



Complete Genome Sequence of *Sphingomonas* sp. Strain NIBR02145

Hyun Young Kim,^{a,b} Jin Young Park,^c Jin Nam Kim,^d Jaewoong Yu,^d  Bongsang Kim^{d,e,f}

^aDepartment of Oral Microbiology and Immunology, School of Dentistry, Seoul National University, Seoul, Republic of Korea

^bDental Research Institute, School of Dentistry, Seoul National University, Seoul, Republic of Korea

^cBiological Resources Research Department, National Institute of Biological Resources, Incheon, Republic of Korea

^deGnome, Inc, Seoul, Republic of Korea

^eDepartment of Agricultural Biotechnology, Seoul National University, Seoul, Republic of Korea

^fResearch Institute of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea

ABSTRACT *Sphingomonas* sp. strain NIBR02145 is a putative chemoheterotrophic strain that was isolated from soil in Wando-gun, Republic of Korea. The NIBR02145 genome was sequenced with PacBio next-generation sequencing technology. The 5,010,245-bp circular genome has a GC content of 66.79% and harbors 4,561 coding sequences, 6 rRNAs, and 52 tRNAs.

The genus *Sphingomonas* was first proposed in 1990 by Yabuuchi et al. (1). The strains of the genus are Gram-negative, strictly aerobic, chemoheterotrophic bacteria (2, 3). *Sphingomonas* strains are found in a variety of habitats, including soil, water, plant roots, and even extreme environments (4–6). They are attracting great attention due to their properties that can be applied to biotechnology, such as biosynthetic abilities (7, 8) and biodegrading abilities (9). Therefore, whole-genome-level study of *Sphingomonas* strains is expected to make significant contributions not only to the possibility of industrial applications but also to ecological research. In addition, although there are 3,960 *Sphingomonas* taxonomies in NCBI, only 139 of them have genome data, which limits comparative and phylogenetic analyses of *Sphingomonas* strains.

A sample of topsoil was obtained at a depth of <10 cm in Wando-gun, Jeollanam-do, Republic of Korea (34°18'52.4"N, 126°42'50.5"E). Strain NIBR02145 was cultured in Reasoner's 2A (R2A) medium (Merck, Millipore, Germany) and incubated at 25°C for 3 days, and cells from a single colony were collected in a 5-mL Eppendorf tube for DNA extraction. Genomic DNA was extracted using a MagCore genomic DNA bacterial kit (RBC Bioscience, Taiwan), and a Covaris g-TUBE device was used to shear the DNA according to the instructions of the manufacturer. DNA was sheared into fragments of >15 kb and purified using 0.45× AMPure XP magnetic beads (Beckman Coulter, USA). Library construction was performed using the Pacific Biosciences (PacBio) SMRTbell library preparation kit v1.0 (PacBio). DNA damage repair and end repair were performed using the SMRTbell damage repair kit (PacBio), and then reagents from the SMRTbell enzyme cleanup kit (PacBio) were added to the library. The library was placed in 0.75% BluePippin and electrophoresed to obtain fragments of ≥15 kb, and then the library was sequenced on the PacBio RS II platform.

The number of sequenced raw reads was 1,747,074 (14,482,769,372 bp), and the N_{50} value was 9,716 bp. Reads of <1,000 bp were filtered using SeqKit v2.3.0, and then *de novo* assembly was conducted with Flye v2.8 (10). The assembly resulted in one circular complete genome of 5,010,245 bp, with a GC content of 66.79%. The quality of the genome assembly was assessed with Benchmarking Universal Single-Copy Orthologs (BUSCO) v5.2.2 with the bacteria_odb data set (11). The result showed that 123 of the 124 BUSCOs were complete single-copy BUSCOs. The genome was annotated using the NCBI Prokaryotic Genome

Editor J. Cameron Thrash, University of Southern California

Copyright © 2023 Kim 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 Jaewoong Yu, jwyu@egnome.co.kr, or Bongsang Kim, kimbongsang@egnome.co.kr.

The authors declare no conflict of interest.

Received 3 February 2023

Accepted 1 June 2023

TABLE 1 Statistical information for the genome sequence of *Sphingomonas* sp. strain NIBR02145

Parameter	Value
No. of PacBio reads (>1,000 bp)	1,712,399
Genome coverage (×)	2,755
Genome size (bp)	5,010,245
No. of contigs	1
GC content (%)	66.79
No. of coding sequences	4,561
No. of tRNAs	52
No. of rRNAs	6
No. of ncRNAs	3

Annotation Pipeline (PGAP) v6.5 (12). The genome of NIBR02145 contained 4,561 coding sequences, 6 rRNAs, 52 tRNAs, and 3 noncoding RNAs (ncRNAs). The statistical information for the genome sequence of NIBR02145 is presented in Table 1. The 16S rRNA gene of NIBR02145 was compared with the NCBI 16S rRNA gene database using BLASTn (13). In the BLASTn comparison, the 16S rRNA gene of *Sphingomonas kyeonggiensis* THG-DT81 (GenBank accession number [NR_134182.1](https://doi.org/10.1348-0421.1990.tb00996.x)) showed the greatest similarity (98.5%). In addition, average nucleotide identity (ANI) values were calculated with pyani v0.2.12 (14). The greatest ANI value was 87.3% for *S. kyeonggiensis* THG-DT81. Default parameters were used for all software unless otherwise specified. All of these results suggested that NIBR02145 is closely related to *Sphingomonas kyeonggiensis*.

Data availability. The complete genome sequence of *Sphingomonas* sp. strain NIBR02145 has been deposited in GenBank under the accession number [CP114791](https://doi.org/10.1111/j.1348-0421.1990.tb00996.x). The raw data have been deposited in the SRA under the accession number [SRR22805145](https://doi.org/10.1111/j.1348-0421.1990.tb00996.x).

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institute of Biological Resources (NIBR), funded by the Ministry of Environment (MOE) of the Republic of Korea (grant NIBR202231203).

REFERENCES

1. Yabuuchi E, Yano I, Oyaizu H, Hashimoto Y, Ezaki T, Yamamoto H. 1990. Proposals of *Sphingomonas paucimobilis* gen. nov. and comb. nov., *Sphingomonas parapaucimobilis* sp. nov., *Sphingomonas yanoikuyae* sp. nov., *Sphingomonas adhaesiva* sp. nov., *Sphingomonas capsulata* comb. nov., and two genospecies of the genus *Sphingomonas*. *Microbiol Immunol* 34:99–119. <https://doi.org/10.1111/j.1348-0421.1990.tb00996.x>.
2. Ryan MP, Adley CC. 2010. *Sphingomonas paucimobilis*: a persistent Gram-negative nosocomial infectious organism. *J Hosp Infect* 75:153–157. <https://doi.org/10.1016/j.jhin.2010.03.007>.
3. Lee J, Shin SC, Kim SJ, Kim B-K, Hong SG, Kim EH, Park H, Lee H. 2012. Draft Genome Sequence of a *Sphingomonas* sp., an Endosymbiotic Bacterium Isolated from an Arctic Lichen *Umbilicaria* sp. *J Bacteriol* 194:3010–3011. <https://doi.org/10.1128/JB.00360-12>.
4. Farias ME, Revale S, Mancini E, Ordonez O, Turjanski A, Cortez N, Vazquez MP. 2011. Genome sequence of *Sphingomonas* sp. S17, isolated from an alkaline, hyperarsenic, and hypersaline volcano-associated lake at high altitude in the Argentinean Puna. *J Bacteriol* 193:3686–3687. <https://doi.org/10.1128/JB.05225-11>.
5. Marizcurrena JJ, Morel MA, Brana V, Morales D, Martinez-Lopez W, Castro-Sowinski S. 2017. Searching for novel photolyases in UVC-resistant Antarctic bacteria. *Extremophiles* 21:409–418. <https://doi.org/10.1007/s00792-016-0914-y>.
6. Takeuchi M, Sakane T, Yanagi M, Yamasato K, Hamana K, Yokota A. 1995. Taxonomic study of bacteria isolated from plants: proposal of *Sphingomonas rosa* sp. nov., *Sphingomonas pruni* sp. nov., *Sphingomonas asaccharolytica* sp. nov., and *Sphingomonas mali* sp. nov. *Int J Syst Bacteriol* 45:334–341. <https://doi.org/10.1099/00207713-45-2-334>.
7. Sa-Correia I, Fialho AM, Videira P, Moreira LM, Marques AR, Albano H. 2002. Gellan gum biosynthesis in *Sphingomonas paucimobilis* ATCC 31461: genes, enzymes and exopolysaccharide production engineering. *J Ind Microbiol Biotechnol* 29:170–176. <https://doi.org/10.1038/sj.jim.7000266>.
8. Zhu L, Wu X, Li O, Qian C, Gao H. 2012. Cloning and characterization of genes involved in nostoxanthin biosynthesis of *Sphingomonas elodea* ATCC 31461. *PLoS One* 7:e35099. <https://doi.org/10.1371/journal.pone.0035099>.
9. Hesham AE-L, Mawad AMM, Mostafa YM, Shoreit A. 2014. Biodegradation ability and catabolic genes of petroleum-degrading *Sphingomonas koreensis* strain ASU-06 isolated from Egyptian oily soil. *Biomed Res Int* 2014:127674. <https://doi.org/10.1155/2014/127674>.
10. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
11. Manni M, Berkeley MR, Seppey M, Simao FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* 38:4647–4654. <https://doi.org/10.1093/molbev/msab199>.
12. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
13. McGinnis S, Madden TL. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res* 32:W20–W25. <https://doi.org/10.1093/nar/gkh435>.
14. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods* 8:12–24. <https://doi.org/10.1039/C5AY02550H>.