Draft Genome Sequence of Strain B 225, an Iron-Depositing Isolate of the Genus *Novosphingobium*

Burga Braun,[a] Sven Künzel,[b] Ulrich Szewzyk[a]

[a]Technische Universität Berlin, Berlin, Germany
[b]Max Planck Institute for Evolutionary Biology, Plön, Germany

ABSTRACT Here, we report the draft genome sequence of *Novosphingobium* sp. strain B 225, an iron-depositing bacterium isolated from a phenazone-amended naturally grown biofilm. This biofilm was grown in the Unteres Odertal National Park, Germany. Illumina NextSeq sequencing was used to determine the genome of the strain.

Wastewater effluent poses a potential risk to aquatic ecosystems due to persistent organic contaminants which are not eliminated during water treatment. Microbial activity plays a major role in the degradation of environmental pollutants, and members of the genus *Novosphingobium* have been previously described as contaminant-degrading organisms (1–3).

Based on 16S rRNA gene sequence similarity, strain B 225 belongs to the genus *Novosphingobium* and shows the highest level of similarity to two *Novosphingobium* strains, namely, *Novosphingobium hassiacum* strain W-51 (98%) and *Novosphingobium aromaticivorans* DSM 12444 (98%), based on BLASTn searches (4), and 97.6% similarity to *Novosphingobium ginsenosidimutans* by EzBioCloud (5) searches.

The Gram-negative rod-shaped *Novosphingobium* strain B 225 was isolated from a 4-month-old iron- and manganese-depositing biofilm grown in a continuous-flow biofilm reactor exposed in the Bogengraben River of the Unteres Odertal National Park, Germany. The reactor was continuously fed with river water and phenazone (0.5 g liter^{-1}). The biofilm was spread on ATA medium [2 g liter^{-1} MnCO$_3$ hydrate, 0.15 g ml^{-1} trisodium citrate, 0.2 g liter^{-1} Fe(NH$_4$)$_2$(SO$_4$)$_2$, 0.1 g liter^{-1} cycloheximide, 2 ml vitamin solution (6), 2 ml trace element solution SL 9 (7), 0.94 g phenazone, 20 g agar, and 1 liter filtered water from Lake Daminke] based on Mn agar (8) and incubated for 2 weeks at room temperature. Small dark-brown colonies were selected, and their iron and manganese deposition abilities were confirmed according to Schmidt et al. (9).

Genomic DNA was extracted using the GeneMATRIX soil DNA purification kit (Roboklon, Berlin, Germany). The paired-end library was prepared by using the TruSeq DNA high-throughput (HT) sample prep kit (Illumina Netherlands, Eindhoven, The Netherlands), and mate-pair libraries were established with the Nextera mate-pair sample preparation kit (Illumina Netherlands). The paired-end library was prepared following the Illumina Nextera XT DNA library prep kit protocol. Genome sequencing was done on an Illumina NextSeq 500 sequencer using a NextSeq mid-output kit version 2 with 300-cycle chemistry by generating 10,292,075 raw reads. Demultiplexing was done with bcl2fastq version 2.18.0.12, and quality filtering of raw reads was performed using Trimmomatic version 0.36 (10). The reads were then checked for ambiguous base calls and low complexity, employing the DUST algorithm (11), and filtered accordingly with an R script in Microsoft R Open version 3.3.2 (http://www.r-project.org/), followed by preassembly with SPAdes version 3.10.0 (12) using default k-mer lengths up to 99 bp. Scaffolds ≥500 bp of this preassembly were subject to extension and second-round scaffolding with the SSPACE standard version 3.0 (13).

Address correspondence to Burga Braun, Burga.braun@tu-berlin.de.

Scaffolds ≥2,500 bp were assigned to genome bins by MetaBAT version 0.32.4 (14), and functional annotation of the draft genomes was performed with Prokka version 1.12b (15).

The draft genome included 73 contigs, with an N_{50} assembly quality of 107,389 bp and an L_{50} value of 11 bp. The shortest sequence was 2,912 bp, and the longest sequence was 400,854 bp. The total size of the draft genome was 3,712,107 bp, with a GC content of 64%. Annotation included 3,500 coding regions for 3,561 genes, 463 signal peptides, 3 rRNAs (16S and 23S), 49 tRNAs, 1 transfer-messenger RNA (tmRNA), and 8 miscellaneous RNAs (miscRNAs).

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number MWLM00000000. The version described in this paper is version MWLM01000000.

ACKNOWLEDGMENTS

REFERENCES

1. Pieper C, Risse D, Schmidt B, Braun B, Szewzyk U, Rotard W. 2010. Investigation of the microbial degradation of phenazone-type drugs and their metabolites by natural biofilms derived from river water using liquid chromatography/tandem mass spectrometry (LC-MS/MS). Water Res 44:4559–4569. https://doi.org/10.1016/j.watres.2010.05.028.

2. Sohn JH, Kwon KK, Kang JH, Jung HB, Kim SJ. 2004. *Novosphingobium pentaromativorans* sp. nov., a high-molecular-mass polycyclic aromatic hydrocarbon-degrading bacterium isolated from estuarine sediment. Int J Syst Evol Microbiol 54:1483–1487. https://doi.org/10.1099/ijs.0.02945-0.

3. Yuan J, Lai Q, Zheng T, Shao Z. 2009. *Novosphingobium indicum* sp. nov., a polycyclic aromatic hydrocarbon-degrading bacterium isolated from a deep-sea environment. Int J Syst Evol Microbiol 59:2084–2088. https://doi.org/10.1099/ijs.0.002873-0.

4. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.

5. Yoon S, Ha S, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67:1613–1617. https://doi.org/10.1099/ijsem.0.001755.

6. Schlegel HG, Zaborosch C. 1992. Allgemeine Mikrobiologie Thieme Verlag, Stuttgart, Germany.

7. Tschech A, Pfennig N. 1984. Growth yield increase linked to caffeate reduction in *Acetobacterium woodii*. Arch Microbiol 137:163–167. https://doi.org/10.1007/BF00414460.

8. Mulder EG, Van Veen W. 1963. Investigations on the *Sphaerotilus-Leptothrix* group. Antonie Van Leeuwenhoek 29:121–153. https://doi.org/10.1007/BF02046045.

9. Schmidt B, Sánchez LA, Fretschner T, Kreps G, Ferrero MA, Siñeriz F, Szewzyk U. 2014. Isolation of *Sphaerotilus-Leptothrix* strains from iron bacteria communities in Tierra del Fuego wetlands. FEMS Microbiol Ecol 90:454–466. https://doi.org/10.1111/1574-6941.12406.

10. 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.

11. Morgulis A, Gertz EM, Schäffer AA, Agarwala R. 2006. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. J Comput Biol 13:1028–1040. https://doi.org/10.1089/cmb.2006.13.1028.

12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

13. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579. https://doi.org/10.1093/bioinformatics/btq683.

14. Kang DD, Froula J, Egan R. 2014. A robust statistical framework for reconstructing genomes from metagenomic data. bioRxiv https://doi.org/10.1101/011460.

15. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.