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 Xu2, Yong Liang1, Ning Chen,2* and Tingyi Wen1*

*Corresponding author:

Mailing address for Tingyi Wen: Department of Industrial Microbiology and Biotechnology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. Phone: 86-10-62526173. Fax: 86-10-62522397. E-mail: wenty@im.ac.cn.

Correspondence can also be addressed to Ning Chen: College of Biotechnology, Tianjin University of Science and Technology, Tianjin, China. Phone: 86-22-60601251. Fax: 86-22-60602198. E-mail: ningch@tust.edu.cn.
Abstract

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 the 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, the microbial genome data have increased explosively (11). However, knowledge on the industrial valuable strains is yet sufficient, especially on 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 was 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 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 ribosomal RNA 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 (1,057,022-2,316,118 bp) was found to be inversed in the chromosome of strain TA208 compared to that in the 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. *Bam*HI, M. *Bam*HIII, R. *Bam*HI, M. H2I (7), and a novel type II system composed of ORFs BAMTA208\_19835 and BAMTA208\_19845. Strain TA208 harbors a complete set of genes required for natural competence, but doesn’t 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).

**Acknowledgements**

This research was funded by the Key Project of Chinese Academy of Sciences
References


