

Draft Genome Sequence of the Marine Sediment-Derived Actinomycete *Streptomyces xinghaiensis* NRRL B24674^T

Xinqing Zhao* and Tianhong Yang

School of Life Science and Biotechnology, Dalian University of Technology, Dalian 116024, China

Received 29 June 2011/Accepted 18 July 2011

Actinobacteria are a rich source of novel natural products. We recently characterized *Streptomyces xinghaiensis* NRRL B24674^T as a novel species of marine origin. Here we report the draft genome sequence of this species. This is the first validly published marine streptomycete for which a genome sequence has been presented.

Streptomyces xinghaiensis NRRL B24674^T, a novel species of *Streptomyces* (10), was isolated from a marine sediment sample collected from Xinghai Bay, Dalian, China. Herein we present a draft genome sequence of *S. xinghaiensis*.

The nucleotide sequence was determined using Roche/454 pyrosequencing and Illumina/Solexa sequencing by synthesis technology. The Mate-paired reads generated by the Solexa sequencer were assembled by SOAP (6), and the contigs were split into fragments. Then 454 reads and the Solexa sequence fragments were assembled with the Newbler assembler. Coding sequences (CDSs) were predicted by Prodigal (4). Functional assignment was obtained by performing a sequence similarity search with BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) against the COG (<http://www.ncbi.nlm.nih.gov/COG/>) reference database and the nonredundant GenBank CDS database. A set of in-house Perl scripts and EMBOSS (<http://emboss.sourceforge.net/>) were used for sequence manipulation. Functional annotation was based on BLASTP with the KEGG databases. tRNA genes were directly predicted using the tRNAscan-SE tool (8). The phylogenetic tree was reconstructed by the CVTtree server (7).

The *S. xinghaiensis* draft genome contains 7,618,725 bp with a GC content of 72.5%, representing approximately 92.7% of the 8.2-Mb estimated size of the genome. The genome consists of one linear chromosome with 6 rRNA operons, 65 tRNA genes, and 6,654 CDSs. For the CDSs, 5,563 proteins could be assigned to COG families. Four thousand nine hundred eighty putative proteins were matched to *Streptomyces* spp., and 1,091 CDSs encode proteins with no match to any known proteins in the databases.

Genome analysis revealed a number of genes related to biosynthesis of secondary metabolites. At least 15 clusters involved in secondary metabolism were identified; these include one gene cluster that highly resembles the gene cluster of ribostamycin (9), an aminoglycoside antibiotic. Other gene clusters for the biosynthesis of polyketides, terpenes, and non-ribosomal peptide antibiotics were also inferred. Many putative antibiotic biosynthesis genes show low identity with the known ones, which indicates that the species is a potential producer of novel natural products.

The *S. xinghaiensis* draft genome also contains genes encoding cellulases, xylanases, and chitinases, as well as genes encoding protease, lipase, peptidase, and esterase. Genes involved in the resistance to heavy metals mercury, copper, and nickel were also identified, indicating the potential application of this strain in biomass bioconversion and environmental bioremediation.

Although the genomic sequences of marine sediment-derived strains *Streptomyces* strain PP-C42 (2) and *Streptomyces griseoaurantiacus* M045 (5) were reported recently, *S. xinghaiensis* is the first validly published marine streptomycete for which a genome sequence has been reported. Comparison of the genomes of these marine sediment-derived strains with those of terrestrial origin will provide insight into the environmental adaptation and evolution of *Streptomyces* species. The genome mining of *S. xinghaiensis* will further explore the chemical diversity and genetic diversity of this species for discovery of novel compounds and enzymes for biotechnology applications (1, 3).

Nucleotide sequence accession numbers. The genome sequence has been deposited at DDBJ/EMBL/GenBank under accession no. AFRP00000000. The version described in this paper is the first version, deposited under AFRP01000000.

This work was supported by a grant from the Open Project of the State Key Laboratory of Bioreactor Engineering, China.

We are grateful to the Tianjin Biochip Corporation, China, and for the help of Junjing-Wang for bioinformatic analysis.

REFERENCES

1. Corre, C., G. L. Challis. 2009. New natural product biosynthetic chemistry discovered by genome mining. *Nat. Prod. Rep.* **26**:977–986.
2. Fan, L., et al. 2011. The draft genome sequence of a marine *Streptomyces* sp. strain PP-C42 isolated from the Baltic Sea. *J. Bacteriol.* **193**:3691–3692.
3. Gulder, T. A., and B. S. Moore. 2009. Chasing the treasures of the sea—bacterial marine natural products. *Curr. Opin. Microbiol.* **12**:252–260.
4. Hyatt, D., et al. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**:119.
5. Li, F., et al. 2011. Draft genome sequence of the marine bacterium *Streptomyces griseoaurantiacus* M045, which produces novel manumycin-type antibiotics with a pABA core component. *J. Bacteriol.* **193**:3417–3418.
6. Li, R., Y. Li, K. Kristiansen, and J. Wang. 2008. SOAP: short oligonucleotide alignment program. *Bioinformatics* **5**:713–714.
7. Qi, J., H. Luo, and B. L. Hao. 2004. CVTtree: a phylogenetic tree reconstruction tool based on whole genomes. *Nucleic Acids Res.* **32**:W45–W47.
8. Schattner, P., A. N. Brooks, and T. M. Lowe. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* **33**:W686–W689.
9. Subba, B., et al. 2005. The ribostamycin biosynthetic gene cluster in *Streptomyces ribosidificus*: comparison with butirosin biosynthesis. *Mol. Cells* **20**:90–96.
10. Zhao, X. Q., et al. 2009. *Streptomyces xinghaiensis* sp. nov., isolated from marine sediment. *Int. J. Syst. Evol. Microbiol.* **59**:2870–2874.

* Corresponding author. Mailing address: School of Life Science and Biotechnology, Dalian University of Technology, Linggong Road 2, Dalian 116024, China. Phone: 86-411-84706308. Fax: 86-411-84706329. E-mail: xqzhao@dlut.edu.cn.