



# Complete Genome Sequence of *Alphaproteobacteria* Bacterium Strain SO-S41, Isolated from Forest Soil

Motoyuki Watanabe,<sup>a,b</sup> Kensuke Igarashi,<sup>b</sup> Souichiro Kato,<sup>a,b</sup> Yoichi Kamagata,<sup>a,c</sup>  Wataru Kitagawa<sup>a,b</sup>

<sup>a</sup>Graduate School of Agriculture, Hokkaido University, Sapporo, Hokkaido, Japan

<sup>b</sup>Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo, Hokkaido, Japan

<sup>c</sup>Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan

**ABSTRACT** The complete genome of hydrogen peroxide-sensitive alphaproteobacterial strain SO-S41 was sequenced. The complete genome contains a single chromosome, is 4,443,179 bp in length, contains a total of 4,632 genes, and has a G+C content of 66.2%.

The high hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) level generated during agar medium preparation decreases the total colony count of inoculated environmental microorganisms (1–3). The strain SO-S41, which is highly sensitive to H<sub>2</sub>O<sub>2</sub>, was exclusively cultivated on an agar medium with a diminished amount of H<sub>2</sub>O<sub>2</sub>. Even if the medium is used, this slow-growing bacterium requires more than 1 week to form observable tiny colonies (1). This strain was isolated from the soil in the forest belonging to Hokkaido University, Sapporo, Hokkaido, Japan, and our previous study indicated that it might be phylogenetically novel at the order level (1). Here, the complete genome sequence of the strain SO-S41 is reported in order to obtain further understanding of its physiology.

The strain SO-S41 was grown in peptone-yeast-glucose (PYG) liquid medium (2) for 2 weeks at 20°C, and then its genomic DNA was extracted using the phenol-chloroform method (4). Short-read sequencing was performed using an iSeq 100 platform with paired-end 150-bp reads, and the sequencing library was prepared with a Nextera DNA Flex library prep kit (Illumina). The reads were trimmed with Trimmomatic v0.39 (SLIDINGWINDOW:6:30 MINLEN:78, and other parameters at the defaults) (5). In total, 923,927 paired-end and 167,116 single-end reads were obtained.

Long-read sequencing was performed using a MinION Mk1B platform, and the sequencing library was prepared with an SQK-RAD004 kit (Oxford Nanopore Technologies). The reads were base-called using the high-accuracy mode of MinKNOW v4.1.22 and then trimmed with Nanofilt v2.7.1 (-q 10, -l 500, -h 75, and other parameters at the defaults) (6), and the quality was assessed with NanoStat v1.5.0 (6). In total, 20,906 reads (*N*<sub>50</sub> read length of 1,350 bases, total read length of 25,855,567 bases) were obtained.

The hybrid assembly of iSeq and MinION reads was performed using SPAdes v3.14.1. (-k 21,33,41,55,77, --nanopore, --isolate, and other parameters at the defaults) (7, 8). Both paired-end and single-end reads from the iSeq sequencing were utilized in the hybrid assembly. As a result, three contigs which are longer than 200 nucleotides (nt) were generated (lengths of 3,781,019 nt, 600,517 nt, and 55,596 nt; coverages of 30×, 29×, and 29×). Outward-directing primers were designed from the ends of contigs and used in the PCR, using genomic DNA as the template. Amplification confirmed that three contigs are parts of the single circular chromosome. The region between contigs was determined by sequencing amplicons using an Applied Biosystems 3500xL genetic analyzer (Thermo Fisher Scientific). The assembled circular chromosome was annotated using Prokka v1.14.6 (9) with default parameters. The genome was rotated manually to the start of the *dnaA* gene.

The genome consists of a single chromosome of 4,443,179 bp with a 66.2% G+C content. The chromosome contains a total of 4,362 genes, 4,304 protein-coding genes, 53

**Citation** Watanabe M, Igarashi K, Kato S, Kamagata Y, Kitagawa W. 2021. Complete genome sequence of *Alphaproteobacteria* bacterium strain SO-S41, isolated from forest soil. *Microbiol Resour Announc* 10:e00536-21. <https://doi.org/10.1128/MRA.00536-21>.

**Editor** Frank J. Stewart, Montana State University

**Copyright** © 2021 Watanabe et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Wataru Kitagawa, [w-kitagawa@aist.go.jp](mailto:w-kitagawa@aist.go.jp).

**Received** 25 May 2021

**Accepted** 29 June 2021

**Published** 22 July 2021

tRNA genes, 4 rRNA genes, and 1 transfer-messenger RNA (tmRNA) gene. The genome contains major H<sub>2</sub>O<sub>2</sub> scavenging enzyme genes such as two catalase genes and two alkyl hydroperoxide reductase genes.

**Data availability.** This whole-genome sequence has been deposited in DDBJ under the accession no. [AP024629](https://doi.org/10.1128/AEM.00807-18). The raw sequence reads have been deposited in the DRA under accession no. [DRA011975](https://doi.org/10.1128/AEM.02741-14).

## ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI grant 18H05295.

## REFERENCES

1. Kato S, Yamagishi A, Daimon S, Kawasaki K, Tamaki H, Kitagawa W, Abe A, Tanaka M, Sone T, Asano K, Kamagata Y. 2018. Isolation of previously uncultured slow-growing bacteria by using a simple modification in the preparation of agar media. *Appl Environ Microbiol* 84:e00807-18. <https://doi.org/10.1128/AEM.00807-18>.
2. Tanaka T, Kawasaki K, Daimon S, Kitagawa W, Yamamoto K, Tamaki H, Tanaka M, Nakatsu CH, Kamagata Y. 2014. A hidden pitfall in agar media preparation undermines microorganism cultivability. *Appl Environ Microbiol* 80:7659–7666. <https://doi.org/10.1128/AEM.02741-14>.
3. Kawasaki K, Kamagata Y. 2017. Phosphate-catalyzed hydrogen peroxide formation from agar, gellan, and κ-carrageenan and recovery of microbial cultivability via catalase and pyruvate. *Appl Environ Microbiol* 83:e01366-17. <https://doi.org/10.1128/AEM.01366-17>.
4. Sambrook J, Russell RW. 2001. *Molecular cloning: a laboratory manual*, 3rd ed, vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
6. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
7. Prijbelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes de novo assembler. *Curr Protoc Bioinformatics* 70:e102. <https://doi.org/10.1002/cpbi.102>.
8. Antipov D, Korobeynikov A, McLean JS, Pevzner PA. 2016. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. *Bioinformatics* 32:1009–1015. <https://doi.org/10.1093/bioinformatics/btv688>.
9. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.