



# Complete Genome Sequence of Lignin-Degrading *Streptomyces* sp. Strain S6, Isolated from an Oil Palm Plantation in Malaysia

Nurtasbiyah Yusof,<sup>a</sup> Anahuda Abdullah Tahir,<sup>a</sup> Fatimah Azizah Riyadi,<sup>a</sup> Nurul Syazwani Ahmad Sabri,<sup>a</sup> Fazrena Nadia Md Akhir,<sup>a</sup> Nor'azizi Othman,<sup>b</sup> Kah Lock Chin,<sup>c</sup> Kuang Yu Tseng,<sup>d</sup> Siang Tean Chin,<sup>c</sup> Zuriati Zakaria,<sup>a</sup> Cheng Yu Lai,<sup>d</sup> Hirofumi Hara<sup>e</sup>

<sup>a</sup>Department of Environmental Engineering and Green Technology, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

<sup>b</sup>Department of Mechanical Precision Engineering, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

<sup>c</sup>Analisa Resources (M) Sdn Bhd, Shah Alam, Selangor, Malaysia

<sup>d</sup>Inong Agriculture Co., Ltd., Taipei, Taiwan, Republic of China

<sup>e</sup>Department of Chemical Process Engineering, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

**ABSTRACT** *Streptomyces* spp. are bacteria that are responsible for the degradation of aromatic compounds and produce secondary metabolites. Here, we present a complete genome sequence of *Streptomyces* sp. strain S6, which was isolated from an oil palm plantation, with a 7.8-Mbp linear chromosome, a GC content of 72%, and 4,266 coding sequences.

*Streptomyces* spp. play significant roles in producing numerous extracellular enzymes to degrade macromolecules such as lignocellulose. They also produce diverse bioactive compounds, including antibiotics, which has led to the study of their genomes (1, 2). Ion Torrent sequencing was combined with Nanopore sequencing to enhance the *de novo* assembly of a complete *Streptomyces* sp. genome.

*Streptomyces* sp. strain S6, which was isolated as a lignin degrader from an oil palm plantation in Malaysia, was grown for 7 days at 30°C on W minimal medium with 10 mM Kraft lignin as the sole carbon source. Genomic DNA was extracted using the QIAamp DNA minikit (Qiagen). For the Ion X5 XL system (Thermo Fisher Scientific), a 400-bp library was constructed using the Ion Xpress Plus fragment library kit (Thermo Fisher Scientific) according to the manufacturer's protocol. The quality of the purified library was checked using an Agilent 2100 bioanalyzer, and the library was diluted before proceeding to template preparation using the Ion Chef system (Thermo Fisher Scientific). Enriched Ion Sphere particles were loaded onto an Ion 530 chip, and sequencing was performed for 4 h. Torrent Suite software (Thermo Fisher Scientific) was used for raw data analysis, alignment, and variant calling. For MinION sequencing (Oxford Nanopore Technologies), genomic DNA was sheared using Covaris g-TUBEs and repaired with the NEBNext end repair/DA-tailing module and NEB blunt/TA ligase master mix. The library was constructed using ligation sequencing kit 1D (Oxford Nanopore Technologies) and NEB T4 DNA ligase without size selection. The purified library was loaded onto a primed MinION flow cell (FLO-MIN106) and sequenced on a GridION X5 sequencer for 48 h. Oxford Nanopore Technologies Albacore software (version 2.1.3) was used for base-calling analysis of unprocessed data.

Short reads from the Ion X5 XL system were quality trimmed using CLC Genomics Workbench software (version 11.0.1; CLC bio, Aarhus, Denmark) with the following parameters: quality score limit, 0.05; maximum number of ambiguous nucleotides, 2; and discarded reads, <400 nucleotides. The sequencing quality of long reads from the

**Citation** Yusof N, Abdullah Tahir A, Riyadi FA, Ahmad Sabri NS, Md Akhir FN, Othman N, Chin KL, Tseng KY, Chin ST, Zakaria Z, Lai CY, Hara H. 2020. Complete genome sequence of lignin-degrading *Streptomyces* sp. strain S6, isolated from an oil palm plantation in Malaysia. *Microbiol Resour Announc* 9:e01332-19. <https://doi.org/10.1128/MRA.01332-19>.

**Editor** Julie C. Dunning Hotopp, University of Maryland School of Medicine

**Copyright** © 2020 Yusof 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 Hirofumi Hara, [hara@utm.my](mailto:hara@utm.my).

**Received** 20 November 2019

**Accepted** 2 January 2020

**Published** 30 January 2020

MinION sequencing was checked using FastQC version 0.11.8, and potential remaining adaptors were trimmed with Porechop. Long reads were first assembled into a draft genome using Canu version 1.8 (3). Short reads were mapped onto the long-read assembly using minimap2 version 2.16 (4) and polished with Racon version 1.3.2 (5), with the entire process being repeated 10 times. The quality of the genome in each round was assessed using CheckM version 1.0.12 (6), and the seventh round was chosen as the final version.

The total genome size is 7,875,470 bp, in a single linear chromosome, with a GC content of 72%. The accuracy of the complete genome sequence is 92.14%, and no plasmid sequence was detected using PlasmidFinder version 2.1 (7). Annotation using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) predicted 4,266 coding sequences, 18 rRNA genes, and 65 tRNAs (8). Rapid Annotations using Subsystems Technology (RAST) revealed 365 genes for the metabolism of aromatic compounds and carbohydrates that could be responsible for lignocellulose degradation (9).

**Data availability.** The complete genome sequence of *Streptomyces* sp. S6 has been deposited at DDBJ/ENA/GenBank under accession number [CP040654](https://www.ncbi.nlm.nih.gov/nuccore/CP040654), BioProject number [PRJNA514949](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA514949), and BioSample number [SAMN10739602](https://www.ncbi.nlm.nih.gov/biosample/SAMN10739602), with SRA accession numbers [SRR9056892](https://www.ncbi.nlm.nih.gov/sra/SRR9056892) (MinION), [SRR8466423](https://www.ncbi.nlm.nih.gov/sra/SRR8466423) (Ion Torrent X5 XL), and [SRP180904](https://www.ncbi.nlm.nih.gov/sra/SRP180904) (fast5 files for MinION reads).

#### ACKNOWLEDGMENT

We thank the Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia (Kuala Lumpur, Malaysia), for financially supporting and facilitating our research.

#### REFERENCES

1. Wibberg D, Al-Dilaimi A, Busche T, Wedderhoff I, Schrempf H, Kalinowski J, Ortiz de Oru  Lucana D. 2016. Complete genome sequence of *Streptomyces reticuli*, an efficient degrader of crystalline cellulose. *J Biotechnol* 222:13–14. <https://doi.org/10.1016/j.jbiotec.2016.02.002>.
2. Wang P, Liu Z, Huang ZY. 2018. Complete genome sequence of soil actinobacteria *Streptomyces cavourensis* TJ430. *J Basic Microbiol* 58: 1083–1090. <https://doi.org/10.1002/jobm.201800181>.
3. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
4. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
5. Vaser R, Sovi  I, Nagarajan N, Siki  M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27:737–746. <https://doi.org/10.1101/gr.214270.116>.
6. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
7. Carattoli A, Zankari E, Garc a-Fern andez A, Voldby Larsen M, Lund O, Villa L, M ller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>.
8. Tatusova T, DiCuccio Badretdin M, Chetvernin AV, Nawrocki EP, Zaslavsky Lomsadze LA, 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>.
9. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.