Complete Genome Sequence of *Pseudomonas putida* BS3701, a Promising Polycyclic Aromatic Hydrocarbon-Degrading Strain for Bioremediation Technologies

Andrey Filonov,[a,b,c] ![ORCID] Yanina Delegan,[a,b] Irina Puntus,[a,b,c] Leonid Valentovich,[d] Artur Akhremchuk,[d] Olesia Evdokimova,[d] Tatyana Funtikova,[a,b] Marina Zakharova,[a,b] Lenar Akhmetov,[a,b] Anna Vetrova,[a,b] Marina Titok[e]

[a]Institute of Biochemistry and Physiology of Microorganisms, Pushchino, Moscow Region, Russian Federation
[b]Federal Research Center Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences, Pushchino, Moscow Region, Russian Federation
[c]Tula State University, Tula, Russian Federation
[d]Institute of Microbiology of the National Academy of Sciences of Belarus, Minsk, Belarus
[e]Belarus State University, Minsk, Belarus

ABSTRACT The strain *Pseudomonas putida* BS3701 was isolated from soil contaminated with coke by-product waste (Moscow Region, Russian Federation). It is capable of degrading crude oil and polycyclic aromatic hydrocarbons (PAHs). The *P. putida* BS3701 genome consists of a 6,337,358-bp circular chromosome and two circular plasmids (pBS1141 with 107,388 bp and pBS1142 with 54,501 bp).

Pseudomonads are known for numerous strains able to degrade various organic compounds, including both aliphatic and aromatic hydrocarbons (1). The biodegradation of polycyclic aromatic hydrocarbons (PAHs) is often controlled by plasmids, most of which are found in representatives of the genus *Pseudomonas* (2). The strain *Pseudomonas putida* BS3701 (plasmids pBS1141 and pBS1142) was isolated from soil contaminated with coke by-product waste (sewage from the Moscow Coke and Gas Plant, Vidnoye, Moscow Region, Russian Federation) (3). It is capable of degrading crude oil and PAHs (strain certificate VKM B-2380D). For long-term storage, the strain was kept in glycerol (40%) stocks at −70°C. For short-term maintenance, the strain was cultured on LB agar plates at 24°C.

Genomic DNA was isolated from a fresh culture biomass (a single colony) of *Pseudomonas putida* BS3701 grown on LB agar using a DNeasy blood and tissue kit (catalog number 69506; Qiagen). Sequencing was performed using a MinION sequencer (Oxford Nanopore Technologies [ONT]) at the Center of Analytical and Genetic Engineering Research (Minsk, Belarus). The library was prepared with a ligation sequencing kit (catalog number SQK-LSK109). Guppy v. 3.6.0 software was used for base calling.

Additionally, the same DNA sample was sequenced with an Illumina MiSeq platform using a MiSeq reagent kit v. 2 (2×251 bp). A paired-end library for sequencing was prepared with the MuSeek library preparation kit (catalog number K1361; Thermo Fisher).

The Nanopore reads were binned and filtered by quality (Q score, >10) using Barapost v. 2020-06-26 (4), which yielded a total of 159.24 Mbp distributed in 17,240 reads with N_{50}/N_{90} values of 15,504/4,540 bp. Prepared reads were assembled into 3 contigs using Flye v. 2.7.1-b1590 (5). The contigs were checked for circularization using the Tablet software (6). The Illumina reads were assembled using SPAdes v. 3.14.1 (7). The Illumina reads were used to correct Nanopore errors using Bowtie 2 v. 2.3.5.1 (8) and Pilon v. 1.23 (9) software. Default parameters were used for all software.

The *P. putida* BS3701 genome consists of a 6,337,358-bp circular chromosome (GC content, 61.6%) and two circular plasmids. The first one, namely, pBS1141, has

107,388 bp and a GC content of 57.8%, and the second one, pBS1142, has 54,501 bp and a GC content of 57.4%. Chromosome and plasmid circularization was specified by ends overlapping.

The strain BS3701 genome was annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v. 4.6 (10). The chromosome contained 5,756 coding sequences, 7 rRNA clusters (5S, 16S, 23S), and 79 tRNAs. Previously (11), *Pseudomonas putida* BS3701 was shown to be capable of growth using naphthalene as the sole carbon and energy source. The genes of naphthalene degradation (naphthalene 1,2-dioxygenase large and small subunits, 2-hydroxymuconic semialdehyde hydrolase, catechol 2,3-dioxygenase, and salicylate 1-monooxygenase) were found on plasmid pBS1141. The mating experiment (12) indicated that pBS1142 is a helper plasmid, which could assist plasmid pBS1141 to transfer during conjugation.

The antiSMASH search for secondary metabolite clusters found 9 clusters on the main chromosome, including clusters of lipopolysaccharides and fengycin production (13). The genome sequence data of *Pseudomonas putida* BS3701 will enhance our understanding of the metabolic capabilities of *Pseudomonas* strains.

Data availability. This genome project has been deposited at GenBank under the accession numbers CP059052.1 for the chromosome, CP059053.1 for plasmid pBS1141, and CP059054.1 for plasmid pBS1142. Other data can be found under BioSample number SAMN15517094 and BioProject number PRJNA645807. The SRA accession numbers are SRX8724918 (Illumina) and SRX8724919 (Nanopore).

REFERENCES

1. Brzeszcz J, Kaszycki P. 2018. Aerobic bacteria degrading both n-alkanes and aromatic hydrocarbons: an undervalued strategy for metabolic diversity and flexibility. Biodegradation 29:359–407. https://doi.org/10.1007/s10532-018-9837-x.
2. Izmalkova T, Sazonova O, Nagornih M, Sokolov S, Kosheleva I, Boronin A. 2013. The organization of naphthalene degradation genes in Pseudomonas putida strain AK5. Res Microbiol 164:244–253. https://doi.org/10.1016/j.resmic.2012.12.007.
3. Ovchinnikova A, Vetrova A, Filonov A, Boronin A. 2009. Phenanthren biodegradation and interaction of *Pseudomonas putida* BS3701 and *Burkholderia* sp. BS3702 in plant rhizosphere. Mikrobiologiia 78:484–490. https://doi.org/10.1134/S0026261709040067.
4. Sikolenko M, Valentovich L. 2020. Barapost: binning of nucleotide sequences according to taxonomic annotation. IEEE/ACM Trans Comput Biol Bioinform https://doi.org/10.1109/TCBB.2020.3009780.
5. Kolmogorov M, Yuan J, Lin Y, Pevzner P. 2019. Assembly of long error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.
6. Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD, Marshall D. 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform 14:193–202. https://doi.org/10.1093/bib/bbs012.
7. Bankevich A, Nurk S, Antipov D, Gurevich A, Dvorkin M, Kulikov AS, Lesin V, Nikolenko S, Pham S, Prjibelski A, Pyshkin A, Sirotkin A, 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.
8. Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
9. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
10. 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.
11. Puntus I, Filonov A, Kosheleva I, Gaiazov R, Karpov A, Boronin A. 1997. Isolation and characterization of microorganisms degrading polycyclic aromatic hydrocarbons. Microbiology 66:222.
12. Balashova N, Kosheleva I, Golovchenko N, Boronin A. 1999. Phenanthrene metabolism by Pseudomonas and Burkholderia strains. Process Biochem 35:291–296. https://doi.org/10.1016/S0032-9592(99)00069-2.
13. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 47:W81–87. https://doi.org/10.1093/nar/gkz310.