INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

TAXONOMIC DESCRIPTION

Kweon *et al.*, *Int. J. Syst. Evol. Microbiol.* 2022;72:005381
DOI 10.1099/ijsem.0.005381

MICROBIOLOGY SOCIETY

Arsenicicoccus cauae sp. nov., isolated from the blood of a pediatric gastroenteritis patient

Oh Joo Kweon[1]†, Kyung Hyun Kim[2]†, Hyung Min Kim[2], Shehzad Abid Khan[2], Ji Hun Jeong[3], Yong Kwan Lim[1], Hye Ryoun Kim[1], Che Ok Jeon[2]* and Mi-Kyung Lee[1]*

Abstract

A Gram-stain-positive coccus was isolated from the blood of a paediatric patient suffering from gastroenteritis. The taxonomic position of this catalase-positive, non-motile, non-spore-forming facultative anaerobe designated as strain MKL-02^T was investigated using a polyphasic approach. Colonies grown on tryptic soy agar with 10% sheep blood were circular, creamy yellow, and convex. Phylogenetic analysis based on 16S rRNA gene and whole-genome sequences revealed that this strain was most closely related to *Arsenicicoccus bolidensis* CCUG 47306^T within the cluster of the genus *Arsenicicoccus*. Average nucleotide identity and digital DNA–DNA hybridization values between strain MKL-02^T and *A. bolidensis* DSM 15745^T, *A. dermatophillus* DSM 25571^T and *A. piscis* DSM 22760^T were 89.5 and 37.0%, 79.6 and 22.4%, and 75.9 and 21.0%, respectively. The genomic size of strain MKL-02^T was 3423857 bp with a 72.7 mol% G+C content. Growth was observed at 10–45 °C (optimum, 37–40 °C) and pH 6.0–10.0 (optimum, pH 7.0), in the presence of 0–10% (w/v) NaCl (optimum, 0.5%). Cells of strain MKL-02^T were non-motile cocci and 0.50–0.60 µm long, as determined by transmission electron microscopy. The strain was catalase-positive and oxidase-negative. The major fatty acid type (>10% of total) was $C_{15:0}$. The polar lipid profile consisted of two unidentified phospholipids, three unidentified lipids and an unidentified aminophospholipid. The strain contained MK-8 (H_4) as the predominant menaquinone. Based on phylogenetic and phenotypic considerations, it is proposed that strain MKL-02^T be classified as a new species, named *Arsenicicoccus cauae* sp. nov. The type strain is MKL-02^T (=NCCP 16967^T=JCM 34624^T).

The genus *Arsenicicoccus* is characterized by Gram-stain-positive, catalase-positive, non-motile, non-spore-forming facultative anerobic cocci [1–3]. Initially, this genus was first proposed as a member of the family *Intrasporangiaceae*, with *Arsenicicoccus bolidensis* as a type species [1], recovered from arsenic enrichment and reported to be able to reduce arsenic [1, 4]. After reclassification, the genus *Arsenicicoccus* is currently classified within the family *Dermatophilaceae* [5]. The genus *Arsenicicoccus* includes three validly published taxa isolated from the foot skin of a flamingo, the gastrointestinal tract of fish, and lake sediment. In the present study, a presumably novel bacterium belonging to the genus *Arsenicicoccus* was isolated from human blood obtained for investigating the cause of fever in a paediatric patient and characterized taxonomically using a polyphasic approach. The name *Arsenicicoccus cauae* sp. nov. is proposed for this strain; the type strain is MKL-02^T.

ISOLATION AND ECOLOGY

Strain MKL-02^T was isolated from the blood sample of a 17-month-old male patient with symptoms of fever, diarrhoea, non-projectile vomiting, and abdominal pain. Molecular testing for viral pathogens including rotavirus, norovirus, enteric adenovirus, and astrovirus as well as stool culture for human gastrointestinal tract bacterial pathogens showed negative results. However,

Author affiliations: [1]Department of Laboratory Medicine, Chung-Ang University College of Medicine, 84 Heukseok-ro, Dongjak-gu, Seoul, Republic of Korea; [2]Department of Life Science, Chung Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, Republic of Korea; [3]Department of Laboratory Medicine, Daejeon Eulji Medical Center, Eulji University, 95 Dunsanseo-ro, Seo-gu, Daejeon, Republic of Korea.
***Correspondence:** Mi-Kyung Lee, cpworld@cau.ac.kr; Che Ok Jeon, cojeon@cau.ac.kr
Keywords: *Arsenicicoccus cauae*; taxonomy, new taxa; phylogenetic analysis.
Abbreviations: ANI, average nucleotide identity; BAP, blood agar plate; DDH, DNA–DNA hybridization; ML, maximum-likelihood; MP, maximum-parsimony; NJ, neighbour-joining; TSA, tryptic soy agar; TSB, tryptic soy broth.
The GenBank accession numbers for the 16S rRNA gene and genome sequences of strain MKL-02^T are MN629181 and WLVL00000000, respectively. The GenBank accession numbers for the genomes of *A. dermatophillus* DSM 25571^T and *A. piscis* DSM 22760^T are JAKZHU000000000 and JAKZHV000000000, respectively.
†These authors contributed equally to this work
Two supplementary figures and one supplementary table are available with the online version of this article.

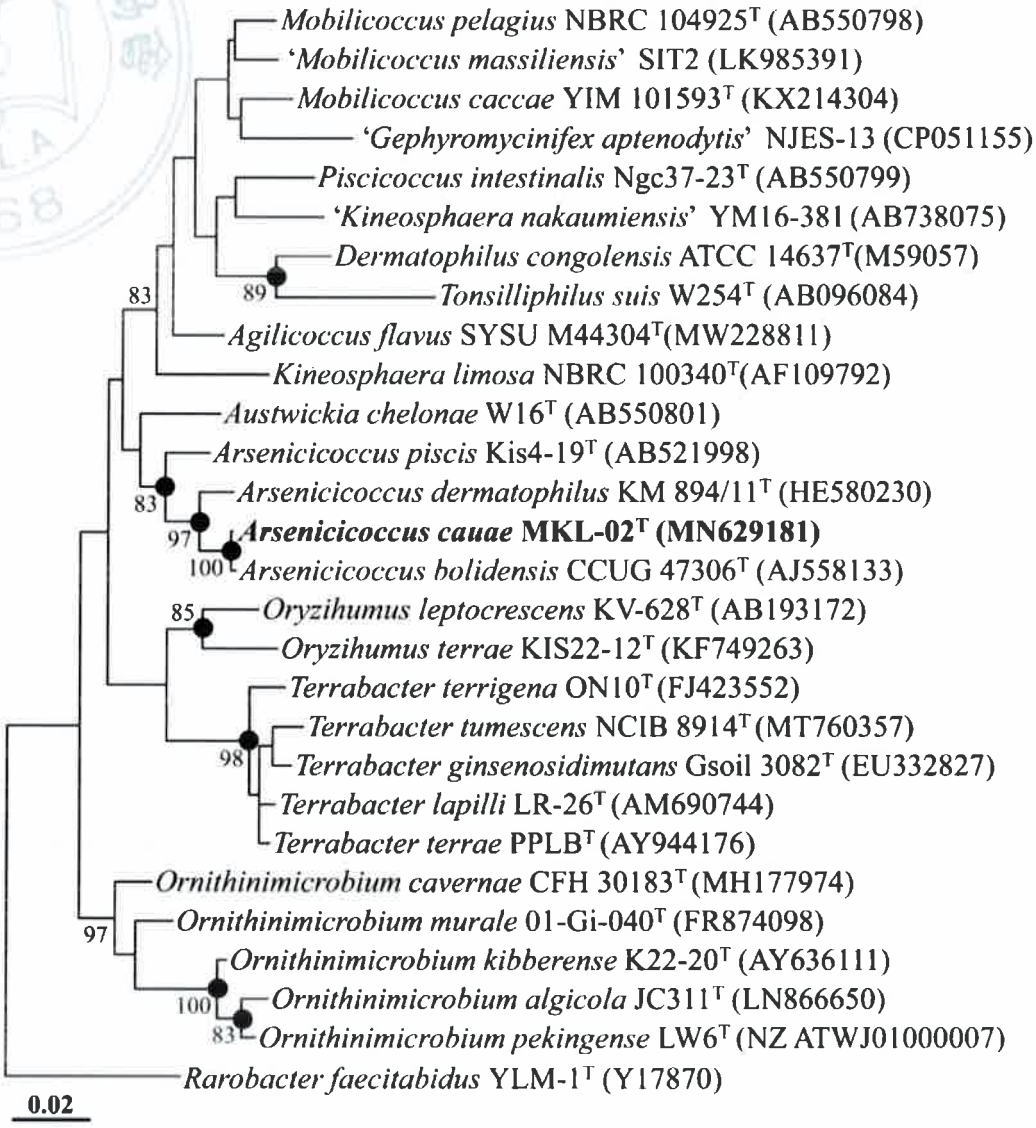

Fig. 1. Maximum-likelihood tree showing the phylogenetic relationships of strain MKL-02^T and its closely related taxa, based on 16S rRNA gene sequences. Bootstrap values over 70% are shown on the nodes as percentages of 1000 replicates. Filled circles (●) indicate that the corresponding nodes were also formed in the trees generated using the neighbour-joining and maximum-parsimony algorithms. *Rarobacter faecitabidus* YLM-1^T (Y17870) was used as an outgroup. The scale bar equals 0.01 changes per nucleotide position.

bacteria were detected in the peripheral blood culture performed using an aerobic paediatric blood culture bottle (BacT/Alert PF Plus, bioMérieux) after approximately 48 h of incubation at 35 °C using the BacT/Alert 3D blood culture system (bioMérieux). The blood in the culture bottle was subcultured onto blood agar plates (BAPs; Synergy Innovation) and chocolate agar plates (Synergy Innovation) and then incubated at 37 °C and 5% CO_2 for 72 h. Colonies on both media were dry, pale yellow, and pinpoint. The isolates were Gram-stain-positive clusters of cocci and presented catalase- and coagulase-positivity. Strain MKL-02^T was the only isolate that grew on BAP medium without any accompanying bacteria. Further investigations to determine the taxonomic status of strain MKL-02^T were conducted.

16S rRNA GENE PHYLOGENY

The 16S rRNA gene of strain MKL-02^T was amplified with the universal primers, 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1492R (5′-TACGGYTACCTTGTTACGACTT-3′) for 1446 bp amplicons [6]. Direct Sanger sequencing of the amplicon was performed using the Applied Biosystems 3500 Dx Genetic Analyzer (Thermo Fisher Scientific). The 16S rRNA gene sequence of

Table 1. General genome features of strain MKL-02^T and the type strains of the genus *Arsenicicoccus**

Strain: 1, MKL-02^T; 2, *A. bolidensis* DSM 15745^T; 3, *A. dermatophilus* DSM 25571^T; 4, *A. piscis* DSM 22760^T.

Genomic feature		1*	2	3*	4*
Genome size (kb)		3423.857	3995.323	3518.211	3602.737
Coverage		281.8×	–	50.0×	76.0×
No. of contigs		60	86	11	12
N50 values (kb)		160.7	94.1	2878.9	613.9
G+C content (mol%)		72.7	71.8	72.6	71.3
No. of total genes		3182	3811	3213	3366
No. of total protein coding genes		3031	3594	3028	3176
No. of pseudogenes		95	155	125	137
Total tRNA genes		48	48	50	44
GenBank accession no.		WLVL00000000	AUFG00000000	JAKZHU000000000	JAKZHV000000000
			Digital DDH value (%)†		
ANI value (%)†	1	–	37.0	22.4	21.0
	2	89.5	–	22.7	21.0
	3	79.6	79.6	–	21.2
	4	75.9	76.1	75.6	–

*All genomes were sequenced in this study and their genomic features were obtained from GenBank annotated by the NCBI prokaryotic genome annotation pipeline (www.ncbi.nlm.nih.gov/genome/annotation_prok/).
†DDH, DNA–DNA hybridization; ANI, average nucleotide identity.

strain MKL-02^T was compared with those of all other reported type strains using the nucleotide similarity search program in the EzBioCloud server (www.ezbiocloud.net/identify) [7] and BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) [8]. The 16S rRNA gene sequences of closely related species of the genus *Arsenicicoccus* were obtained from GenBank (www.ncbi.nlm.nih.gov/genbank/) [9]. Phylogenetic trees were reconstructed based on the 16S rRNA gene sequences using the maximum-likelihood (ML), neighbour-joining (NJ), and maximum-parsimony (MP) algorithms in the MEGA11 software [10]. Bootstrap values were determined based on 1000 replications.

The comparison of 16S rRNA gene sequences between strain MKL-02^T and other type strains showed that strain MKL-02^T was most closely related to *A. bolidensis* CCUG 47306^T (99.8% similarity), *Arsenicicoccus dermatophilus* KM 894/11^T (98.2%), and *Arsenicicoccus piscis* Kis4-19^T (97.4%). A phylogenetic tree based on the ML algorithm showed that strain MKL-02^T formed a phylogenic lineage with *A. bolidensis* CCUG 47306^T (100% bootstrap value) within the genus *Arsenicicoccus* (Fig. 1). Phylogenetic trees based on the NJ and MP algorithms also confirmed that strain MKL-02^T was tightly clustered with *A. bolidensis* CCUG 47306^T within the genus *Arsenicicoccus* (Fig. S1, available in the online version of this article).

GENOME FEATURES

The genomic DNA of strain MKL-02^T, *A. dermatophilus* DSM 25571^T and *A. piscis* DSM 22760^T was extracted using the Wizard Genomic DNA Purification Kit (Promega), according to the manufacturer's instructions. For the genome sequencing of strain MKL-02^T, the concentration of the genomic DNA of strain MKL-02^T was measured using the Quat-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific). A DNA library was constructed using the TruSeq Nano DNA LT Library Prep Kit (Illumina, San Diego, CA) according to the manufacturer's instructions. The library was quantified using Bioanalyzer 2100 (Agilent Technologies) and a DNA 7500 Kit (Agilent Technologies). Whole-genome sequencing was performed on the Illumina MiSeq platform (2×300 bp; Illumina). The genomic DNA of strains DSM 25571^T and DSM 22760^T was sequenced using an Oxford Nanopore MinION platform. The sequencing reads of strains DSM 25571^T and DSM 22760^T were *de novo*-assembled using Unicycler (version 0.4.7) [11]. Any possible contamination of genomic data with other organisms was screened using the ContEst16S algorithm (ChunLab), in which 16S rRNA gene fragments are screened to determine whether the genome assembly is contaminated [12]. The sequenced genomes of strain MKL-02^T, *A. dermatophilus* DSM 25571^T and *A. piscis* DSM 22760^T have been deposited in GenBank under the accession numbers of WLVL00000000, JAKZHU000000000 and JAKZHV000000000, respectively. For

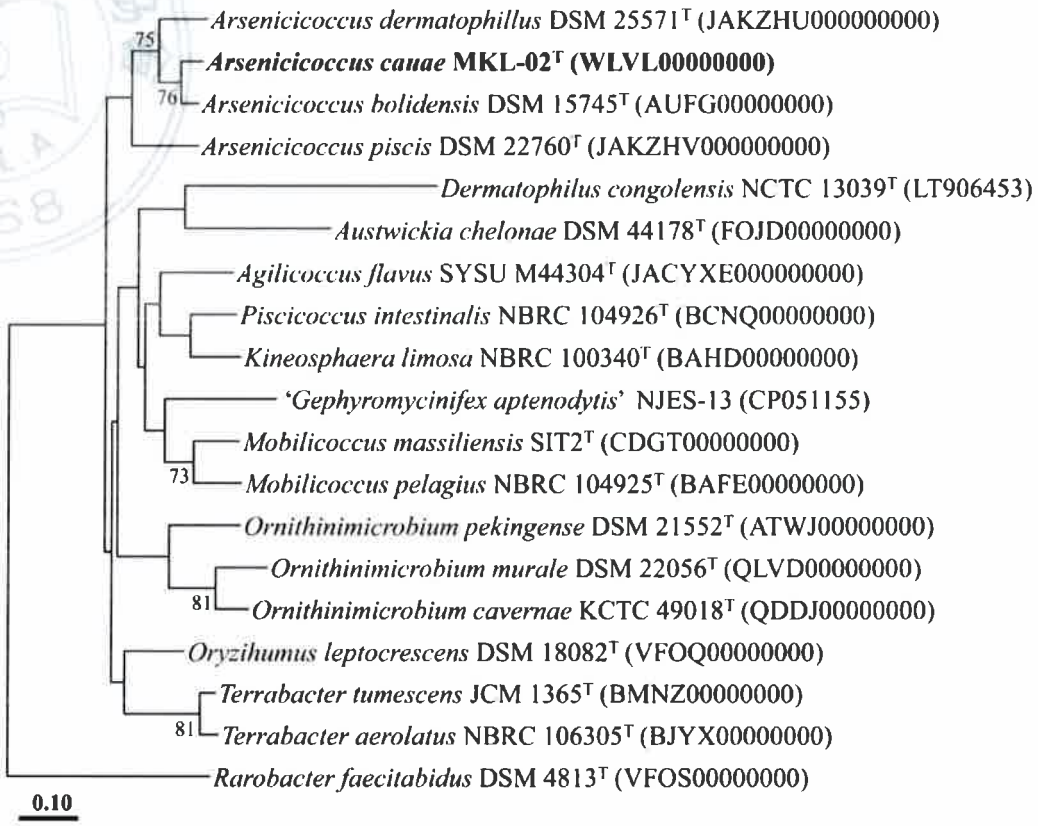

Fig. 2. Phylogenomic tree based on the 81 concatenated core housekeeping genes showing the phylogenetic relationships of strain MKL-02^T and its closely related taxa. Bootstrap values are shown on the nodes as percentages of 1000 replicates for values over 70%. *Rarobacter faecitabidus* DSM 4813^T (VFOS00000000) was used as an outgroup. The scale bar equals 0.1 changes per nucleotide position.

genome-based phylogenomic analysis, the up-to-date bacterial core gene (UBCG2) pipeline (http://leb.snu.ac.kr/ubcg2) [13] was used to extract 81 core housekeeping genes from the genomes of strain MKL-02^T and closely related taxa. An ML tree was reconstructed with bootstrap values (1000 replications) based on the concatenated 81 core housekeeping genes using MEGA software. The average nucleotide identity (ANI) and digital DNA–DNA hybridization (DDH) values between strain MKL-02^T and the most closely related species were determined using the EzTaxon-e server (www.ezbiocloud.net/tools/ani) [14] and the server-based Genome-to-Genome Distance Calculator version 2.1 (http://ggdc.dsmz.de/distcalc2.php), respectively [15].

The genome sizes of strain MKL-02^T, *A. dermatophilus* DSM 25571^T and *A. piscis* DSM 22760^T were 3.42, 4.00 and 3.60 Mb, respectively, and their sequencing summaries and general genomic features are presented in Table 1. The genomic DNA G+C contents of strains MKL-02^T, DSM 25571^T and DSM 22760^T were 71.9, 72.6 and 71.3mol%, respectively, which were similar to that of *A. bolidensis* DSM 15745^T, the type species of the genus *Arsenicicoccus*. ANI and digital DDH values between strain MKL-02^T and *A. bolidensis* DSM 15745^T (GenBank accession no., AUFG00000000), *A. dermatophillus* DSM 25571^T, and *A. piscis* DSM 22760^T were 89.5 and 37.0%, 79.6 and 22.4%, and 75.9 and 21.0%, respectively, which were clearly lower than the thresholds (ANI 95% and DDH 70%) for prokaryotic species delineation [16]. A phylogenomic tree based on the 81 core housekeeping gene sequences showed that strain MKL-02^T formed a phyletic lineage with *A. bolidensis* DSM4813^T and was clustered with the members of the genus *Arsenicicoccus* (Fig. 2). In conclusion, the phylogenetic and genome relatedness analyses based on 16S rRNA gene and whole-genome sequences indicated that strain MKL-02^T represents a novel species belonging to the genus *Arsenicicoccus*.

The genomic analysis revealed that strain MKL-02^T has complete gene sets linked to glycolysis, the pentose phosphate pathway, and the tricarboxylic acid cycle to metabolize carbon compounds, indicating that they can grow aerobically. In addition, strain MKL-02^T harbours genes encoding respiratory nitrate reductase, which may contribute to the facultative anaerobic growth of strain MKL-02^T. Because strain MKL-02^T was isolated from a human blood sample, genes associated with pathogenicity and virulence were analysed. The genome of strain MKL-02^T harboured six copies of genes encoding haemolysin family proteins (locus_tags: GGG17_10345, 01110, 01115, 12290, 15220 and 15225), which was in agreement with the phenotypic result of strain MKL-02^T showing haemolysis on tryptic soy agar (TSA) containing 10% sheep blood.

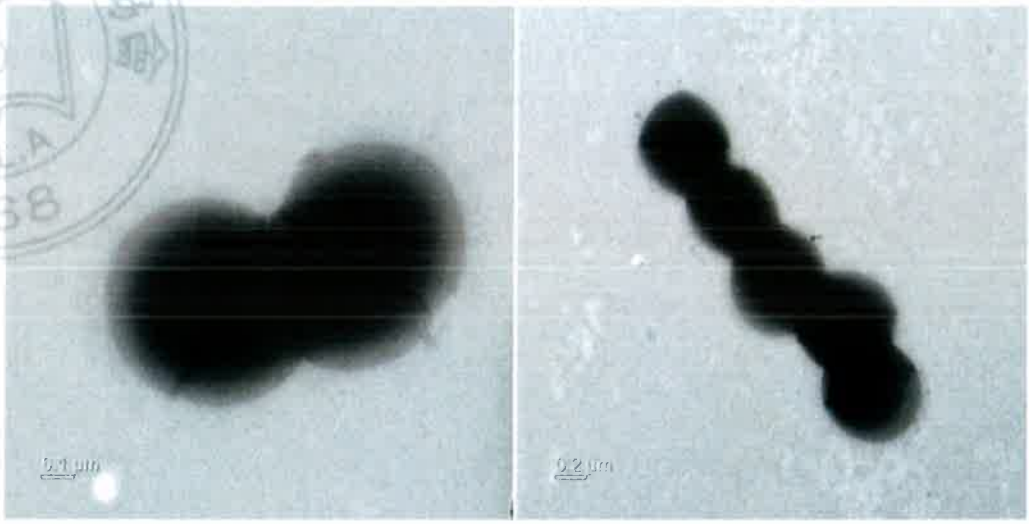

Fig. 3. Transmission electron microscopy images showing the general morphology of negatively stained cells of strain MKL-02^T grown on BAP for 5 days at 37 °C.

An MviN family virulence factor-coding gene (locus_tag: GGG17_07455) reported to regulate virulence in the host [17] was identified in the genome of strain MKL-02^T. Bacterial toxin-antitoxin systems have also been considered potential virulence factors [18]. Two complete toxin-antitoxin systems, including Phd/YefM family and VapBC (locus_tags: GGG17_04150 and 04155 and GGG17_04120 and 11665, respectively), three incomplete toxin genes (locus_tag: GGG17_12175, 12315, and 13725), and two antitoxin genes (locus_tag: GGG17_07570 and 07660), were identified in the genome of strain MKL-02^T. A gene cluster consisting of *brkA* and *brkB* in *Bordetella pertussis* was reported to confer resistance to complement-dependent killing caused by human serum [19]. However, strain MKL-02^T was shown to have three copies of only a *brkB* gene homolog (locus_tag: GGG17_00710, 04380 and 05040). It is not known whether this homolog confers the observed resistance of strain MKL-02^T to human serum; further studies are necessary to explore this resistance.

PHYSIOLOGY AND CHEMOTAXONOMY

Based on the phylogenetic analysis, the cells of strain MKL-02^T, *A. bolidensis* DSM 15745^T, *A. dermatophilus* DSM 25571^T, and *A. piscis* DSM 22760^T were selected and obtained from their corresponding collection centres for use as reference strains for the comparison of phenotypic properties and fatty acid composition. They were harvested during their exponential growth phase (optical density of approximately 0.8 at 600 nm) under their optimal growth conditions.

The growth of strain MKL-02^T was evaluated at 37 °C for 5 days in several bacteriological agar media (all from BD), including marine agar, Luria–Bertani, R2A, nutrient agar, TSA) TSA with 10% sheep blood, TSA with 10% glucose, and TSA with 10% starch. To determine the optimal growth of strain MKL-02^T, the isolate was cultivated at different temperatures (0–40 °C with 5 °C intervals) on TSA with 10% sheep blood medium. In addition, strain MKL-02^T was cultivated at different pH values from 3.0 to 11.0 at intervals of 1.0 pH unit. The pH values of tryptic soy broth (TSB) were adjusted using the following buffers: for pH 3.0–5.0, sodium citrate; for pH 6.0–7.0, Na_2HPO_4-NaH_2PO_4; for pH 8.0–9.0 Tris-HCl; for pH 10.0–11.0 sodium carbonate, and if necessary, adjusting the pH values again after autoclaving for 15 min at 121 °C and then adding 10% sheep blood. The NaCl tolerance was tested in TSB with 10% sheep blood (0–20% (w/v) at an interval of 1%) [20]. Gram staining was carried out according to the manufacturer's instructions (bioMérieux). Characterization of the cell morphology was performed after culturing the strain for 3 days at 37 °C using transmission electron microscopy (Zeiss LEO906). Oxidase and catalase tests were carried out according to standard methods [21]. The hydrolysis of casein, Tween 20, Tween 80, aesculin, tyrosine and starch was examined according to previously described methods [22]. Additional biochemical and enzymatic activities were examined to identify the profiles of assimilation and conversion capacities of strain MKL-02^T and closely related reference strains for different substrates using the API 20NE and API ZYM kits (bioMérieux) according to the manufacturer's instructions. For the phenotypic and physiological characteristics of the closely related reference strains, the results of hydrolysis, enzyme activity, assimilation, and antibiotic resistance were obtained in the present study, and the remaining data were from previous studies [1–3].

Strain MKL-02^T could grow on TSA with 10% sheep blood under aerobic and anaerobic conditions whereas it could only grow on TSA with 10% glucose under anaerobic conditions. These data suggested that strain MKL-02^T was a facultative anaerobe. Colonies

Table 2. Differential phenotypic and physiological characteristics of strain MKL-02^T and type strains in the genus *Arsenicicoccus*

Strains: 1, MKL-02^T (this study); 2, *Arsenicicoccus bolidensis* DSM 15745^T [1]; 3, *Arsenicicoccus dermatophilus* DSM 25571^T [2]; 4, *Arsenicicoccus piscis* DSM 22760^T [3]. All strains are negative for enzymatic activities of lipase (C14), trypsin, β-glucuronidase, α-mannosidase and α-fucosidase, indole production, glucose fermentation, and assimilation of L-arabinose, D-mannose, *N*-acetylglucosamine and adipic acid; as well as hydrolysis* of Tween 80, Tween 20, and starch. All strains are positive for enzymatic activities of arginine dihydrolase and urease. Symbols: +, positive; −, negative; NA, not available.

Characteristics	1*	2	3	4
Isolation source	Human	Lake sediment (containing mine waste)	Skin of greater flamingos	Fish
Cell morphology	Coccus	NA	NA	NA
Cell size (μm)	0.6	NA	1	0.7
Flagellation	−	−	−	−
Catalase	+	+	+	+
Colony colour	Yellow–white	Yellow	Yellow–white	Greyish cream
Range for growth:				
Temperature (°C)	30–37	5–37	NA	15–37
pH	5–10	6–9	NA	5–8
NaCl	0–7	0–7	NA	0–5
Nitrate reduction	+	+	+	−
Hydrolysis of:				
Tyrosine, casein	−	−	−	+
Aesculin	+	+	−	+
Enzyme activity of:*				
Alkaline phosphatase, esterase lipase (C8)	+	−	+	−
Esterase (C4), acid phospatase	+	−	+	+
Leucine arylamidase	+	−	−	−
Valine arylamidase	+	+	−	−
Cystine arylamidase	−	+	−	−
β-Chymotrypsin, β-galactosidase, β-galactosidase, β-glucosidase, β-glucosidase	−	−	−	+
Naphthol-AS-BI-phosphohydrolase	+	+	+	−
N-Acetyl-β-glucosaminidase, gelatinase	−	−	+	−
Assimilation of:*				
D-Glucose, D-mannose, D-mannitol	+	−	−	−
Potassium gluconate, trisodium citrate	−	−	+	−
Capric acid, phenylacetic acid	−	+	−	−
Maltose	+	+	−	+
Malic acid	+	+	+	−

*These data were obtained from this study under the same conditions

of strain MKL-02^T on TSA with 10% sheep blood agar after incubation for 72 h at 37 °C were circular, cream yellow, convex, and 1.3 mm in diameter. Strain MKL-02^T consisted of Gram-stain-positive and non-motile cocci. Cells of strain MKL-02^T were 0.50–0.60 μm long (Fig. 3). Growth occurred at 30–37 °C with an optimal growth temperature of 37 °C in TSA with 10% sheep blood. The NaCl concentration range for growth was 0–7% NaCl in TSB with 10% sheep blood, with an optimal concentration of

Table 3. Cellular fatty acid composition (%) of strain MKL-02^T and the type strains of the genus *Arsenicicoccus*

Strain: 1, MKL-02^T; 2, *Arsenicicoccus bolidensis* DSM 15745^T; 3, *Arsenicicoccus dermatophilus* DSM 25571^T; 4, *Arsenicicoccus piscis* DSM 22760^T. Data are expressed as percentages of the total fatty acids; fatty acids amounting to <0.5% in all strains are not shown. Major components are indicated with bold text. TR, Trace amount (<0.5%); –, not detected.

Fatty acid	1	2	3	4
Saturated:				
$C_{10:0}$	1.1	2.4	–	–
$C_{14:0}$	9.2	2.8	2.3	1.1
$C_{16:0}$	4.3	7.3	2.0	3.8
$C_{17:0}$	–	2.5	1.4	5.2
$C_{18:0}$	TR	6.2	0.5	2.6
Unsaturated:				
$C_{15:1}\omega6c$	TR	3.4	1.1	2.6
$C_{16:1}\omega9c$	1.2	–	–	–
$C_{17:1}\omega8c$	TR	7.5	**14.1**	6.8
anteiso-$C_{17:1}\omega9c$	–	1.5	2.0	1.9
Branched:				
iso-$C_{10:0}$	–	3.5	1.0	2.5
iso-$C_{11:0}$	8.4	–	–	6.4
iso-$C_{13:0}$	2.4	0.7	0.8	0.8
iso-$C_{14:0}$	6.3	7.8	3.5	3.2
iso-$C_{15:0}$	**31.0**	**23.1**	**28.0**	**25.5**
iso-$C_{16:0}$	1.2	0.9	4.2	3.9
iso-$C_{19:0}$	1.6	–	2.0	–
anteiso-$C_{13:0}$	0.9	1.4	0.8	–
anteiso-$C_{15:0}$	**11.6**	1.4	6.8	**22.5**
anteiso-$C_{16:0}$	–	1.0	–	1.7
anteiso-$C_{17:0}$	TR	1.3	1.4	5.0
Hydroxy:				
$C_{12:0}$ 2-OH	0.6	–	–	–
$C_{12:0}$ 3-OH	1.1	TR	–	TR
iso-$C_{11:0}$ 3OH	2.2	0.8	–	1.5
iso-$C_{12:1}$ 3OH	0.7	–	–	–
iso-$C_{13:0}$ 3OH	1.5	–	–	–
Summed features:*				
3	4.8	9.8	6.6	1.3
5	TR	5.4	–	0.6
8	TR	7.3	–	0.9
9	1.4	–	3.0	–

*Summed features are fatty acids that cannot be resolved reliably from another fatty acid using the chromatographic conditions chosen. The MIDI system groups these fatty acids together as one feature with a single percentage of the total. Summed feature 3, $C_{16:1}\omega7c$ and/or $C_{16:1}\omega6c$; summed feature 5, $C_{18:0}$ ante/$C_{18:2}\omega6,9c$ and/or $C_{18:2}\omega6,9c$ /$C_{18:0}$ ante; summed feature 8, $C_{18:1}\omega7c$ and/or $C_{18:1}\omega6c$, summed feature 9, $C_{16:0}$ 10-methyl and/or iso-$C_{17:1}\omega9c$

3.0%. Growth was found to occur over a pH range of 5.0–10.0 in TSB with 10% sheep blood, and the optimal pH was 7.0. Strain MKL-02^T was catalase-positive and oxidase-negative. The results of assimilation and conversion capacities of strain MKL-02^T and the closely related reference strains for different substrates are listed in Table 1. Strain MKL-02^T and related reference strains are all negative for enzymatic activities of lipase (C14), trypsin, β-glucuronidase, α-mannosidase, and α-fucosidase; indole production; glucose fermentation; and the hydrolysis of Tween 80, Tween 20, and starch. All strains are positive for enzymatic activities of arginine dihydrolase and urease. Assimilation of L-arabinose D-mannose, *N*-acetylglucosamine, and adipic acid was not detected. However, strain MKL-02^T showed many phenotypic characteristics that differed from its related type strains, including different growth conditions and several other enzymatic and assimilation activities (Table 2). Data for phenotypic and physiological characteristics are from this study and other studies.

For the analysis of fatty acid composition, strain MKL-02^T and related type strains were cultured in TSB medium at 37 °C for 3 days under their optimal growth conditions and harvested during their exponential growth phases (optical density: 0.8 at 600 nm). Fatty acid analysis was based on previously described methods [23], including the subsequent steps of saponification, methylation, and fatty acid extraction from the harvested cells followed by gas chromatography for the investigation of fatty acid methyl esters. The results were analysed with the RTSBA6 database of the Microbial Identification System (Sherlock version 6.0B) [23]. The polar lipids of strain MKL-02^T, *A. bolidensis* DSM 15745^T and *A. dermatophilus* DSM 25571^T were analysed by two-dimensional thin-layer chromatography (TLC) following the approach previously described by Minnikin *et al.* [24]. The solvent systems used were chloroform–methanol–water (65:25:4, v/v) for developing the first dimension and chloroform–methanol–acetic acid–water (80:12:15:4, v/v) for developing the second dimension. The detection of polar lipids was carried out using the following reagents: 10% ethanolic molybdophosphoric acid specific for total polar lipids, ninhydrin specific for aminolipids, Dittmer–Lester reagent specific for phospholipids, and α-naphthol specific for glycolipids. Four TLC plates were prepared for each sample. The respiratory quinones for strain MKL-02^T were analysed according to the method of Minnikin *et al.* [25], utilizing an LC-20A high-performance liquid chromatography system (Shimadzu) equipped with a diode array detector (SPD-M20A, Shimadzu) and a reversed-phase column (250×4.6 mm; Kromasil).

The predominant fatty acids of strain MKL-02^T are composed of saturated fatty acid $C_{14:0}$ (9.2%), iso-$C_{11:0}$ (8.4%), iso-$C_{14:0}$ (6.3%), iso-$C_{15:0}$ (31.0%) and anteiso-$C_{15:0}$ (11.6%). In contrast to other *Arsenicicocus* species, strain MKL-02^T has higher amounts of $C_{14:0}$, iso-$C_{11:0}$ and iso-$C_{15:0}$, and lower amounts of $C_{17:1}$ ω8*c* and anteiso-$C_{17:0}$. Details of the cellular fatty acid composition of strain MKL-02^T and related strains are listed in Table 3. Major polar lipids detected in strain MKL-02^T were two unidentified phospholipids, three unidentified lipids, and an unidentified aminophospholipid, which were different from those in other *Arsenicicoccus* species (Fig. S2) [3]. In particular, the absence of diphosphatidylglycerol and phosphatidylglycerol in strain MKL-02^T clearly differentiated the strain from the closely related *Arsenicicoccus* type strains, *A. bolidensis* DSM 15745^T and *A. dermatophilus* DSM 25571^T (Fig. S2). The isoprenoid quinone analysis showed that strain MKL-02^T contained MK-8(H_4) as the predominant menaquinone, which was similar to the other reference strains of the genus *Arsenicicoccus* [1–3].

Based on the results of the phylogenetic analysis using 16S rRNA gene and whole-genome sequences, strain MKL-02^T is closely related to *A. bolidensis* DSM 15745^T. In addition to the 16S rRNA gene sequence similarities and genomic sequence similarities, including DDH and ANI, strain MKL-02^T is distinct from the type strains belonging to the genus *Arsenicicoccus* in morphology, cellular fatty acid composition, polar lipids, and some distinguishable phenotypic and physiological characteristics, including conditions for growth, enzyme activities, and assimilation abilities. Therefore, strain MKL-02^T can be considered a novel species belonging to the genus *Arsenicicoccus*, for which the name *Arsenicicoccus cauae* sp. nov. is proposed.

DESCRIPTION OF *ARSENICICOCCUS CAUAE* SP. NOV.

Arsenicicoccus cauae (cau'ae. N.L. gen. n. *cauae* of CAU referring to Chung-Ang University or Chung-Ang University Hospital, Seoul, Republic of Korea, where the type strain was isolated).

Cells are Gram-stain-positive, facultative anaerobic, non-spore forming, and non-motile cocci (0.50–0.60 µm). Colonies on TSA with 10% sheep blood agar after incubation for 72 h at 37 °C are circular, cream yellow, convex, and 1.3 mm in diameter. In TSB with 10% sheep blood, growth is observed at a temperature of 10–45 °C (optimal temperature 37–40 °C) in the presence of 0–10% (w/v) NaCl (optimal concentration 0.5%) and at a pH of 6.0–10.0 (optimal pH 7.0). The strain is positive for catalase activity and negative for oxidase activity. Nitrate reduction is detectable but not indole production and glucose fermentation. Moreover, aesculin hydrolysis is detectable but not the hydrolysis of tyrosine, Tween 20, Tween 80, casein and starch. Alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, acid phosphatase and naphthol-AS-BI-phosphohydrolase, arginine dihydrolase and urease activities are present. However, lipase (C14), cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase, and α-fucosidase activities are absent. It can assimilate D-glucose, D-mannitol, maltose and malic acid. However, it cannot assimilate L-arabinose, D-mannose, *N*-acetyl-glucosamine, potassium gluconate, capric acid, adipic acid, trisodium citrate and phenylacetic acid. The major fatty acid type (>10% of total) is $C_{15:0}$. The polar lipid profile consists of two unidentified

phospholipids, three unidentified lipids and an unidentified aminophospholipid. The strain contains MK-8 (H_4) as the predominant menaquinone.

The type strain is MKL-02^T. It has been isolated from the blood of paediatric gastrointeritis patient in the Republic of Korea. The genomic DNA of this strain has 72.7 mol% G+C content.

Funding information
This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1A5A1018052) (M.-K.L.). This research was also supported by the Chung-Ang University Research Grants in 2021 (O.J.K.). The funders did not have any additional role in the study design, data acquisition and analysis, decision to publish, or preparation of the manuscript.

Author contributions
O.J.K.: data curation, formal analysis, investigation, methodology, and writing (original draft). K.H.K.: data curation, investigation, methodology, visualization, and writing (original draft). H.M.K.: data curation, investigation, methodology, software, and visualization. S.A.K.: data curation, investigation, methodology, software, and visualization. J.H.J. and Y.K.L.: investigation, software, and visualization. H.R.K.: investigation and methodology. C.O.J.: conceptualization, writing (review and editing), project administration, supervision, and validation. M.-K.L.: conceptualization, writing (review and editing), funding acquisition, project administration, supervision, and validation.

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

Ethical statement
According to the Chung-Ang University Hospital Institutional Review Board (IRB) policy, IRB review of the study and the need for obtaining informed consent from the patient for the publication were waived. (IRB No. 2011-007-19340).

References

1. **Collins MD, Routh J, Saraswathy A, Lawson PA, Schumann P, *et al.*** *Arsenicicoccus bolidensis* gen. nov., sp. nov., a novel actinomycete isolated from contaminated lake sediment. *Int J Syst Evol Microbiol* 2004;54:605–608.
2. **Gobeli S, Thomann A, Wyss F, Kuehni-Boghenbor K, Brodard I, *et al.*** *Arsenicicoccus dermatophilus* sp. nov., a hypha-forming bacterium isolated from the skin of greater flamingos (*Phoenicopterus roseus*) with pododermatitis. *Int J Syst Evol Microbiol* 2013;63:4046–4051.
3. **Hamada M, Iino T, Iwami T, Tamura T, Harayama S, *et al.*** *Arsenicicoccus piscis* sp. nov., a mesophilic actinobacterium isolated from the intestinal tract of a fish. *Actinomycetologica* 2009;23:40–45.
4. **Routh J, Saraswathy A, Collins MD.** *Arsenicicoccus bolidensis* a novel arsenic reducing actinomycete in contaminated sediments near the Adak mine (northern Sweden): impact on water chemistry. *Sci Total Environ* 2007;379:216–225.
5. **Nouioui I, Carro L, García-López M, Meier-Kolthoff JP, Woyke T, *et al.*** Genome-based taxonomic classification of the phylum *Actinobacteria*. *Front Microbiol* 2018;9:2007.
6. **CLSI.** *Interpretive Criteria for Identification of Bacteria and Fungi by Targeted DNA Sequencing*. 2nd edn. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.
7. **Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, *et al.*** Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.
8. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** Basic Local Alignment Search Tool. *J Mol Biol* 1990;215:403–410.
9. **Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW.** GenBank. *Nucleic Acids Res* 2016;44:D67–72.
10. **Tamura K, Stecher G, Kumar S.** MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Mol Biol Evol* 2021;38:3022–3027.
11. **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.
12. **Lee I, Chalita M, Ha S-M, Na S-I, Yoon S-H, *et al.*** ContEst16S: an algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *Int J Syst Evol Microbiol* 2017;67:2053–2057.
13. **Kim J, Na S-I, Kim D, Chun J.** UBCG2: Up-to-date bacterial core genes and pipeline for phylogenomic analysis. *J Microbiol* 2021;59:609–615.
14. **Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J.** A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* 2017;110:1281–1286.
15. **Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M.** Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
16. **Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, *et al.*** Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461–466.
17. **Benjamin WH, Yother J, Hall P, Briles DE.** The *Salmonella typhimurium* locus mviA regulates virulence in Itys but not Ityr mice: functional mviA results in avirulence; mutant (nonfunctional) mviA results in virulence. *J Exp Med* 1991;174:1073–1083.
18. **Bustamante P, Vidal R.** Repertoire and diversity of toxin - antitoxin systems of crohn's disease-associated adherent-invasive *Escherichia coli*. new insight of this emergent *E. coli* pathotype. *Front Microbiol* 2020;11:807.
19. **Fernandez RC, Weiss AA.** Cloning and sequencing of a *Bordetella pertussis* serum resistance locus. *Infect Immun* 1994;62:4727–4738.
20. **Gomori G.** *Preparation of Buffers for Use in Enzyme Studiesbuffers for use in enzyme studies. Handbook of Biochemistry and Molecular Biology*. CRC Press, 2010, pp. 721–724.
21. **Gerhardt P.** *Methods for General and Molecular Bacteriology*. Washington, D.C: American Society for Microbiology; 1994.
22. **MacFaddin JF.** *Biochemical Tests for Identification of Medical Bacteritests for identification of medical bacteria*. Williams & Wilkins Co, 1976.
23. **Sasser M.** MIDI technical note 101. Identification of bacteria by gas chromatography of cellular fatty acids. Newark, DE: MIDI, 1990, pp. 1–7.
24. **Minnikin DE, Patel PV, Alshamaony L, Goodfellow M.** Polar lipid composition in the classification of nocardia and related bacteria. *Int J Syst Bacteriol* 1977;27:104–117.
25. **Minnikin DE, O'Donnell AG, Goodfellow M, Alderson G, Athalye M, *et al.*** An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *J Microbiol Methods* 1984;2:233–241.