



A Novel Species of the Genus *Arsenicococcus* Isolated From Human Blood Using Whole-Genome Sequencing

Ji Hun Jeong , M.D., Ph.D.¹, Oh Joo Kweon , M.D.², Hye Ryoum Kim , M.D., Ph.D.², Tae-Hyoung Kim , M.D., Ph.D.³, Sung-min Ha , Ph.D.⁴, and Mi-Kyung Lee , M.D., Ph.D.²

¹Department of Laboratory Medicine, Daejeon Eulji Medical Center, Eulji University, Daejeon, Korea; ²Department of Laboratory Medicine, Chung-Ang University College of Medicine, Seoul, Korea; ³Department of Urology, Chung-Ang University College of Medicine, Seoul, Korea; ⁴Department of Integrative Biology and Physiology, University of California, Los Angeles, USA

Whole-genome sequencing (WGS) is an easily accessible and valuable tool in clinical microbiology, which can be used for identifying novel and rare species. We isolated gram-positive cocci from the blood of a pediatric patient, which could not be phenotypically identified using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) (BioMérieux, Marcy-l'Étoile, France). We could not identify the isolate to the species level using 16S ribosomal RNA (rRNA) sequencing. WGS was performed using the Illumina MiSeq platform (Illumina, San Diego, CA, USA); however, the subsequent genomic sequence database search using the TrueBac ID-Genome system (ChunLab, Inc., Seoul, Korea) did not yield any hits with an average nucleotide identity value >95.0%, which is the cut-off for species-level identification. Phylogenetic analysis suggested that the isolate belonged to a new *Arsenicococcus* species, forming a subcluster with *Arsenicococcus bolidensis*. Our data demonstrate that WGS allows a more accurate annotation of microbial genomes than other clinical microbiology tools, such as MALDI-TOF MS and 16S rRNA sequencing. This is the first report of the isolation of a novel *Arsenicococcus* species from a clinical sample.

Key Words: Novel *Arsenicococcus* species, 16S rRNA sequencing, Whole-genome sequencing

Received: April 29, 2020

Revision received: June 9, 2020

Accepted: November 26, 2020

Corresponding author:

Mi-Kyung Lee, M.D., Ph.D.
Department of Laboratory Medicine,
Chung-Ang University College of Medicine,
102 Heukseok-ro, Dongjak-gu, Seoul
06973, Korea.
Tel: +82-2-6299-2719
Fax: +82-2-6298-8630.
E-mail: cpworld@cau.ac.kr



© Korean Society for Laboratory Medicine

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

The genus *Arsenicococcus* comprises gram-positive, facultatively anaerobic, catalase-positive, and non-spore-forming bacteria that have been isolated from environmental sources and animal specimens [1-3]. *Arsenicococcus* was first proposed in 2004 by Collins, *et al.* [1], who described *Arsenicococcus bolidensis* isolated from mine sediment [4]. *Arsenicococcus* is a member of the family *Intrasporangiaceae* and initially comprised a single species, *A. bolidensis*. Subsequently, two additional *Arsenicococcus* species have been reported, *A. piscis* and *A. dermatophilus*, from the intestinal tract of the Japanese fish (*Sillago japonica*) and flamingo foot skin, respectively [2, 3].

We investigated a case of bloodstream infection caused by gram-positive cocci that were not accurately identified by rou-

tine testing. The isolate was identified as belonging to a novel *Arsenicococcus* species using whole-genome sequencing (WGS). A database search and phylogenetic tree analysis were conducted using the whole-genome sequence, and the identified features of the novel species, including its phenotype and antimicrobial susceptibility, are described. The need for informed consent for conducting this study was waived by the Institutional Review Board of Chung-Ang University Hospital, Seoul, Korea (No. 2011-007-19340).

A 17-month-old boy presenting with fever, nonprojectile vomiting, cyclic abdominal pain, and decreased urination was admitted to the emergency department (ED) of Chung-Ang University Hospital, Seoul, Korea, in May 2019. At admission, the

patient's temperature was 37.6°C, and blood analysis indicated a white blood cell count of $10.46 \times 10^9/L$, with a 70.0% of neutrophils. The C-reactive protein serum concentration was normal (<0.1 mg/L), and the patient was discharged from the ED after 5-6 hours following clinical improvement after hydration therapy. However, the patient's gastrointestinal symptoms aggravated soon thereafter; he presented with persistent vomiting and new symptoms, including diarrhea. In the evening, the patient was readmitted to the ED and diagnosed as having gastroenteritis. Blood analysis showed no significant changes compared to previous ones. Microbial analyses of stool specimens, including PCR analysis for detecting viral pathogens (rotavirus, norovirus, enteric adenovirus, and astrovirus) and a stool culture for detecting bacterial pathogens (*Salmonella*, *Shigella*, and *Vibrio* sp.), showed negative results. Pathogenic microorganisms were also absent in urine specimens; however, bacteria were detected in the patient's blood. After treatment with cefotaxime for three days, the patient was discharged from the hospital following clinical improvement.

An aerobic blood culture was performed using a pediatric blood culture bottle and the BacTAlert 3D blood culture system (BioMérieux, Inc., Durham, NC, USA); positive signals were detected (BacT/Alert PF Plus; BioMérieux, Inc., Marcy-l'Étoile, France) after approximately 48 hours of incubation at 35.0°C. Subculture on blood and chocolate agar plates were also prepared. After 24 hours, colonies on both media showed a slow

growth rate, with a dry, pale yellow, pinpoint morphology (Fig. 1). These isolates were gram-positive clusters of cocci and presented delayed catalase- and coagulase-positive results. The Vitek 2 system (BioMérieux, Marcy-l'Étoile, France) identified the isolate as *Facklamia hominis*, with an identification probability of 99%, whereas the Vitek MS system (BioMérieux) could not identify the isolated species even though *F. hominis* is included in the Vitek MS database (Version 3.2). *Facklamia* species are gram-positive, non-spore-forming, α -hemolytic, catalase-negative, and facultative anaerobic cocci. Because the Vitek 2 system identification result did not match the known phenotypic characteristics of the identified species, the isolate was further analyzed

Table 1. Whole-genome sequence assembly results (based on the TrueBac ID-Genome system) used to identify the novel *Arsenicococcus* species

Species (taxon hit)	ANI (%)	ANI coverage (%)	16S rRNA gene identity (%)*
<i>Arsenicococcus bolidensis</i>	90.05	76.3	99.79
<i>Arsenicococcus dermatophilus</i>	NA	NA	98.16
<i>Arsenicococcus piscis</i>	NA	NA	97.43
<i>Ornithinmicrobium murale</i>	83.47	3.7	96.59
<i>Ornithinmicrobium pekingense</i>	83.94	7.3	95.56

*Obtained from whole-genome sequences.

Abbreviations: ANI, average nucleotide identity; NA, not available; rRNA, ribosomal RNA.

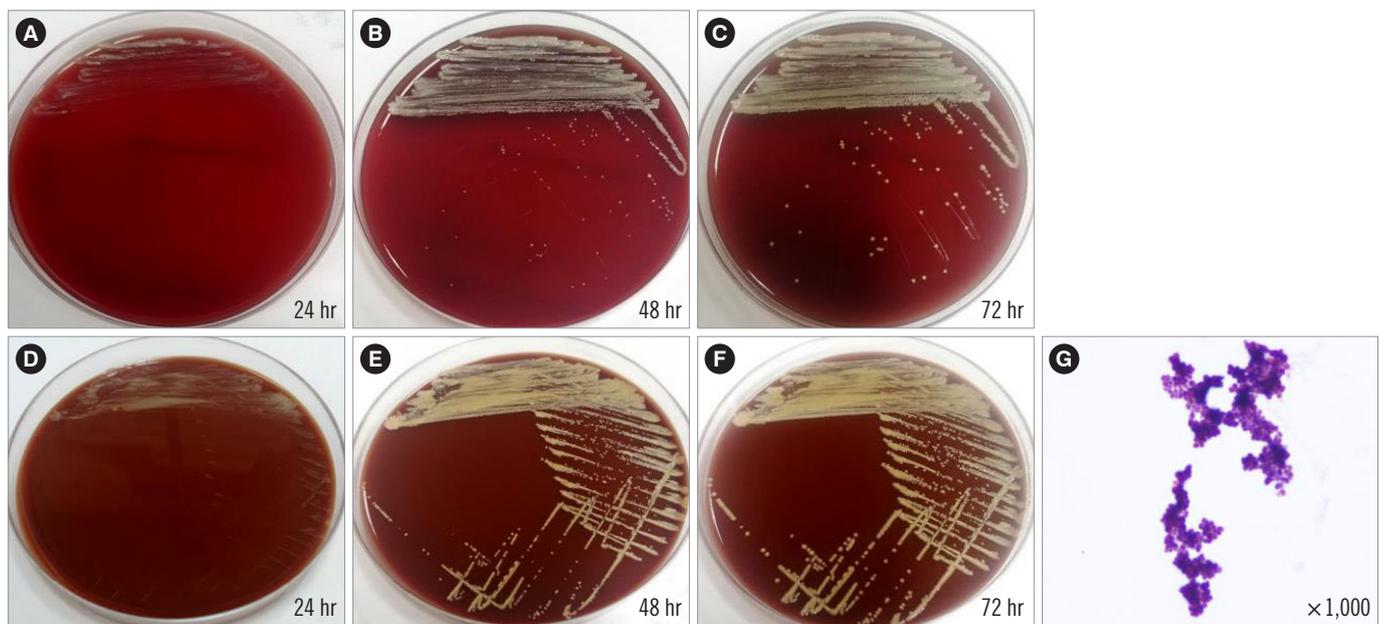


Fig. 1. Subculturing of the isolate on blood (A, B, C) and chocolate (D, E, F) agar plates. G. Gram staining of the isolate.

using 16S ribosomal RNA (rRNA) gene sequencing (1,446 bp) with universal primers (27F and 1492R) for species-level identification, according to the CLSI document MM18-A [5].

Sequencing of the 16S rRNA gene in conjunction with a GenBank Basic Local Alignment Search Tool (BLAST) search identified the pathogen as *A. bolidensis* (strain y63), with the highest sequence identity of 99.79% (1,440/1,446 bp); *Dermabacter* sp. (99.1%; 1,435/1,448 bp), *A. piscis* (97.38%; 1,411/1,449 bp), and *A. dermatophilus* (98.38%; 1,339/1,361 bp; partial iden-

tity) had lower sequence identities. *A. bolidensis* (strain CCUG 47306) showed only partial sequence identity (1,435/1,438 bp). For proper species identification, at least a 99.0% 16S rRNA sequence identity is required, with a >0.8% separation between species. Consequently, this isolate could not be categorized at the species level using 16S rRNA sequencing. As the matches obtained in our database search were not common strains, additional analysis was required to accurately identify the genus of the isolate (*Arsenicococcus* or *Dermabacter*). The 16S rRNA se-

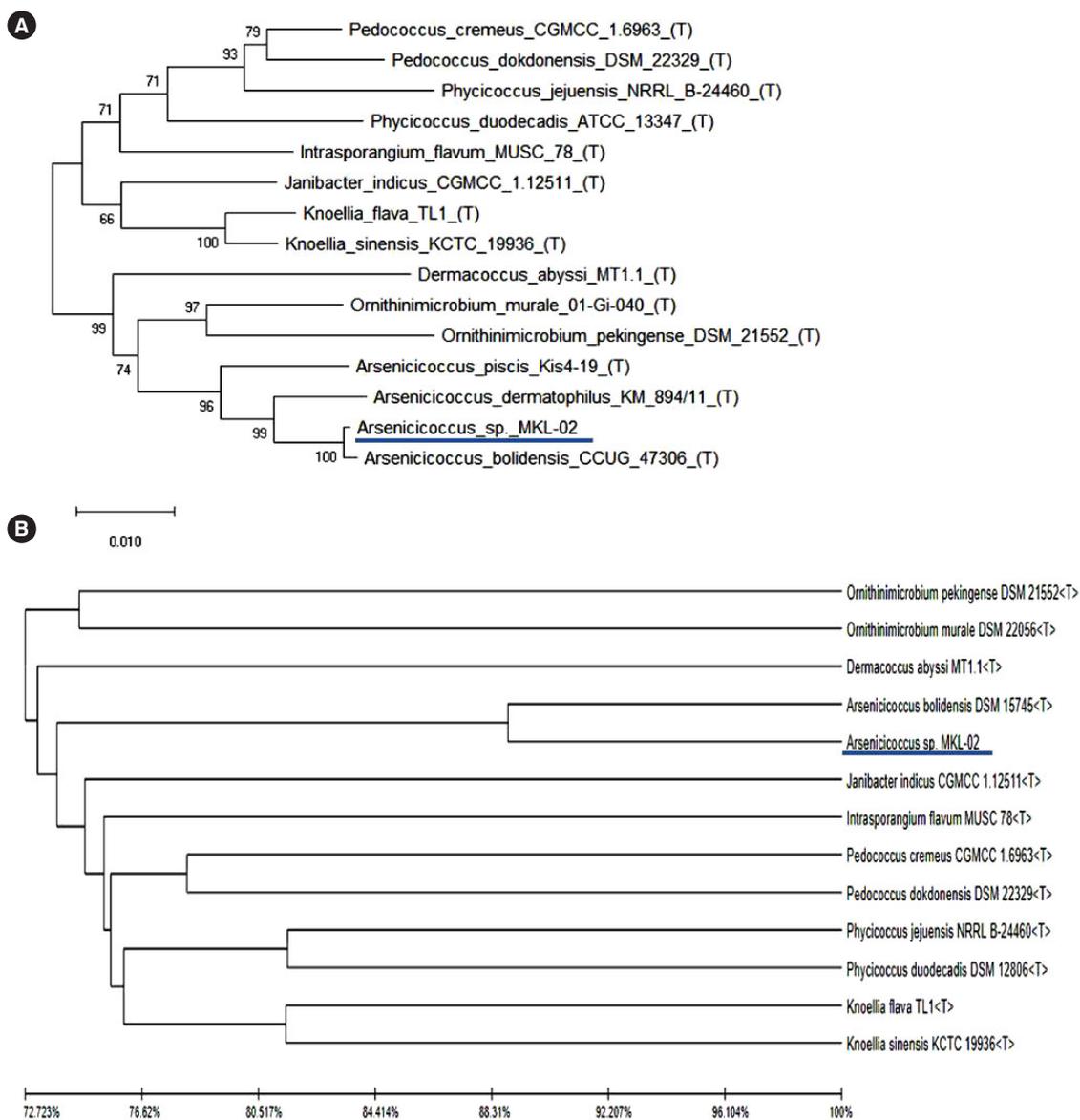


Fig. 2. Phylogenetic analysis based on (A) 16S rRNA sequences using the neighbor-joining method or (B) whole-genome sequencing combined with unweighted pair group and arithmetic mean clustering methods as well as average nucleotide identity values. The isolated pathogen (MKL-02) is underlined.

Abbreviation: rRNA, ribosomal RNA.

quence of the isolate was submitted to the GenBank database (Accession number MN629181).

For a more accurate identification of the isolate, WGS using the Illumina MiSeq platform (Illumina, San Diego, CA, USA) and data analysis using the TrueBac ID-Genome system (www.truebacid.com; ChunLab, Inc., Seoul, Korea) were performed [6]. The genome size of the pathogen obtained by WGS was 3,423,857 bp. We further determined a GC content of 71.98%, with an average sequencing depth of 281.8 \times , a total of 62 contigs, and an N50 value of 160,716 bp. The TrueBac ID database search showed that *A. bolidensis* had the highest 16S rRNA sequence identity (99.79%), exceeding the cut-off value (98.7%); however, we did not detect any pathogen with an average nucleotide identity (ANI) value >95.0%, which is the algorithmic cut-off for species-level identification. *A. bolidensis* showed the highest ANI value, 90.05% (ANI coverage of 76.3%). Details of the database search for identification of the isolated pathogen are presented in Table 1. The whole-genome sequence of the isolated bacterium was submitted to the BioProject database (Accession number PRJNA581338; Table 1).

A phylogenetic tree analysis using the neighbor-joining method with 16S rRNA gene sequences revealed that the newly isolated species (MKL-02) formed a subcluster with *A. bolidensis* CCUG 47306^T, with 98% bootstrap support (Fig. 2A). A similar subcluster was determined by phylogenetic tree analysis using WGS combined with unweighted pair group and arithmetic mean clustering methods as well as ANI values. The pathogen of interest (MKL-02) was most closely related to *A. bolidensis* DSM 15745^T; however, it was found to be phylogenetically distinct from *A. bolidensis* (Fig. 2B).

Antimicrobial susceptibility was tested using the Vitek 2 system, AST-ST01. As the Vitek 2 system identified the isolate as *F. hominis*, the antimicrobial susceptibility was interpreted according to the CLSI M100 guidelines and interpretive criteria for *Streptococcus pneumoniae* isolates [7]. The Vitek 2 system data indicated that the isolate is resistant to cefotaxime (≥ 8 $\mu\text{g/mL}$) and ceftriaxone (4 $\mu\text{g/mL}$). In addition, an antimicrobial susceptibility test was performed by the disk diffusion method, and the results were compared with the minimum inhibitory concentration (MIC) results. In the disk diffusion analysis, the zones of inhibition for cefotaxime and ceftriaxone were 34 and 35 mm, respectively. As there are no cefotaxime and ceftriaxone disk diffusion standard values for *S. pneumoniae*, we applied the CLSI M100 guidelines for *Streptococcus* species (viridans group or beta-hemolytic group), which defined this isolate as susceptible to both antibiotics.

As no susceptibility criteria (MIC or disk diffusion values) are available for *Arsenicococcus* species, we decided to define the isolate as clinically sensitive to cefotaxime based on the fact that the patient showed an improved clinical course after cefotaxime treatment.

Our contradictory results can be explained by the “90-60 rule.” While 90% of the susceptibility results are expected to predict success, only 60% of the resistance results are expected to do the same [8].

In clinical microbiology, automated phenotypical identification systems and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) have been widely adopted for routine bacterial identification. However, discrepancies among methods are frequently encountered, even if the identification score is high enough to identify organisms at the species level. This problem may originate from overlapping biochemical or proteomic profiles among different strains, the limited number of databases/libraries in the analyzer, or both. Likewise, 16S rRNA gene sequence similarities above cut-off levels (98.7–99.0%) do not guarantee the correct identification of certain strains, as nearly identical 16S rRNA gene sequences have been reported in different species [6, 9]. These species identification difficulties encountered using molecular or phenotypic typing can occur more frequently for strains rarely found in clinical specimens, as in the present case. In such cases, genome-based identification can be useful. Unlike 16S rRNA gene sequencing, WGS provides clear-cut criteria for bacterial classification [10, 11]. The ANI value is most widely used for analyzing genomic sequence similarity and is thus regarded as a possible next-generation gold standard for species delineation. It represents the mean of the similarity values between homologous genomic regions shared by two genes. It is now generally accepted that ANI values of 95–96% equate to a DNA-DNA hybridization value of 70%, which is the accepted standard for bacterial species identification [12].

In conclusion, we identified a clinical isolate to be a gram-positive coccus belonging to a novel *Arsenicococcus* species (MKL-02), which is most closely related to *A. bolidensis*. Genome-based methods, such as WGS, allow accurate bacterial identification in cases of discrepancies between more commonly used methods of bacterial identification, such as phenotyping, 16S rRNA sequence analysis, and MALDI-TOF MS. Genome-based identification can serve as a useful tool in clinical microbiology as the cost of WGS continues to decrease and database coverage expands.

ACKNOWLEDGEMENTS

None.

AUTHOR CONTRIBUTIONS

Research conception & design: Lee MK. Data acquisition: Jeong JH, Kweon OJ, Kim HR, Kim TH, and Lee MK. Data analysis and interpretation: Jeong JH, Kweon OJ, and Lee MK. Figures: Ha SM. Drafting of the manuscript: Jeong JH. Critical revision of the manuscript: Lee MK. Approval of the final manuscript: all authors.

CONFLICTS OF INTEREST

None declared.

RESEARCH FUNDING

None declared.

ORCID

Ji Hun Jeong	https://orcid.org/0000-0002-5586-7889
Oh Joo Kweon	https://orcid.org/0000-0003-4751-7384
Hye Ryouon Kim	https://orcid.org/0000-0002-9229-9665
Tae-Hyoung Kim	https://orcid.org/0000-0002-0257-3449
Sung-min Ha	https://orcid.org/0000-0003-3945-8329
Mi-Kyung Lee	https://orcid.org/0000-0003-1824-476X

REFERENCES

- Collins MD, Routh J, Saraswathy A, Lawson PA, Schumann P, Welinder-Olsson C, et al. *Arsenicococcus bolidensis* gen. nov., sp. nov., a novel actinomycete isolated from contaminated lake sediment. *Int J Syst Evol Microbiol* 2004;54:605-8.
- Hamada M, Iino T, Iwami T, Tamura T, Harayama S, Suzuki KI. *Arsenicococcus piscis* sp. nov., a mesophilic actinobacterium isolated from the intestinal tract of a fish. *Actinomycetologica* 2009;23:40-5.
- Gobeli S, Thomann A, Wyss F, Kuehni-Boghenbor K, Brodard I, Perreten V. *Arsenicococcus dermatophilus* sp. nov., a hypha-forming bacterium isolated from the skin of greater flamingos (*Phoenicopterus roseus*) with pododermatitis. *Int J Syst Evol Microbiol* 2013;63:4046-51.
- Routh J, Saraswathy A, Collins MD. *Arsenicococcus bolidensis* a novel arsenic reducing actinomycete in contaminated sediments near the Adak mine (northern Sweden): impact on water chemistry. *Sci Total Environ* 2007;379:216-25.
- CLSI. Interpretive criteria for identification of bacteria and fungi by DNA target sequencing; approved guideline. CLSI MM18-AE. Wayne, PA: Clinical and Laboratory Standards Institute. 2008.
- Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, et al. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461-6.
- CLSI. Performance standards for antimicrobial susceptibility testing. 29th ed. CLSI supplement M100-S29. Wayne, PA: Clinical and Laboratory Standards Institute. 2019.
- Doern GV and Brecher SM. The clinical predictive value (or lack thereof) of the results of in vitro antimicrobial susceptibility tests. *Journal of Clinical Microbiology* 2011;49:S11-S4.
- Yoon SH, Ha SM, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110:1281-6.
- Ha SM, Kim CK, Roh J, Byun JH, Yang SJ, Choi SB, et al. Application of the whole genome-based bacterial identification system, TrueBac ID, using clinical isolates that were not identified with three matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) systems. *Ann Lab Med* 2019;39:530-6.
- Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346-51.
- Richter M and Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 2009;106:19126-31.