JOURNAL OF BACTERIOLOGY, Nov. 2011, p. 6389–6390 Vol. 193, No. 22
0021-9193/11/$12.00 doi:10.1128/JB.06066-11

Complete Genome Sequence of the Industrial Strain *Bacillus megaterium* WSH-002

Liming Liu,[1,2]† Yang Li,[4,5]† Jing Zhang,[1,2] Wei Zou,[1,2] Zhemin Zhou,[4,5] Jie Liu,[3]
Xiaomin Li,[4,6] Lei Wang,[4] and Jian Chen[1,2]*

State Key Laboratory of Food Science and Technology, Jiangnan University, 1800 Lihu Road, Wuxi, Jiangsu 214122, China[1]; School of Biotechnology and Key Laboratory of Industrial Biotechnology, Ministry of Education, Jiangnan University, 1800 Lihu Road, Wuxi, Jiangsu 214122, China[2]; Jiangsu Jiangshan Pharmaceutical Co., Ltd., Jingjiang 214500, China[3]; TEDA School of Biological Sciences and Biotechnology, Nankai University, Tianjin 300457, China[4]; Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, Tianjin 300457, China[5]; and Tianjin Biochip Corporation, 23 Hongda Street, TEDA, Tianjin 300457, China[6]

Received 25 August 2011/Accepted 2 September 2011

***Bacillus megaterium*, an industrial strain, has been widely used in protein production and the vitamin C industry. Here we reported a finished, annotated, and compared 4.14-Mbp high-quality genome sequence of *B. megaterium* WSH-002, which is the companion strain for *Ketogulonicigenium vulgare* in the vitamin C industry and is stocked in our laboratory.**

Bacillus megaterium, a potential industrial strain, has exhibited many advantages in the industrial production of recombinant proteins (1, 8) and vitamins (2). In this study, *B. megaterium* WSH-002, the companion strain to improve *Ketogulonicigenium vulgare* growth and 2-keto-L-gulonic acid synthesis in the vitamin C industry (12, 16), was sequenced.

The *B. megaterium* WSH-002 genome was sequenced by a combined strategy of the Sanger shotgun approach and 454 single-end sequencing technology. A genomic library containing a 5-kb insert was constructed, and 9,604 single-end reads were generated using the Sanger shotgun method, giving 15.0-fold coverage of the genome. Using the 454 Newbler (454 Life Sciences, Branford, CT), about 97% of the 258,353 single-end reads were assembled into two large scaffolds, including 62 contigs. The relationship of the contigs was determined by multiplex PCR, and the gaps between the contigs were closed by PCR amplification, primer walking, or shotgun sequencing with an ABI 3730 sequencer. The complete genome of WSH-002 achieves an error rate of less than 1 in the range of a 10-kb sequence through sequence assembly and quality assessment by the Phred/Phrap/Consed software package (6). Protein-coding genes were predicted by combining the results obtained with Glimmer 3.02 (4) and ZCURVE (7), followed by manual inspection. Both tRNA and rRNA genes were identified by tRNAscan-SE (11) and RNAmmer (10), respectively. Functional annotation was performed by searching against an in-house protein database of the microbial genome.

The 4.14-Mb genome of *B. megaterium* WSH-002 contains four replicons, a circular chromosome (4.04 Mb) encoding 5,186 predicted open reading frames (ORFs), and three circular plasmids, named pBME_100 (0.074 Mb), pBME_200 (9,699 bp), and pBME_300 (7,006 bp), with mean GC contents of 39.1%, 36%, 32.2%, and 33.2%, respectively. There are 5,482 protein-encoding genes, 99 tRNAs, and 10 rRNA operons. Among them, 2,460 functional descriptions, 1,327 gene abbreviations, and 856 EC numbers were assigned to the WSH-002 genome by function annotation. Furthermore, we assigned 782 genes to the Kyoto Encyclopedia of Genes and Genomes metabolic pathways (13, 17). Of the functional genes, at least 98 were found to be potentially involved in spore formation and germination, such as *spo0A*, *spo0B*, *gerAB*, *gerAC*, *gerBB*, etc. (3). Moreover, 99 tRNA-coding genes and 409 genes related to heterologous protein secretion systems were identified (8, 15).

Compared with the previously published genomes of vitamin B_{12} producers *B. megaterium* DSM 319 and ATCC 12872/QMB1551 (5, 9, 14), the genome of strain WSH-002 carries most of the core genes and pathways of *B. megaterium* (5,269 ORFs). A notable characteristic of strain WSH-002 is that about 9.26% of its genes are responsible for the synthesis and secretion of proteins (508 ORFs).

In summary, the complete *B. megaterium* WSH-002 genome sequence and its curated annotation provide a better-defined genetic background for gene expression and regulation studies, especially heterologous proteins production and genome scale metabolic network construction.

Nucleotide sequence accession numbers. The sequence of the *B. megaterium* WSH-002 genome has been deposited in GenBank. The accession number for the chromosome is CP003017, and those for plasmids pBME_100, pBME_200, and pBME_300 are CP003018, CP003019, and CP003020, respectively.

---

* Corresponding author. Mailing address: State Key Laboratory of Food Science and Technology, Jiangnan University, 1800 Lihu Road, Wuxi, Jiangsu 214122, China. Phone and fax: 86-510-85197875. E-mail: jchen@jiangnan.edu.cn.

† Contributed equally to this paper.

This work was supported by the Key Program of the National Natural Science Foundation of China (20836003), the National Basic Research Program of China (973) (grants 2007CB714306 and 2009CB522603), the Program for New Century Excellent Talents in University (NCET-10-0456), and the Enterprise University Research Prospective Program, Jiangsu Province (BY2009112).

REFERENCES

1. **Biedendieck, R., M. Gamer, D. Jahn, and M. Malten.** 2005. *Bacillus megaterium*—a versatile tool for production and secretion of heterologous proteins. 12th European Congress on Biotechnology, Copenhagen, Denmark.
2. **Biedendieck, R., et al.** 2010. Metabolic engineering of cobalamin (vitamin B12) production in *Bacillus megaterium*. Microb. Biotechnol. **3:**24–37.
3. **Christie, G., and C. R. Lowe.** 2007. Role of chromosomal and plasmid-borne receptor homologues in the response of *Bacillus megaterium* QM B1551 spores to germinants. J. Bacteriol. **189:**4375–4383.
4. **Delcher, A. L., K. A. Bratke, E. C. Powers, and S. L. Salzberg.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23:**673–679.
5. **Eppinger, M., et al.** 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. J. Bacteriol. **193:**4199–4213.
6. **Gordon, D., C, Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
7. **Guo, F. B., H. Y. Ou, and C. T. Zhang.** 2003. ZCURVE: a new system for recognizing protein-coding genes in bacterial and archaeal genomes. Nucleic Acids Res. **31:**1780–1789.
8. **Hollmann, R., et al.** 2006. *Bacillus megaterium* as production system for recombinant proteins. Chem. Ing. Tech. **78:**289–294.
9. **Kunnimalaiyaan, M., and P. S. Vary.** 2005. Molecular characterization of plasmid pBM300 from *Bacillus megaterium* QM B1551. Appl. Environ. Microbiol. **71:**3068–3076.
10. **Lagesen, K., et al.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108.
11. **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.
12. **Lu, S., J. Wang, J. Yao, and Z. Yu.** 2003. Study on the effect of mutated *Bacillus megaterium* in two-stage fermentation of vitamin C. Plasma Sci. Technol. **5:**2011–2016.
13. **Okuda, S., et al.** 2008. KEGG Atlas mapping for global analysis of metabolic pathways. Nucleic Acids Res. **36:**W423–W426.
14. **Scholle, M. D., C. A. White, M. Kunnimalaiyaan, and P. S. Vary.** 2003. Sequencing and characterization of pBM400 from *Bacillus megaterium* QM B1551. Appl. Environ. Microbiol. **69:**6888–6898.
15. **Stammen, S., et al.** 2010. High-yield intra- and extracellular protein production using *Bacillus megaterium*. Appl. Environ. Microbiol. **76:**4037–4046.
16. **Zhang, J., J. Liu, Z. Shi, L. Liu, and J. Chen.** 2010. Manipulation of *B. megaterium* growth for efficient 2-KLG production by *K. vulgare*. Process Biochem. **45:**602–606.
17. **Zhou, T.-T., et al.** 2010. *MetaGen*: a promising tool for modeling metabolic networks from KEGG. Prog. Biochem. Biophys. **37:**63–68.