Complete Genome Sequence of *Gordonia polyisoprenivorans* 135, a Promising Degrader of Aromatic Compounds

Ekaterina Frantsuzova,[a] Viktor Solomentsev,[a,b] Anna Vetrova,[a] Vasili Travkin,[c] Inna Solyanikova,[a,c] [ID] Yanina Delegan[a]

[a]Institute of Biochemistry and Physiology of Microorganisms, Federal Research Center "Pushchino Scientific Center for Biological Research of Russian Academy of Sciences" (FRC PSCBR RAS), Pushchino, Moscow Region, Russia
[b]State Research Center for Applied Microbiology and Biotechnology, Obolensk, Moscow Region, Russia
[c]Belgorod State University, Belgorod, Russian Federation

ABSTRACT *Gordonia polyisoprenivorans* 135 is a promising degrader of aromatic hydrocarbons. It can utilize phenanthrene, anthracene, benzoate, chlorobenzoates, and phenol. The genome of strain 135 was completely sequenced; it consists of a single 5,988,360-bp circular chromosome (GC content of 67.01%).

Polycyclic aromatic hydrocarbons (PAHs) are chemically stable carcinogenic and mutagenic compounds of crude oil (1). The ability to utilize PAHs is characteristic for many *Actinobacteria* members, but not many strains of the genus *Gordonia* are currently known to possess this property. A description of the genetic organization of the PAH catabolism pathway is presented in the work of Lin et al. (2).

Gordonia polyisoprenivorans strain 135 (previously *Rhodococcus rhodnii* 135) was isolated in 1998 from soil contaminated with oil, diesel fuel, and chlorophenols (latitude, 53°10′39′′N; longitude, 50°04′44′′E; Samara, Russia) using the enrichment culture method (3, 4) on Evans medium (5) with *p*-hydroxybenzoate as the sole source of carbon and energy. The strain is capable of utilizing phenanthrene, anthracene (6), benzoate (7), chlorobenzoates (8), and phenol (4). For long-term storage, the strain was kept in glycerol (40%) stocks at −70°C. For short-term maintenance, the strain was cultured on a Lysogeny broth agar plate at 27°C.

Genomic DNA was isolated from a fresh culture biomass (a colony) of *Gordonia polyisoprenivorans* 135 grown on LB agar using a DNeasy blood and tissue kit (Qiagen; 69506). Sequencing was performed using a MinION sequencer with flow cell FLO-MIN106 (Oxford Nanopore Technologies [ONT]). A library was prepared with a ligation kit (SQK-LSK109). Guppy 3.2.4 was used for base calling, which yielded a total of 1,335.8 Mb distributed in 332,510 reads with a Q of >10 (N_{50} is 14,599 bp).

Additionally, the same DNA sample was sequenced with an MGI platform (DNBSEQ-G400) using the DNBSEQ-G400RS high-throughput sequencing set (FCL PE150) (2 × 150 bp). A paired-end library was prepared with the MGIEasy universal DNA library prep set. We obtained 8,939,552 paired-end reads of <150 bp. The MGI and Nanopore reads were used for hybrid assembly with SPAdes 3.15.4 (9). The Nanopore reads with a length of >2,000 bp were assembled using Flye 2.9.1 (10), and a single circular contig was obtained. Next, SPAdes contigs were combined into replicons in SnapGene 6.1 (from Dotmatics; available at http://snapgene.com) using Flye data as the reference. The MGI reads were used to correct Nanopore or assembly errors using Bowtie 2 2.3.5.1 (11) and Pilon 1.24 (12) software. Default parameters were used for all software.

The *Gordonia polyisoprenivorans* 135 genome consists of a single 5,988,360-bp circular chromosome (GC content of 67.01%). Chromosome circularization was specified by ends overlapping. The average nucleotide identity (ANI) value was calculated using the EzBioCloud ANI calculator (13). DNA-DNA hybridization (DDH) was calculated using

1

the Genome-to-Genome Distance Calculator 3.0 (GGDC) (14). The ANI value with the type strain *Gordonia polyisoprenivorans* NBRC16320 (BAEI00000000.1) was 98.68%, and the DDH value was 88.40%.

The *G. polyisoprenivorans* 135 genome was annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) 4.6 (15). The chromosome contained 5,168 coding sequences, 3 rRNA clusters (5S, 16S, and 23S), 49 tRNAs, and 3 noncoding RNAs (ncRNAs). The genes of the PAH catabolic pathway were found in the genome of the strain 135, but their order was different from that of the strain *Gordonia* sp. CC-NAPH129-6 (2). The genome sequence data of *Gordonia polyisoprenivorans* 135 will enhance our understanding of the metabolism of PAH-degrading *Gordonia* strains.

The antiSMASH 6.0 (16) search for secondary metabolite clusters found 15 clusters on the chromosome, including clusters of ε-poly-L-lysine, ectoine, and aminopolycarboxylic-acid production.

Data availability. This genome project has been deposited at GenBank under the accession number CP116236. BioSample number SAMN32738803, and BioProject number PRJNA923796. SRA accession numbers were SRX19143368 for Oxford Nanopore data and SRX19144143 for DNBSEQ data.

REFERENCES

1. Delegan Y, Sushkova S, Minkina T, Filonov A, Kocharovskaya Y, Demin K, Gorovtsov A, Rajput VD, Zamulina I, Grigoryeva T, Dudnikova T, Barbashev A, Maksimov A. 2022. Diversity and metabolic potential of a PAH-degrading bacterial consortium in technogenically contaminated haplic chernozem, Southern Russia. Processes 10:2555. https://doi.org/10.3390/pr10122555.
2. Lin C-L, Shen F-T, Tan C-C, Huang C-C, Chen B-Y, Arun AB, Young C-C. 2012. Characterization of Gordonia sp. strain CC-NAPH129-6 capable of naphthalene degradation. Microbiol Res 167:395–404. https://doi.org/10.1016/j.micres.2011.12.002.
3. Emelyanova EV, Souzina NE, Polivtseva VN, Reshetilov AN, Solyanikova IP. 2017. Survival and biodegradation activity of Gordonia polyisoprenivorans 135: basics of a biosensor receptor. Appl Biochem Microbiol 53:580–586. https://doi.org/10.1134/S0003683817050039.
4. Solyanikova IP, Golovlev EL, Lisnyak OV, Golovleva LA. 1999. Isolation and characterization of catechol 1,2-dioxygenases from Rhodococcus rhodnii strain 135 and Rhodococcus rhodochrous strain 89: comparison with analogous enzymes of the ordinary and modified ortho-cleavage pathways. Biochemistry (Mosc) 64:824–831.
5. Evans CGT, Herbert D, Tempest DW. 1970. The continuous cultivation of microorganisms. 2. Construction of a chemostat. Methods Microbiol 2:e327. https://doi.org/10.1016/S0580-9517(08)70227-7.
6. Baboshin MA, Baskunov BP, Finkel'shtei±n ZI, Golovlev EL, Golovleva LA. 2005. The microbial transformation of phenanthrene and anthracene. Mikrobiologiia 74:357–364.
7. Suzina NE, Sorokin VV, Polivtseva VN, Klyueva VV, Emelyanova EV, Solyanikova IP. 2022. from rest to growth: life collisions of Gordonia polyisoprenivorans 135. Microorganisms 10:465. https://doi.org/10.3390/microorganisms10020465.
8. Solyanikova IP, Emelyanova EV, Shumkova ES, Egorova DO, Korsakova ES, Plotnikova EG, Golovleva LA. 2015. Peculiarities of the degradation of benzoate and its chloro- and hydroxy-substituted analogs by actinobacteria. Int Biodeter Biodegr 100:155–164. https://doi.org/10.1016/j.ibiod.2015.02.028.
9. 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.
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. Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
12. 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.
13. Yoon SH, Ha SM, Lim JM, Kwon SJ, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110:1281–1286. https://doi.org/10.1007/s10482-017-0844-4.
14. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60.
15. 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.
16. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. https://doi.org/10.1093/nar/gkab335.