



# Draft Genome Sequence of *Streptomyces morookaense* DSM 40503, an 8-Azaguanine-Producing Strain

Yueyu Hei,<sup>a</sup> Ziyi Li,<sup>a</sup> Yuhan Zhou,<sup>a</sup> Chenhao Hu,<sup>a</sup> Jin-Ming Gao,<sup>a</sup>  Jianzhao Qi<sup>a</sup>

<sup>a</sup>Shaanxi Key Laboratory of Natural Products & Chemical Biology, College of Chemistry & Pharmacy, Northwest A&F University, Yangling, Shaanxi, China

**ABSTRACT** Here, we describe the genome of *Streptomyces morookaense* DSM 40503, an 8-azaguanine-producing strain. The genome is the basis for future study and presents an underexplored taxonomy and biosynthetic potential, which expands our understanding of the diversity of microorganisms that produce nitrogen heterocyclic compounds.

The compound 8-azaguanine is an important natural guanine nucleoside analog. Naturally derived 8-azaguanine was initially reported as an antifungal compound (1). Subsequent studies revealed that 8-azaguanine has significant anticancer effects (2) and can inhibit protein synthesis (3), making it a drug candidate for early anticancer therapy (4). *Streptomyces morookaense* DSM 40503 is a type strain stored in the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ) (5), which we ordered in 2015 and preserved in our laboratory. The whole-genome sequencing of this strain was performed in 2018. Its other name is *Streptovercillium* sp. JCM 4673 (6), which was reported to produce 8-azaguanine (7). This strain was grown on GYM streptomyces medium 65 at 28°C and 7.2 pH before the addition of agar. GYM streptomyces medium contained 4.0 g glucose, 4.0 g yeast extract, 10.0 g malt extract, 2.0 g CaCO<sub>3</sub>, and 12.0 g agar. GYM liquid medium without agar is used for mycelium culture.

The culture solution was centrifuged at 5,000 × *g* for 10 min to obtain mycelium. About 20 mg mycelium was put into a 1.5-ml microcentrifuge tube and resuspended with 200 μl Tris-EDTA (TE) buffer. DNA extraction and purification were done using a PureLink microbiome DNA purification kit (Thermo Fisher, USA) following the manufacturer's instructions. A DNA library was prepared using the MagMAX-96 DNA multi-sample kit (Thermo Fisher, USA) and protocol (Illumina, CA) and was subsequently paired-end sequenced on an Illumina NextSeq platform using a 2 × 150-bp sequencing kit. Since the original sequencing data of Illumina X Ten had some data of lower quality, in order to make the subsequent assembly more accurate, the original data were quality cut with seqtk trimfq version 1.3 (<https://github.com/lh3/seqtk>) and TrimGalore version 0.5.0 (<https://github.com/FelixKrueger/TrimGalore/releases>). *De novo* assembly was carried out using SOAPdenovo version 2.04 (8). Predictions of genes, coding DNA sequences (CDS), and RNA genes were carried out using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (9). Tandem repeats were predicted with tandem repeats finder (TRF) software (10). The biosynthetic gene cluster for 8-azaguanine was predicted with antiSMASH version 5.0 (11). All software runs used default parameters.

The *S. morookaense* genome was recovered at 924× coverage as 4,760,411 reads, which were assembled into 238 contigs. The draft genome sequence has an *N*<sub>50</sub> value of 127,665 bp and a total of 7,469,426 bp with a 71.66% GC content. A total of 6,855 genes were annotated using PGAP and deposited in the GenBank database. Among them, 6,781 CDS consisted of 6,334 protein-coding genes and 447 pseudogenes, and 74 RNA coding genes were responsible for 66 tRNAs, 3 noncoding RNAs (ncRNAs), 1 16S

**Citation** Hei Y, Li Z, Zhou Y, Hu C, Gao J-M, Qi J. 2020. Draft genome sequence of *Streptomyces morookaense* DSM 40503, an 8-azaguanine-producing strain. *Microbiol Resour Announc* 9:e00518-20. <https://doi.org/10.1128/MRA.00518-20>.

**Editor** J. Cameron Thrash, University of Southern California

**Copyright** © 2020 Hei 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 Jianzhao Qi, [qjz@nwafu.edu.cn](mailto:qjz@nwafu.edu.cn).

**Received** 30 May 2020

**Accepted** 6 July 2020

**Published** 30 July 2020

rRNA, and 4 23S (partial) rRNAs. The total length of 943 tandem repeats is 174,279 bp, accounting for 1.70% of the genome.

*Streptomyces morookaense* DSM 40503 is closely related to *Streptomyces luteoverticillatus* CGMCC 15060 (GenBank accession number [CP034587.1](https://doi.org/10.1093/bioinformatics/btp336)), and their 16S rRNA gene identity was 98.89%. Based on the gene cluster for 8-azaguanine in *Streptomyces pathocidini* NRRL b-24287 (12), the 8-azaguanine gene cluster was found in *Streptomyces morookaense* DSM 40503 with the aid of antiSMASH and 2ndFind (<http://biosyn.nih.gov/jp/2ndfind/>). A total of 52 biosynthetic gene clusters involved in the production of different bioactive compounds were predicted in this genome. This work broadened our understanding of the genomes of bacteria producing nitrogen-containing compounds.

**Data availability.** This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number [JABBXF000000000](https://doi.org/10.1093/bioinformatics/btp336). The version described in this paper is version [JABBXF000000000.1](https://doi.org/10.1093/bioinformatics/btp336). The FASTQ files of the raw reads were deposited in the NCBI SRA under accession number [PRJNA625712](https://doi.org/10.1093/bioinformatics/btp336). The 8-azaguanine gene cluster accession number is [MT344128](https://doi.org/10.1093/bioinformatics/btp336).

## ACKNOWLEDGMENTS

This work was supported by the Natural Science Basic Research Plan in the Shaanxi province of China (2019JQ-046), the China Postdoctoral Science Foundation (2019M653760), and the National Natural Science Foundation of China (31800031).

## REFERENCES

1. Anzai K, Nagatsu J, Suzuki S. 1961. Pathocidin, a new antifungal antibiotic, I. Isolation, physical and chemical properties, and biological activities. *J Antibiot (Tokyo)* 14:340–342.
2. Zimmerman EF, Greenberg SA. 1965. Inhibition of protein synthesis by 8-azaguanine. I. Effects on polyribosomes in HeLa cells. *Mol Pharmacol* 1:113–125.
3. Zimmerman EF, Holler BW, Pearson GD. 1967. Azaguanine inhibition of protein synthesis. II. Effects of poly (U) in *Bacillus cereus*. *Biochim Biophys Acta Nucleic Acids Protein Synthesis* 134:402–410. [https://doi.org/10.1016/0005-2787\(67\)90017-2](https://doi.org/10.1016/0005-2787(67)90017-2).
4. Nelson JA, Carpenter JW, Rose LM, Adamson DJ. 1975. Mechanisms of action of 6-thioguanine, 6-mercaptopurine, and 8-azaguanine. *Cancer Res* 35:2872–2878.
5. DSMZ. *Streptomyces morookaense*. <https://www.dsmz.de/collection/catalogue/details/culture/DSM-40503>.
6. UniProt. <https://www.uniprot.org/taxonomy/1970>.
7. Sugiyama M, Takeda A, Paik SY, Nimi O, Nomi R. 1986. Acetylation of blasticidin S by its producing actinomycetes. *J Antibiot (Tokyo)* 39: 827–832. <https://doi.org/10.7164/antibiotics.39.827>.
8. Li R, Yu C, Li Y, Lam T-W, Yiu S-M, Kristiansen K, Wang J. 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25: 1966–1967. <https://doi.org/10.1093/bioinformatics/btp336>.
9. Tatusova T, DiCuccio M, Badretdin A, Chetvermin 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>.
10. Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580. <https://doi.org/10.1093/nar/27.2.573>.
11. 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–W87. <https://doi.org/10.1093/nar/gkz310>.
12. Zhao G, Guo Y-Y, Yao S, Shi X, Lv L, Du Y-L. 2020. Nitric oxide as a source for bacterial triazole biosynthesis. *Nat Commun* 11:1614. <https://doi.org/10.1038/s41467-020-15420-8>.