JOURNAL OF BACTERIOLOGY, July 2011, p. 3417–3418 Vol. 193, No. 13
0021-9193/11/$12.00 doi:10.1128/JB.05053-11

Draft Genome Sequence of the Marine Bacterium *Streptomyces griseoaurantiacus* M045, Which Produces Novel Manumycin-Type Antibiotics with a pABA Core Component▽

Fuchao Li,[1,2] Peng Jiang,[2] Huajun Zheng,[3] Shengyue Wang,[3] Guoping Zhao,[3] Song Qin,[2]* and Zhaopu Liu[1]*

Key Laboratory of Marine Biology, Jiangsu Province, Nanjing Agricultural University, Nanjing, China[1]; Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China[2]; and Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, China[3]

Received 11 April 2011/Accepted 25 April 2011

***Streptomyces griseoaurantiacus* M045, isolated from marine sediment, produces manumycin and chinikomycin antibiotics. Here we present a high-quality draft genome sequence of *S. griseoaurantiacus* M045, the first marine *Streptomyces* species to be sequenced and annotated. The genome encodes several gene clusters for biosynthesis of secondary metabolites and has provided insight into genomic islands linking secondary metabolism to functional adaptation in marine *S. griseoaurantiacus* M045.**

Owing to inhabiting a complex marine environment, marine streptomycetes are robust sources of new bioactive compounds (4, 9, 5). *Streptomyces griseoaurantiacus* M045, isolated from marine sediment, produces manumycin and chinikomycin antibiotics (11). Chinikomycins are the first-discovered truly natural manumycin antibiotics with a *para* orientation of the side chains (*p*-aminobenzoic acid core component, pABA) (8, 11, 6, 16). Herein we present a draft genome sequence of marine *S. griseoaurantiacus* M045.

The nucleotide sequence was determined using a 454 GS FLX sequencer (14). A total of 607,480 reads including up to 276,865,258 bp were obtained, which represented a 36-fold coverage of the genome. Assembly was performed using the GS De Novo Assembler software program (454). In addition, 570-fold coverage of pair-end sequences (2 × 120 bp) produced by Illumina Solexa technology was mapped to the genome sequence to promote sequence quality and fill in gaps. Finally, we obtained the *S. griseoaurantiacus* M045 draft genome of 7,712,377 bp distributed in 46 contigs with a GC content of 72.73%.

Putative protein-coding sequences were predicted using the Glimmer (3) and GeneMark (13) software programs. Functional annotation was based on BLASTP with the KEGG and NR databases. tRNA genes were directly predicted using the tRNAscan-SE tool (12). The signal peptide cleavage sites, transmembrane topologies, and lipoproteins were predicted using the SignalP 3.0 software program with hidden Markov models (2), ConPred II (1), and LipoP 1.0 (10), respectively. Genomic islands were predicted on the IslandViewer web service (http://www.pathogenomics.sfu.ca/islandviewer) using the IslandPath-DIMOB (7) and SIGI-HMM (18) methods.

The genome consists of one linear chromosome with 6 rRNA operons, 65 tRNA genes, and 6,839 protein-coding genes (CDSs). For the CDSs, 5,003 proteins could be assigned to COG (cluster of orthologous groups) families (17). Six hundred fourteen hypothetical proteins have no match to any known proteins in the databases. At least 660 multigene (paralog) families containing 2,693 predicted proteins were identified. Two-component regulatory systems, which detect and respond to changes in the marine environment, are widely distributed on the chromosome. As for the subcellular localization of the proteins, 234 proteins were identified as secreted proteins, 1,352 proteins as transmembrane proteins, and 106 proteins as lipoproteins. In addition, 18 genomic islands occur in the genome; the genomic islands link secondary metabolism to functional and environmental adaptation in marine *S. griseoaurantiacus* M045 (15).

Genome analysis revealed a number of genes related to biosynthesis of secondary metabolites. The gene cluster for manumycin is located on one genomic island and contains a total of 28 genes, including 3-amino-4-hydroxybenzoic acid (3,4-AHBA) synthase, 3,4-AHBA carrier protein, 3-oxoacyl-(acyl carrier protein) synthase, and 5-aminolevulinate synthase. Chinikomycins contain a pABA core component but share a partial gene cluster with manumycin that consists of two triene polyketides and a 2-amino-3-hydroxycyclopent-2-enone biosynthetic pathway.

S. griseoaurantiacus M045 is the first marine streptomycete for which a genome sequence has been reported. We believe this work provides insight into the functional adaptation and combinatorial biosynthesis of bioactive molecules produced by marine streptomycetes.

* Corresponding author. Mailing address for Zhaopu Liu: Key Laboratory of Marine Biology, Jiangsu Province, Nanjing Agricultural University, Nanjing 210095, China. Phone and fax: 86-25-84396678. E-mail: sea@njau.edu.cn. Mailing address for Song Qin: Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China. Phone: 86-532-82898500. Fax: 86-532-82898919. E-mail: sqin@qdio.ac.cn.

▽ Published ahead of print on 6 May 2011.

Nucleotide sequence accession numbers. The genome sequence has been deposited at DDBJ/EMBL/GenBank under the accession no. AEYX00000000. The version described in this paper is the first version, AEYX01000000.

This work was supported by the National Natural Science Foundation of China (31000037), the Key Innovative Project of the Chinese Academy of Science (KSCX2-YW-G-073 and KZCX2-YW-216), the Natural Science Foundation of Shandong Province (ZR2009EQ004), and the Foundation of the Key Laboratory of Marine Bioactive Substance and Modern Analytical Techniques, SOA (MBSMAT-2010-07).

REFERENCES

1. **Arai, M., et al.** 2004. ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. Nucleic Acids Res. **32:**W390–W393.
2. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340:**783–795.
3. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636–4641.
4. **Fenical, W.** 1993. Chemical studies of marine bacteria: developing a new resource. Chem. Rev. **93:**1673–1683.
5. **Fenical, W., and P. R. Jensen.** 2006. Developing a new resource for drug discovery: marine actinomycete bacteria. Nat. Chem. Biol. **2:**666–673.
6. **Floss, H. G.** 2006. From ergot to ansamycins—45 years in biosynthesis. J. Nat. Prod. **69:**158–169.
7. **Hsiao, W., I. Wan, S. J. Jones, and F. S. Brinkman.** 2003. IslandPath: aiding detection of genomic islands in prokaryotes. Bioinformatics **19:**418–420.
8. **Hu, Y., and H. G. Floss.** 2004. Further studies on the biosynthesis of the manumycin-type antibiotic, asukamycin, and the chemical synthesis of protoasukamycin. J. Am. Chem. Soc. **126:**3837–3844.
9. **Jensen, P. R., T. J. Mincer, P. G. Williams, and W. Fenical.** 2005. Marine actinomycete diversity and natural product discovery. Antonie Van Leeuwenhoek **87:**43–48.
10. **Juncker, A. S., et al.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci. **12:**1652–1662.
11. **Li, F., et al.** 2005. Chinikomycins A and B: isolation, structure elucidation, and biological activity of novel antibiotics from a marine *Streptomyces* sp. isolate M045. J. Nat. Prod. **68:**349–353.
12. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
13. **Lukashin, A. V., and M. Borodovsky.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107–1115.
14. **Margulies, M., et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.
15. **Penn, K., et al.** 2009. Genomic islands link secondary metabolism to functional adaptation in marine actinobacteria. ISME J. **3:**1193–1203.
16. **Rui, Z., et al.** 2010. Biochemical and genetic insights into asukamycin biosynthesis. J. Biol. Chem. **285:**24915–24924.
17. **Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28:**33–36.
18. **Waack, S., et al.** 2006. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinformatics **7:**142.