

Whole-Genome Shotgun Assembly and Analysis of the Genome of *Streptomyces mobaraensis* DSM 40847, a Strain for Industrial Production of Microbial Transglutaminase

Huilin Yang, Tiansheng He, Wenxia Wu, Weitao Zhu, Bai Lu, Wanju Sun

Key Laboratory of Functional Small Organic Molecule, Ministry of Education and College of Life Science, Jiangxi Normal University, Nanchang, Jiangxi, China

Here, we report the draft annotated genome sequence of *Streptomyces mobaraensis* strain DSM 40847, which is used in industry to produce microbial transglutaminase. The genome sequence will allow for the characterization of the molecular mechanisms underlying the beneficial properties of this organism.

Received 27 February 2013 Accepted 4 March 2013 Published 4 April 2013

Citation Yang H, He T, Wu W, Zhu W, Lu B, Sun W. 2013. Whole-genome shotgun assembly and analysis of the genome of *Streptomyces mobaraensis* DSM 40847, a strain for industrial production of microbial transglutaminase. *Genome Announc.* 1(2):e00143-13. doi:10.1128/genomeA.00143-13.

Copyright © 2013 Yang et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Huilin Yang, yanghl@jxnu.edu.cn.

Transglutaminases (TGases) (protein-glutamine γ -glutamyl-transferases, EC 2.3.2.13) are a family of enzymes that catalyze an acyl transfer reaction between a free amine group and a γ -carboximide group of protein-bound or peptide-bound glutamine (1). TGases derived from eukaryotes, including human blood coagulation factor XIII (2), human tissue TG (2), pig liver TG (3), and fish liver TG (4), are calcium-dependent, poly-subunit enzymes. To date, there are nine distinct TG isoenzymes in mammals that have been identified at the genomic level, of which eight are structurally and functionally related to the function of TGases, viz., TG 1 to 7 and factor XIII (5, 6). In prokaryotes, calcium-independent TGases with only one subunit have been discovered in *Streptomyces* (7) and *Bacillus* (8). Currently, transglutaminases are produced mainly by *Streptomyces mobaraensis* microbial fermentation. Although there are many published papers (9–13) reporting transglutaminase-producing strains of *S. mobaraensis*, not much is known about the genome sequences of these producing strains. We first sequenced the genome of *Streptomyces mobaraensis* DSM 40847, a strain for industrial production of microbial transglutaminase.

The genome was sequenced using the Illumina Solexa HiSeq2000 instrument at Beijing Genomics Institute (BGI) (Shenzhen, China). A library containing 500-bp inserts was constructed. Sequencing was performed with the paired-end strategy of (90, 90)-bp reads to produce 1.2 Gb of filtered sequences, representing an 85.0-fold coverage of the genome. The sequences were assembled into 266 contigs using the Velvet software (14).

Genome annotation was performed at the NCBI Prokaryotic Genomes Automatic Annotation Pipeline. Open reading frames (ORFs) were identified by Glimmer 3.02 (15) and Genemark (16). The resulting translations were used for a BLASTP (17) search against the GenBank NR database, as well as the KEGG (18) and COG (19) databases. tRNA and rRNA genes were identified by tRNAscan-SE (20) and RNAmmer (21), respectively.

The DSM 40847 chromosome is about 7.5 Mbp in length, with an average G + C content of 72.5%. A total of 6,422 protein-coding genes were identified. The genome sequence will represent a valu-

able shortcut, helping scientists to find genes. The transglutaminase gene was found, and the genes encoding the endogenous proteases TAMEP and TAP, which were used to activate protransglutaminase, were also found. The genome sequence will accelerate the progress of *Streptomyces mobaraensis* research.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [AORZ000000000](https://www.ncbi.nlm.nih.gov/nuclink/AORZ000000000). The version described in this paper is the first version, [AORZ010000000](https://www.ncbi.nlm.nih.gov/nuclink/AORZ010000000).

ACKNOWLEDGMENT

This work was supported by the Doctoral Starting Up Foundation of Jiangxi Normal University.

REFERENCES

1. Folk JE. 1980. Transglutaminases. *Annu. Rev. Biochem.* 49:517–531.
2. Chung SI. 1972. Comparative studies on tissue transglutaminase and factor XIII. *Ann. N. Y. Acad. Sci.* 202:240–255.
3. Folk JE, Cole PW, Mullooly JP. 1967. Mechanism of action of guinea pig liver transglutaminase. IV. The trimethylacyl enzyme. *J. Biol. Chem.* 242:4329–4333.
4. Yasueda H, Kumazawa Y, Motoki M. 1994. Purification and characterization of a tissue-type transglutaminase from red sea bream (*Pagrus major*). *Biosci. Biotechnol. Biochem.* 58:2041–2045.
5. Facchiano A, Facchiano F. 2009. Transglutaminases and their substrates in biology and human diseases: 50 years of growing. *Amino Acids* 36:599–614.
6. Beninati S, Piacentini M. 2004. The transglutaminase family: an overview: minireview article. *Amino Acids* 26:367–372.
7. Duran R, Junqua M, Schmitter JM, Gancet C, Goulas P. 1998. Purification, characterisation, and gene cloning of transglutaminase from *Streptovorticillum cinnamomeum* CBS 683.68. *Biochimie* 80:313–319.
8. Suzuki S, Izawa Y, Kobayashi K, Eto Y, Yamanaka S, Kubota K, Yokozeki K. 2000. Purification and characterization of novel transglutaminase from *Bacillus subtilis* spores. *Biosci. Biotechnol. Biochem.* 64:2344–2351.
9. Date M, Yokoyama K, Umezawa Y, Matsui H, Kikuchi Y. 2004. High level expression of *Streptomyces mobaraensis* transglutaminase in *Corynebacterium glutamicum* using a chimeric pro-region from *Streptomyces cinnamomeus* transglutaminase. *J. Biotechnol.* 110:219–226.
10. Liu S, Zhang D, Wang M, Cui W, Chen K, Du G, Chen J, Zhou Z. 2011. The order of expression is a key factor in the production of active trans-

- glutaminase in *Escherichia coli* by co-expression with its pro-peptide. *Microb. Cell Fact.* 10:112.
11. Marx CK, Hertel TC, Pietzsch M. 2008. Purification and activation of a recombinant histidine-tagged pro-transglutaminase after soluble expression in *Escherichia coli* and partial characterization of the active enzyme. *Enzyme Microb. Technol.* 42:568–575.
 12. Wang K, Wang B, Yang HL, Pan L. 2012. New strategy for specific activation of recombinant microbial pro-transglutaminase by introducing an enterokinase cleavage site. *Biotechnol. Lett.* 35:383–388.
 13. Zheng MY, Du GC, Guo WF, Chen J. 2001. A temperature-shift strategy in batch microbial transglutaminase fermentation. *Proc. Biochem.* 36: 525–530.
 14. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829.
 15. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23:673–679.
 16. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29:2607–2618.
 17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.
 18. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y. 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36:D480–D484.
 19. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28:33–36.
 20. 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.
 21. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.