



Complete Genome Sequence of *Stenotrophomonas maltophilia* 1800, a New Bacterial Strain with Potential for Bioremediation of Oil-Contaminated Environments

Annela Semai,^a Frédéric Plewniak,^a Joanna Lledo,^b Gwenolah Annonay,^b Céline Vandecasteele,^b Céline Lopez-Roques,^b
Philippe N. Bertin^a

^aGénétique Moléculaire, Génomique et Microbiologie, Université de Strasbourg, Strasbourg, France

^bINRAE, GeT-PlaGe, Genotoul, Castanet-Tolosan, France

ABSTRACT *Stenotrophomonas maltophilia* strain 1800 was isolated from the effluent of an industrial oil refinery in Algeria. Its genome was sequenced using Illumina MiSeq (2 × 150-bp read pairs) and Oxford Nanopore (long reads) technologies and assembled using Unicycler. It is composed of one chromosome of 4.83 Mb.

Most often described as an opportunistic pathogen (1), *Stenotrophomonas maltophilia* has also been shown to produce various biosurfactants/bioemulsifiers and to degrade diesel or motor oil (2, 3).

Stenotrophomonas maltophilia strain 1800 was isolated using the suspension-dilution method from a petrochemical refinery near Oran, Algeria (35°49'49.6"N, 0°19'31.6"W), on mineral salt medium (MSM) plates supplemented with trace elements and crude oil (1% vol/vol) (4). DNA was extracted from an LB liquid medium culture incubated for 72 h at 30°C using the MasterPure complete DNA and RNA purification kit (Epicentre) and used for both sequencing libraries.

DNA sequencing (DNA-seq) libraries were prepared following Illumina's protocols for the Illumina TruSeq Nano DNA low-throughput (LT) library prep kit. The library quality was assessed using an advanced analytical fragment analyzer (Agilent), and the libraries were quantified by quantitative PCR (qPCR) using the Kapa library quantification kit (Roche). Paired-end sequencing (2 × 150 bp) was performed on an Illumina MiSeq instrument with the Illumina MiSeq Reagent Kit v2 Micro.

Oxford Nanopore Technologies (ONT) libraries were prepared according to the manufacturer's instructions for 1D native barcoding genomic DNA (EXP-NBD104 and SQK-LSK109 kits). The DNA was quantified using the Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit (Life Technologies), and the purity was determined using a NanoDrop instrument (Thermo Fisher Scientific). The size distribution was assessed using the fragment analyzer (AATI) high-sensitivity DNA fragment analysis kit. The DNA was purified using AMPure XP beads (Beckman Coulter) and sheared at 20 kb (speed 32) using the Megaruptor III system (Diagenode). Following DNA damage repair, end repair, a dA-tailing step, and sample-specific index ligation, the library was loaded onto an R9.4.1 revD flow cell and sequenced on a GridION instrument at 0.013 pmol within 48 h using MinKNOW v21.05.12 and Guppy v5.0.12 for base calling.

Default parameters were used for all software unless otherwise specified. Short reads with an average quality lower than 35 were filtered out using BBDuk v38.91 from JGI's BBTools suite (<https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide>). Long reads with an average quality greater than 9 were further filtered using Filtlong v0.2.1 (<https://github.com/rwick/Filtlong>), using the short reads as an external reference for quality assessment. The best reads longer than 20 kb were kept for a cumulative length of 500 Mb.

Editor J. Cameron Thrash, University of Southern California

Copyright © 2022 Semai 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 Philippe N. Bertin, philippe.bertin@unistra.fr.

The authors declare no conflict of interest.

Received 16 November 2021

Accepted 4 February 2022

Published 17 February 2022

Hybrid assembly with Unicycler v0.4.8 (5) of the long and short filtered reads yielded one complete circular 4,837,110 bp-long contig with a GC content of 66.19%. The average coverage estimated using QUAST v5.0.2 (6) was 136×, with 100% of reads mapping back onto the assembly. All 659 marker genes for *Xanthomonadaceae* were found using CheckM v1.1.3 (7) in the assembly, suggesting that the genome was 100% complete. Potential contamination was insignificant (0.11%), and no strain heterogeneity was identified. A GTOTree v1.5.47 phylogenomic analysis using 172 gammaproteobacterium single-copy genes in 204 *S. maltophilia* reference strains and *Stenotrophomonas pictorum* JCM 9942 as an outgroup suggested that *S. maltophilia* 1800 belongs to the *S. maltophilia* complex genogroup F (8, 9).

MicroScope automatic genome annotation (10) predicted 4,671 coding sequences (CDS), 70 tRNAs, and 13 rRNA genes. BLASTP (11) searches using *Geobacillus thermoleovorans* B23 LadA α , LadA β , and LadB as queries identified one putative LadA β alkane monooxygenase in *S. maltophilia* 1800 (12).

Data availability. The raw reads and assembly of the *Stenotrophomonas maltophilia* 1800 genome have been deposited in DDBJ/EMBL/GenBank under BioProject accession number [PRJEB48474](https://www.ncbi.nlm.nih.gov/bioproject/PRJEB48474). The raw reads are available in the SRA under accession numbers [ERR7198268](https://www.ncbi.nlm.nih.gov/sra/ERR7198268) for the MiSeq paired-end reads and [ERR7198269](https://www.ncbi.nlm.nih.gov/sra/ERR7198269) for the ONT long reads. The sequence of the chromosome is available in RefSeq under the accession number [NZ_OU943334](https://www.ncbi.nlm.nih.gov/refseq/NC_020943334).

ACKNOWLEDGMENTS

This work benefited from the financial support of the DMO/DMA project, subsidized by the CNRS national program EC2CO-MicrobiEn. It was performed in collaboration with the GeT core facility, Toulouse, France (<http://get.genotoul.fr>), and was supported by the France Génomique National Infrastructure, funded as part of the “Investissement d’avenir” program managed by Agence Nationale pour la Recherche (contract number ANR-10-INBS-09). We thank the Genoscope (Évry, France) for providing online access to MicroScope, the Microbial Genome Annotation and Analysis Platform (<https://mage.genoscope.cns.fr/microscope/home/index.php>), and the Algerian Ministry for Higher Education and Scientific Research for a scholarship to A. Semai.

REFERENCES

- Trifonova A, Strateva T. 2019. *Stenotrophomonas maltophilia*—a low-grade pathogen with numerous virulence factors. *Infect Dis (Lond)* 51: 168–178. <https://doi.org/10.1080/23744235.2018.1531145>.
- Deepali K, Sneha P, Sucheta P. 2014. Larvicidal activity of rhamnolipid biosurfactant produced by *Stenotrophomonas maltophilia*. *Int J Sci Eng Res* 5: 60–63.
- Larik IA, Qazi MA, Phulpoto AH, Haleem A, Ahmed S, Kanhar NA. 2019. *Stenotrophomonas maltophilia* strain 5DMD: an efficient biosurfactant-producing bacterium for biodegradation of diesel oil and used engine oil. *Int J Environ Sci Technol* 16:259–268. <https://doi.org/10.1007/s13762-018-1666-2>.
- Semai A, Plewniak F, Charrié-Duhaut A, Sayeh A, Gil L, Vandecasteele C, Lopez-Roques C, Leize-Wagner E, Bensalah F, Bertin PN. 2021. Characterisation of hydrocarbon degradation, biosurfactant production, and biofilm formation in *Serratia* sp. Tan611: a new strain isolated from industrially contaminated environment in Algeria. *Antonie Van Leeuwenhoek* 114:411–424. <https://doi.org/10.1007/s10482-021-01527-5>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- 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>.
- 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>.
- Vinuesa P, Ochoa-Sánchez LE, Contreras-Moreira B. 2018. GET_PHYLO-MARKERS, a software package to select optimal orthologous clusters for phylogenomics and inferring pan-genome phylogenies, used for a critical geno-taxonomic revision of the genus *Stenotrophomonas*. *Front Microbiol* 9:771. <https://doi.org/10.3389/fmicb.2018.00771>.
- Mercier-Darty M, Royer G, Lamy B, Charron C, Lemenand O, Gomart C, Fourreau F, Madec J-Y, Jumas-Bilak E, Decousser J-W. 2020. Comparative whole-genome phylogeny of animal, environmental, and human strains confirms the genogroup organization and diversity of the *Stenotrophomonas maltophilia* complex. *Appl Environ Microbiol* 86:e02919-19. <https://doi.org/10.1128/AEM.02919-19>.
- Vallenet D, Calteau A, Dubois M, Amours P, Bazin A, Beuvin M, Burlot L, Bussell X, Fouteau S, Gautreau G, Lajus A, Langlois J, Planel R, Roche D, Rollin J, Rouy Z, Sabatet V, Médigue C. 2020. MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Res* 48:D579–D589. <https://doi.org/10.1093/nar/gkz926>.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
- Boonmak C, Takahashi Y, Morikawa M. 2014. Cloning and expression of three *ladA*-type alkane monooxygenase genes from an extremely thermophilic alkane-degrading bacterium *Geobacillus thermoleovorans* B23. *Extremophiles* 18:515–523. <https://doi.org/10.1007/s00792-014-0636-y>.