 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; 9:1-8.
- 40. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based
 species delimitation with confidence intervals nd improved distance functions.
 BMC Bioinform 2013; 14:60.
- 41. Kim D, Park S, Chun J. Introducing EzAAI: a pipeline for high throughput
 calculations of prokaryotic average amino acid identity. *J Microbiol* 2021; 59:476 480.
- 42. Qin QL, Xie BB, Zhang XY, Chen XL, Zhou BC, et al. A proposed genus
 boundary for the prokaryotes based on genomic insights. J Bacteriol 2014;
 196:2210-2215.
- 636 43. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the
 637 prokaryotic species definition. *Proc Natl Acad Sci USA* 2009; 106:19126–19131.
- 44. Wayne L, Brenner DJ, Colwell RR, Grimont PAD, Kandler O et al. Report of
 the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics.
 Int J Syst Bacteriol 1987; 37:463–464.
- 641 45. Konstantinidis KT, Rosselló-Móra R, Amann R. Uncultivated microbes in need
 642 of their own taxonomy. *ISME J* 2017; 11:2399-2406.

643 46. Barco RA, Garrity GM, Scott JJ, Amend JP, Nealson KH *et al.* A genus
644 definition for bacteria and archaea based on a standard genome relatedness
645 index. *mBio* 2020; 11:e02475–19.

List of figures:





Figure 1. Transmission electron micrographs of cells of strain BS-T2-15^T showing (A) intracellular granules (indicated by black arrows), (B) a dividing cell (by binary

fission), (C) a cell with its polar flagellum, and (D) the insertion point of the flagellum.



654 655 Fig. 2. Phylogenomic tree (concatenated alignment of 340 shared gene clusters) showing the phylogenomic position of strain BS-T2-15^T, its two closely related MAGs and 656 representatives of other related taxa. The tree was built using IQ-Tree and the maximum 657 658 likelihood method with 1,000 replicates for the calculation of the ultrafast bootstrap. Bootstrap 659 values >95% are indicated at branching point by a dot. The tree was rooted using Alcaligenes faecalis ZD02 (GCA_000967305.2) as the outgroup. Bar, 0.04 substitutions per 660 661 position.

662 List of tables:

663

664 Table 1. Differential characteristics of strain BS-T2-15^T and type strains of 665 closely related genera.

666 Taxa : 1. BS-T2- 15^{T} , 2. Leptothrix mobilis DSM 10617^T, 3. Rubrivivax gelatinosus DSM 667 1709^T, 4. Ideonella dechloratans CGUG 30977^T, 5. Aquincola tertiaricarbonis DSM 18512^T.

Data are from Mieszkin *et al.* [4], Lechner *et al.* [7], Malmqvist *et al.* [8], Chen *et al.* [15], Sheu *et al.* [16], Spring *et al.* [17] and Willems *et al.* [18]. +, Positive; –, negative; NA, not available; w, weak growth; ^aSummed feature 3 comprises C16: 1 ω 7c and/or C16: 1 ω 6c. All strains are motile, positive for tween 40 hydrolysis, and negative for urease activity and citrate utilization. [§]As all strains have not been grown under exactly the same conditions nor in the same media, percentages of fatty acids cannot be compared from one strain to another.

Characteristics	1	2	3	4	5
Colony color	White-	Dark-brown	Orange-	White	White
	ivory		brown		
Cell size (µm)	0.6-0.9 x	0.6-0.8 x	0.4-0.7 x 1-3	0.7-1 x	0.8-1.1
	1.7-2.7	1.5-12		2.5-5	x 1.2-
					2.3
Temperature range (optimum)	4-30 (20-	10-37	10-45 (37-	15-30	4-40
(°C)	22)		40)	(30)	(30)
pH growth range (optimum)	2-12 (6)	6.5-8.5	5-9 (6-7)	6-8 (6-7)	5-9 (6-
NoOL (alaraman (antimum) (0)			0.0(4)	0.0	<i>(</i>)
Naci tolerance (optimum) (%,	0-0.5 (0)	NA	0-2 (1)	0-2	0-1 (0)
W/V) Corbon courses utilization					
Acelale	-	-	+	+	+
	+ \\/	_	τ ΝΛ	+	+
	vv	_		-	+
ryiuvale Arabiagaa	+	— N A	+	+	+ NIA
Arabinose	+		_	-	
	+		-	_	NA
CM-cellulose	-	NA	_	+	_
D-Glucose	+	-	-	+	+
D-Maltose	_	-	+	-	+
D-Mannose	-	- <u>,</u>	-	+	+
D-Mannitol	_	+/-	-	-	+
N-Acetylglucosamine	W	-	-	-	+
Indole formation	-	-	+	-	-
Enzymatic activities					
Catalase	+	NA	-	+	+
Oxidase	-	+	+	+	+
Nitrate reductase	+	NA	+	+	-
α-Glucosidase	+	NA	-	-	+
Quinone type	Q-8	Q-8	Q-8, MK-8	Q-8	Q-8
Polar lipids					
Diphosphatidylglycerol (DPG)	+	_	+	+	+
Major fatty acids (%) [§]					
14: 0 2-OH	10.49	_	_	_	_
16: 1 <i>cis</i> -9	_	53.04	_	_	39
16: 1 <i>ω7</i> c	40.45	_	_	_	_
Summed feature 3 ^a	_	_	42.8	40.2	_
DNA G+C content (mol%)	69.6	68	71.9	68.1	70.5
Isolation source	Forest	Freshwater	Acetate	Activated	Aquifer

soilsedimentenrichmentsludge676Table 2. Comparison of whole-cell fatty acid profiles (% of the total) of strain677BS-T2-15^T with type strains of closely related genera.

Taxa: 1. BS-T2-15^T, 2. Leptothrix mobilis DSM 10617^T, 3. Rubrivivax gelatinosus DSM 678 1709^T, 4. Ideonella dechloratans CGUG 30977^T, 5. Aquincola tertiaricarbonis DSM 18512^T. 679 680 Values are percentages of the fatty acids that were assigned to fatty acids in the peaknaming table of the MIS database (MIDI, Microbial ID, Newark, DE 19711 U.S.A.). The 681 682 nomenclature is as follows: the first number indicates the number of carbon atoms in the 683 molecule; 'OH' and 'cyclo' indicate hydroxy or cyclic fatty acids; the second number following 684 the colon indicates the number of double bonds present. The position of the double bond is 685 indicated by the carbon atom position starting from the methyl (ω) end of the molecule. c, cis isomer. Data are from Lechner et al. [7], Malmqvist et al. [8], Sheu et al. [16] and Spring et al. 686 [18]. Major fatty acids (>10% of the total fatty acids) are indicated in bold. 687

Fatty acids	1	2	3	4	5
10: 0 3-OH	5.99	4.54	4.8	2.4	2
12: 0	_	2.43	3.7	2.1	4
12: 0 2-OH	0.47	_	_	2.5	_
12: 0 3-OH	3.17	_	_	4.1	_
14: 0	3.14	0.77	5.4	1.5	2
14: 0 2-OH	10.49	_	_	_	_
15: 0	_	_	_	_	3
15: 1	_	_	_	_	2
15: 1 <i>ω6</i> c	_	_	1.7	_	_
16: 0	25.18	30.29	33.1	31.5	37
16: 1 <i>cis</i> -9	_	53.04	_	_	39
16: 1 <i>ω7c</i>	40.45	_	_	_	_
17: 0	_	_	1.1	_	2
17: 0 cyclo	_	_	_	_	2
17: 0 cyclo <i>ω7c</i>	1.44	_	_	_	_
18: 0	_	1.19	_	_	1
18: 1	_	_	_	_	6
18: 1 <i>ci</i> s-9,11ª	_	6.99	_	_	_
18: 1 <i>ω7</i> c	9.68	_	4.5	12.7	_
19: 1 <i>cis</i> -9,10 ^ь	_	0.84	_	_	_
Summed feature	_	_	42.8	40.2	_
3 _c					

. .

688 –, Not detected

^aSum of both compounds: cis-9octadecenoic acid and cis-11-octadecenoic acid

690 ^b cis-9,10-methyleneoctadecanoic acid

691 ^cSummed feature 3 comprises C16: 1 ω 7c o and/or C16: 1 ω 6c

Table 3. Genome statistics and overall genome relatedness indices (OGRI) between

strain BS-T2-15^T and type strains of closely related genera and closely related MAGs. Taxa : 1. BS-T2-15^T, 2. Leptothrix mobilis DSM 10617^T, 3. Rubrivivax gelatinosus DSM 1709^T, 4. Ideonella dechloratans CGUG 30977^T, 5. Aquincola tertiaricarbonis DSM 18512^T. 6. Azohydromonas lata NBRC 102462^T, 7. Roseateles depolymerans KCTC 42856^T, 8. MAG GCA_013289985.1, 9. MAG_GCA_903845345.1.

	1	2	3	4	5	6	7	8	9
Number of contigs	28	17	36	158	98	342	1	210	631
Size (Mb)	6.27	4.65	5.08	4.51	6.32	7.18	5.68	6	5.4
GC content (%)	69.56	69.00	71.40	69.30	70.20	69.00	66.60	69.55	69.50
Number of CDS	5,712	4,019	4,691	4,172	5,711	6,349	4,894	5,441	5,362
rRNA	3	6	3	7	3	4	12	3	1
tRNA	55	49	46	59	43	46	58	19	17
dDDH (%)	100	20.50	20.40	20.10	20.30	20.30	20.00	30.10	33.40
ANI (%)	100	79.50	79.70	79.70	79.60	78.60	78.50	87.80	88.90
AAI (%)	100	65.34	66.03	66.57	65.48	65.32	64.27	84.14	86.43
POCP (%)	100	43.17	44.84	47.33	49.02	40.89	49.27	72.60	74.76
AF	1.0	0.42	0.48	0.43	0.40	0.32	0.29	0.71	0.80