AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of NAH7-Harboring *Pseudomonas putida* Strain G7

Paige M. Varner,[a] Claudia K. Gunsch[a]

[a]Department of Civil and Environmental Engineering, Duke University, Durham, NC, USA

ABSTRACT *Pseudomonas putida* G7 (111) is an aerobic bacterial species discovered in soil known to harbor the naphthalene-degrading NAH7 plasmid. Here, we report the genome sequence of G7. The genome was assembled by Nanopore sequencing and consisted of a chromosome of 6.4 Mbp with a G+C content of 62.13%.

Pseudomonas putida 111 (commonly referred to as strain G7 or ATCC 17485) is a bacterium known for harboring the naphthalene-degrading plasmid NAH7. It was first isolated by W.R. Sistrom from soil enriched with naphthalene (1). *P. putida* G7 (hereon G7) has been studied often in regard to its motility or as a host of the well-characterized NAH7 plasmid (2–5). We report the complete genome sequence of *P. putida* G7 obtained by Nanopore sequencing, to better understand *Pseudomonas* biology and as a host to a degradative plasmid of engineering significance.

G7 was obtained from ATCC under number 17485. A single colony was picked and cultured on Luria broth (LB) medium at 30°C until reaching its exponential growth phase (OD600 of 0.5). Total genomic DNA was extracted using Qiagen's Genomic DNA Extraction protocol with the Genomic-tip 500/G kit (Qiagen, Venlo, the Netherlands). A quality score of 2.7 for 280/260 absorbance was obtained prior to sequencing. Library preparation was performed using the Oxford Nanopore Technologies (ONT) Rapid Sequencing Kit (SQK-RAD004). Sequencing was performed on the ONT MinION using a R9.4.1 flow cell and fast base calling, and sequencing was monitored with the MinKNOW software v4.0.13 (ONT, Oxford, UK). There were a total of 227,618 postfiltered reads with an average length of 9,506 bp. Reads with a length of 10 kb or more were retained for assembly resulting in 200× coverage. Reads were assembled with Flye v2.8.1 with default parameters which includes one round of polishing (6). The assembly was further polished and circular contigs were reoriented to start at the origin of replication with Medaka model r941_min_fast_g303, v1.2.0f (ONT, Oxford, UK).

Assembly quality and completeness was assessed first by QUAST (v5.0.2) to report basic assembly statistics, such as N50. BUSCO (v4.1.4) was used to assess assembly completeness by searching for single-copy conserved orthologs using the "auto-lineage selection" parameter to select the appropriate odb10 database (7, 8). QUAST and BUSCO results were summarized with multiQC (v1.9) (9). An N50 of 6,376 kb was reported and 96.2% of BUSCOs were complete and single copy. To confirm the species sequenced and check for contamination, the contig from the assembly was queried against the blast nt database using the megablast algorithm (v2.7.1). Results showed a 99.168% identity to *Pseudomonas putida* KF715. Minimap2 (v2.17) was used to assess depth of coverage across the contigs by mapping reads to the polished assemblies (10). To check for mis-assemblies, each assembly was aligned to itself using minimap2 with the "asm5" preset settings for aligning genomes to one another. No mis-assemblies were found.

The complete genome of this strain is 6,375,733 bp with a G+C content of 62.13%. The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) found 6,067 total genes with 5,506 being protein coding and 104 being RNA genes (11). Twelve genes were

Editor David A. Baltrus, University of Arizona

Address correspondence to Claudia K. Gunsch, ckgunsch@duke.edu.

The authors declare no conflict of interest.

Received 3 May 2022
Accepted 3 September 2022
Published 21 September 2022

found to be involved in the chemotaxis of G7 which supports previous studies (3–5). Additionally, multiple Type IV pilus assembly proteins were identified which might aid in the conjugation/transfer of the NAH7 plasmid harbored in G7, making G7 attractive for its potential to be used in genetic bioaugmentation schemes for remediation of polycyclic aromatic hydrocarbons (12).

This is a report of the complete genome sequence of *P. putida* G7. This assembly will facilitate genome-wide comparison studies and development of microbiome engineering strategies for bioremediation.

Data availability. The chromosome sequence and annotation reported here was deposited in the National Center for Biotechnology Information (NCBI) GenBank Sequence Database under accession number CP096581. Raw, unfiltered fast5 sequences are submitted to the NCBI Sequence Resource Archive under BioProject accession number PRJNA831801.

ACKNOWLEDGMENTS

Thank you to Jennifer Modliszewski and the Genomic Analysis and Bioinformatics Core Facility at the Duke Center for Genomic and Computational Biology for assembly of the genome. This work was supported by the National Science Foundation Graduate Research Fellowship Program under Grant number DGE-1644868 and the NIEHS supported Duke University Superfund Research Center (P42ES010356).

REFERENCES

1. Sistrom WR. 1962. Microbial life. Holt, Rinehart and Winston, Ann Arbor, MI.
2. Sota M, Yano H, Ono A, Miyazaki R, Ishii H, Genka H, Top EM, Tsuda M. 2006. Genomic and functional analysis of the IncP-9 naphthalene-catabolic plasmid NAH7 and its transposon Tn4655 suggests catabolic gene spread by a tyrosine recombinase. J Bacteriol 188:4057–4067. https://doi.org/10.1128/JB.00185-06.
3. Law AMJ, Aitken MD. 2003. Bacterial chemotaxis to naphthalene desorbing from a nonaqueous liquid. Appl Environ Microbiol 69:5968–5973. https://doi.org/10.1128/AEM.69.10.5968-5973.2003.
4. Grimm AC, Harwood CS. 1997. Chemotaxis of Pseudomonas spp. to the polyaromatic hydrocarbon naphthalene. Appl Environ Microbiol 63:4111–4115. https://doi.org/10.1128/aem.63.10.4111-4115.1997.
5. Marx RB, Aitken MD. 1999. Quantification of chemotaxis to naphthalene by Pseudomonas putida G7. Appl Environ Microbiol 65:2847–2852. https://doi.org/10.1128/AEM.65.7.2847-2852.1999.
6. 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.
7. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
8. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.
9. Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354.
10. Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. Bioinformatics 32:2103–2110. https://doi.org/10.1093/bioinformatics/btw152.
11. 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.
12. Varner PM, Gunsch CK. 2021. Properties affecting transfer and expression of degradative plasmids for the purpose of bioremediation. Biodegradation 32:361–375. https://doi.org/10.1007/s10532-021-09950-1.