Draft Genome Sequence of *Glycomyces fuscus* TRM 49117, Isolated from a Hypersaline Soil Sample

Yang Wang, Xiaoxia Luo, Lili Zhang

College of Life Science, Key Laboratory of Protection and Utilization of Biological Resources in Tarim Basin of Xinjiang Production & Construction Corps, Tarim University, Alar, China

**ABSTRACT** *Glycomyces* spp. are rare actinomycetes that are potential antibiotic producers. Here, we report the draft genome sequence of *Glycomyces fuscus* TRM 49117. This is the first genome report of a bacterium belonging to the genus *Glycomyces*. The genome information of *G. fuscus* will contribute to studies on the structure and function of antibiotics.

In a previous study, a strain was reported to show antibacterial activity against the rare actinomycete *Glycomyces harbinensis* ATCC 43155T (1). It was previously discovered that the high-temperature, drought-ridden, and high-salinity Lop Nur environment contains a large number of *Glycomyces* actinomycetes. New species may produce new metabolites and active substances. *Glycomyces fuscus* TRM 49117 was originally isolated from a soil sample from the Lop Nur salt lake in Xinjiang, China (39°35’N 89°46’E). It can tolerate a higher salt concentration than can *G. harbinensis* ATCC 43155T or *Glycomyces* spp. from other environments and can grow in more than 10% ISP4 medium, and it contains a more abundant polar lipid than does *G. harbinensis* ATCC 43155T or *Glycomyces* spp. from other environments (2).

The genome was sequenced using an Illumina HiSeq 2000 platform. Two different sequencing libraries, a 500-bp paired-end and 6,000-bp mate pair library, were used. Sequencing yielded 230-fold genome coverage. Assembly of all sequence reads (1,002-Mbp paired-end reads and 808-Mbp mate pair reads) applying the SOAPdenovo 1.05 software (3) resulted in a draft genome comprising 18 scaffolds, with 63 contigs with an N50 size of 193,136 bp and a longest contig of 613,903 bp. The total scaffold size was determined to be 6,556,887 bp, with a GC content of 72.84%. Genome annotation was accomplished using the Rapid Annotations using Subsystems Technology (RAST) server (4). Briefly, protein-coding genes were identified using Glimmer version 3.0 (5). Ribosomal RNAs were predicted by a sequence similarity search using BLAST against an RNA sequence database and/or using the Infernal and Rfam models (6, 7). tRNAs were predicted using tRNAscan-SE (8). The predicted coding sequences (CDSs) were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant (NR), UniProt, TIGR-Fam, Pfam, PRIAM, KEGG, COG, Swiss-Prot, TrEMBL, GO, and InterPro databases. Other noncoding genes were predicted using Infernal 1.0 (9). Additional gene prediction analysis and functional annotation were performed within the Integrated Microbial Genomes-Expert Review (IMG-ER) platform (10). Gene prediction revealed 5,666 protein-coding genes, 9 rRNA operons, and 74 tRNA genes. Among these protein-coding sequences, 5,582 (98.52%) have been assigned to a Clusters of Orthologous Groups (COG) functional category (11).

The complete genome sequence includes a number of gene regions devoted to the biosynthesis of secondary metabolites. The *G. fuscus* TRM 49117 genome sequence was scanned by antiSMASH (12), and we found 4 type I and 1 type II polyketide biosynthesis gene clusters and 6 nonribosomal peptide synthetase biosynthesis gene clusters. Also,
we found 5 gene clusters for a bacteriocin-like bioactive substance, which is usually produced by Gram-positive bacteria and can inhibit other closely related Gram-positive bacteria (13).

The G. fuscus TRM 49117 genome information will serve as a valuable reference to elucidate the potential of hypersaline-derived Glycomyces strains as promising sources for bioactive secondary metabolites. The G. fuscus TRM 49117 genome sequence provides us the opportunity to further understand the evolutionary relationship between Glycomyces strains and to explain the physiological mechanism for its growth in hypersaline soil. The genome sequence of G. fuscus was determined, and the genes possibly involved in the antibacterial activity were determined.

Accession number(s). The G. fuscus TRM 49117 draft genome sequence was deposited in the NCBI database under the accession no. NYMS00000000. The version described in this paper is the first version.

ACKNOWLEDGMENT
This work was supported by grants from the National Natural Science Foundation of China (grants 31560002 and 31500009).

REFERENCES
1. Lee MD, Borders DB, Labeda DP, Fantini AA, Testa RT. February 1986. Preparation of antibiotic LL-D05139/H9252 from cultures of Glycomyces harbinensis, gen. sp. nov. U.S. patent 4568646 A.
2. Han XX, Luo XX, Zhang LL. 2014. Glycomyces fuscus sp. nov. and Glycomyces albus sp. nov., actinomycetes isolated from a hypersaline habitat. Int J Syst Evol Microbiol 64:2437–2441. https://doi.org/10.1099/ijs.0.061788-0.
3. Li R, Zhu H, Ruan J, Qian W, Fung X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J. 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res 20:265–272. https://doi.org/10.1101/gr.097261.109.
4. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsmas K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeeka RA, McNeil JK, Paarmann D, Paczian T, Parrello B, Puch 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.
5. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with Glimmer. Nucleic Acids Res 27:4636–4641. https://doi.org/10.1093/nar/27.23.4636.
6. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. 2003. Rfam: an RNA family database. Nucleic Acids Res 31:439–441. https://doi.org/10.1093/nar/gkg006.
7. Eddy SR. 2002. A memory-efficient dynamic programming algorithm for optimal alignment of a sequence to an RNA secondary structure. BMC Bioinformatics 3:18. https://doi.org/10.1186/1471-2105-3-18.
8. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964.
9. Nawrocki EP, Kolbe DL, Eddy SR. 2009. Infernal 1.0: inference of RNA alignments. Bioinformatics 25:1335–1337. https://doi.org/10.1093/bioinformatics/btp157.
10. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25:2271–2278. https://doi.org/10.1093/bioinformatics/btp393.
11. Kristensen DM, Kannan L, Coleman MK, Wolf YI, Sorokin A, Koonin EV, Mushhegan A. 2010. A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. Bioinformatics 26:1481–1487. https://doi.org/10.1093/bioinformatics/btp229.
12. Blin K, Medema MH, Kazempour D, Fischbach MA, Bretilting R, Takano E, Weber T. 2013. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res 41:W204–W212. https://doi.org/10.1093/nar/gkt449.
13. Wiedemann I, Breukink E, van Kraaij C, Kuipers OP, Bierbaum G, de Kruijff B, Sahl HG. 2001. Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. J Biol Chem 276:1772–1779. https://doi.org/10.1074/jbc.M006770200.