Draft Genome Sequence of the Streptothricin-Producing Strain *Streptomyces* sp. fd2-tb

Biao Tang,[a] Yucong Yu,[a] Xufeng Cen,[b] Yongqiang Zhu,[c] Ruixue Dai,[a] Xianwei Wang,[a] Guoping Zhao,[a,b,c] Xiaoming Ding[a]

Collaborative Innovation Center for Genetics and Development, State Key Laboratory of Genetic Engineering, Department of Microbiology, School of Life Sciences, Fudan University, Shanghai, China[a]; CAS Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China[b]; Shanghai-MOST Key Laboratory of Disease and Health Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, China[c]

B.T. and Y.Y. contributed equally to this study.

Streptomyces sp. fd2-tb can produce streptothricin class antibiotics with broad antimicrobial spectra. To better understand the mechanism of streptothricin biosynthesis and to assess the capacity of this strain in secondary metabolism, we report the draft genome sequence of *Streptomyces* sp. strain fd2-tb.

Address correspondence to Guoping Zhao, gpzhao@sibs.ac.cn, or Xiaoming Ding, xmding74@fudan.edu.cn.

Streptomyces sp. strain fd2-tb, which can produce streptothricins (STs), was isolated from a soil sample collected in Shanghai, China. STs can inhibit protein synthesis in prokaryotic cells and show antibacterial, antifungal and antiviral bioactivities ([1], [2]). STs are used to treat bacterial and fungal diseases in agriculture ([3]). The 16S rRNA sequence of *Streptomyces* sp. fd2-tb revealed that it belongs to the genus *Streptomyces* using the EzTaxon server ([4]). The voucher specimen of this actinomycete was deposited at the German Collection of Microorganisms and Cell Cultures (DSMZ) as DSM 100336.

The genome was sequenced using the Illumina HiSeq2500 platform and *de novo* assembly was performed with Velvet version 1.2.03 ([5]), resulting in 14 scaffolds with 737-fold average coverage. The draft genomic DNA sequences of the paired-end (PE) and mate-pair (MP) DNA libraries were 18,765,858 and 6,698,055 total reads, respectively. The size of the draft genome sequence is 7,646,853 bp, with an average G+C content of 72.03%, and the longest scaffold size assembled is 2,226,054 bp. A total of 7,012 open reading frames (ORFs) were predicted by the RAST server ([6]) and 76 tRNA genes were predicted by tRNA-Scan program. The coding regions constitute 86.7% of the genome.

Based on the analysis of the antiSMASH tool ([7]), 28 gene clusters were revealed for secondary metabolites biosynthesis, including 3 siderophores, 6 terpenes, 4 non-ribosomal peptide synthetase (NRPS), 1 mixed terpene/thiopeptide/lantipeptide/NRPS, 1 mixed NRPS/ladderane, 3 polyketide synthases (PKSs) (1 T1-PKS, 1 T2-PKS, 1 T3-PKS), 2 lantipeptides, 2 bacteriocins, 1 thiopeptide, 2 butyrolactones, 1 melanin, and 2 unspecified clusters. An NRPS gene cluster related to the biosynthesis of streptothricin was revealed with a length of 27,803 bp, including 21 ORFs. Further analysis of the streptothricin is under way.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in the GenBank/EMBL/DDBJ database under the accession number LBMK00000000. The version described in this paper is the first version, LBMK01000000.

REFERENCES

1. **Takemoto T, Inamori Y, Kato Y, Kubo M, Morimoto K, Morisaka K, Sakai M, Sawada Y, Taniyama H**. 1980. Physiological activity of streptothricin antibiotics. Chem Pharm Bull **28**:2884–2891. http://dx.doi.org/10.1248/cpb.28.2884.
2. **Inamori Y, Amino H, Tsuboi M, Yamaguchi S, Tsujibo H**. 1990. Biological activities of racemomycin-B, β-lysine rich streptothricin antibiotic, the main component of streptomyces lavendulae OP-2. Chem Pharm Bull **38**:2296–2298. http://dx.doi.org/10.1248/cpb.38.2296.
3. **Li J, Guo Z, Huang W, Meng X, Ai G, Tang G, Chen Y**. 2013. Mining of a streptothricin gene cluster from streptomyces sp. TP-A0356 genome via heterologous expression. Sci China Life Sci **56**:619–627. http://dx.doi.org/10.1007/s11427-013-4504-2.
4. **Chun J, Lee J-, Jung Y, Kim M, Kim S, Kim BK, Lim Y-**. 2007. EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. Int J Syst Evol Microbiol **57**:2259–2261. http://dx.doi.org/10.1099/ijs.0.64915-0.
5. **Zerbino DR, Birney E**. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res **18**:821–829. http://dx.doi.org/10.1101/gr.074492.107.
6. **Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R**. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res **42**:D206–D214. http://dx.doi.org/10.1093/nar/gkt1226.
7. **Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling 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. http://dx.doi.org/10.1093/nar/gkt449.