JOURNAL OF BACTERIOLOGY, June 2011, p. 3142–3143 Vol. 193, No. 12
0021-9193/11/$12.00 doi:10.1128/JB.00440-11

GENOME ANNOUNCEMENTS

Complete Genome Sequence of *Bacillus amyloliquefaciens* TA208, a Strain for Industrial Production of Guanosine and Ribavirin▽

Guoqiang Zhang,[1,3] Aihua Deng,[1,3] Qingyang Xu,[2] Yong Liang,[1] Ning Chen,[2]* and Tingyi Wen[1]*

Department of Industrial Microbiology and Biotechnology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China[1]; College of Biotechnology, Tianjin University of Science and Technology, Tianjin, China[2]; and Graduate University of the Chinese Academy of Sciences, Beijing, China[3]

Received 31 March 2011/Accepted 8 April 2011

Here, we report the complete genome sequence of *Bacillus amyloliquefaciens* TA208, a strain for industrial production of guanosine and synthesis of ribavirin by assimilation of formamide. Comparison of its genome sequence with those of strains DSM7 and FZB42 revealed horizontal gene transfer represented by unique prophages and restriction-modification systems and indicated significant accumulation of guanosine.

With the development of next-generation sequencing technology, microbial genome data have increased explosively (11). However, knowledge on the industrially valuable strains is not yet sufficient, especially at the genome level. Here we report the complete genome sequence of *Bacillus amyloliquefaciens* TA208, which has been used in production of guanosine and also employed in microbial synthesis of the antiviral drug ribavirin by assimilation of formamide in fermentation. Strain TA208 is an adenine auxotrophic, 8-azaguanine- and methionine sulfoxide-resistant strain generated by conventional mutagenesis (15).

The genome was sequenced using the Illumina HiSeq 2000 at Beijing Genomics Institute (BGI; Shenzhen, China). A library containing 500-bp inserts was constructed. Sequencing was performed with the pair-end strategy of 90-bp reads to produce 419.04 Mb of filtered sequences, representing a 106.4-fold coverage of the genome. The sequences were assembled into 162 contigs and 18 scaffolds using the SOAPdenovo package (8). Gaps were closed by PCR and sequencing of the amplicons by primer walking. The closed genome is a circular chromosome of 3,937,511 bp without plasmids.

Genome annotation was performed at the NCBI Prokaryotic Genomes Automatic Annotation Pipeline, where 4,135 open reading frames (ORFs), six 16S-23S-5S rRNA operons, and 71 tRNAs were identified using Glimmer (4) and Genemark (10), BLAST (13) against the Rfam database (5), and tRNAscan-SE (9), respectively. A 1,259,096-bp fragment (position 1,057,022 to 2,316,118) was found to be inversed in the chromosome of strain TA208 compared to that in strains DSM7 (14) and FZB42 (2) by the Mauve package (3), presumably due to phage infection, since two novel potential prophages flanking this region were identified using Prophage Finder (1). Moreover, six other prophage and prophage remnants were found, providing evidence for horizontal gene transfer. Preliminary analysis revealed sets of restriction-modification systems in strain TA208, including M.BamHI, M.BamHII, R.BamHI, M.H2I (7), and a novel type II system composed of the ORFs BAMTA208_19835 and BAMTA208_19845. Strain TA208 harbors a complete set of genes required for natural competence but does not readily incorporate exogenous DNA (15), which might be caused by the frameshift mutation in *comS* (6) and the complex restriction-modification systems.

The most noticeable change in the genes for purine *de novo* synthesis is the presence of a truncated copy of *purA* in strain TA208, revealing the loss of activity for adenylosuccinate synthetase. Mutations potentially elevating the corresponding activity or abolishing the end product feedback inhibition were also observed in the enzymes encoded by *purL* (D624N), *purM* (Q322K, H334Q), and *purC* (N6S, I89V, H166L) compared to the reference strains. The nonsense mutation in *pbuX* caused the absence of guanine permease in strain TA208, corresponding to the phenotype of 8-azaguanine resistance. However, most of the other genes encoding the enzymes in *de novo* purine synthesis pathway, the nucleoside efflux pumps, and their riboswitch containing promoter regions (12) remain unchanged, indicating the potential for further strain improvement.

Nucleotide sequence accession number. The complete genome sequence of *B. amyloliquefaciens* TA208 has been deposited in GenBank (accession number CP002627).

* Corresponding author. Mailing address for Tingyi Wen: Department of Industrial Microbiology and Biotechnology, Institute of Microbiology, Chinese Academy of Sciences, 1 West Beichen Road, Chaoyang District, Beijing 100101, China. Phone: 86-10-62526173. Fax: 86-10-62522397. E-mail: wenty@im.ac.cn. Mailing address for Ning Chen: College of Biotechnology, Tianjin University of Science and Technology, Tianjin Economic and Technological Development Zone, Tianjin 300457, China. Phone: 86-22-60601251. Fax: 86-22-60602198. E-mail: ningch@tust.edu.cn.

▽ Published ahead of print on 22 April 2011.

This research was funded by the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-EW-J-6) and the National Drug Discovery Program of China (2008ZX09401-05).

REFERENCES

1. **Bose, M., and R. D. Barber.** 2006. Prophage Finder: a prophage loci prediction tool for prokaryotic genome sequences. In Silico Biol. **6:**223–227.
2. **Chen, X. H., et al.** 2007. Comparative analysis of the complete genome sequence of the plant growth–promoting bacterium *Bacillus amyloliquefaciens* FZB42. Nat. Biotechnol. **25:**1007–1014.
3. **Darling, A. C. E., B. Mau, F. R. Blattner, and N. T. Perna.** 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. **14:**1394–1403.
4. **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.
5. **Griffiths-Jones, S., et al.** 2005. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res. **33:**D121–D124.
6. **Hamoen, L. W., H. Eshuis, J. Jongbloed, G. Venema, and D. V. Sinderen.** 1995. A small gene, designated *comS*, located within the coding region of the fourth amino acid activation domain of *srfA*, is required for competence development in *Bacillus subtilis*. Mol. Microbiol. **15:**55–63.
7. **Langea, C., M. Noyer-Weidnera, T. A. Trautner, M. Weinerb, and S. A. Zahlerb.** 1991. M.H2I, a multispecific 5C-DNA methyltransferase encoded by *Bacillus amyloliquefaciens* phage H2. Gene **100:**213–218.
8. **Li, R., et al.** 2010. *De novo* assembly of human genomes with massively parallel short read sequencing. Genome Res. **20:**265–272.
9. **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.
10. **Lukashin, A. V., and M. Borodovsky.** 1998. GeneMark. hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107–1115.
11. **MacLean, D., J. D. G. Jones, and D. J. Studholme.** 2009. Application of 'next-generation' sequencing technologies to microbial genetics. Nat. Rev. Microbiol. **7:**287–296.
12. **Mandal, M., B. Boese, J. E. Barrick, W. C. Winkler, and R. R. Breaker.** 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. Cell **113:**577–586.
13. **McGinnis, S., and T. L. Madden.** 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res. **32:**W20.
14. **Rückert, C., J. Blom, X. Chen, O. Reva, and R. Borriss.** Genome sequence of *B. amyloliquefaciens* type strain DSM7^T reveals differences to plant-associated *B. amyloliquefaciens* FZB42. J. Biotechnol., in press. doi:10.1016/j.jbiotec.2011.01.006,
15. **Zhang, G., et al.** 2011. Enhancing electro-transformation competency of recalcitrant *Bacillus amyloliquefaciens* by combining cell-wall weakening and cell-membrane fluidity disturbing. Anal. Biochem. **409:**130–137.