

Complete Genome Sequence of *Bacillus thuringiensis* subsp. *kurstaki* Strain HD73

Guiming Liu,^{a,b} Lai Song,^b Changlong Shu,^a Pinshu Wang,^a Chao Deng,^a Qi Peng,^a Didier Lereclus,^c Xumin Wang,^b Dafang Huang,^a Jie Zhang,^a Fuping Song^a

State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China^a; CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China^b; INRA, UMR1319 Micalis, Génétique Microbienne et Environnement, La Minière, Guyancourt, France^c

***Bacillus thuringiensis* is a Gram-positive bacterium that produces intracellular protein crystals toxic to a wide variety of insect larvae. We report the complete genome sequence of *Bacillus thuringiensis* subsp. *kurstaki* strain HD73 from the Centre OILB (Institut Pasteur, France), which belongs to serotype 3ab and is toxic to lepidopteran larvae.**

Received 1 February 2013 Accepted 14 February 2013 Published 14 March 2013

Citation Liu G, Song L, Shu C, Wang P, Deng C, Peng Q, Lereclus D, Wang X, Huang D, Zhang J, Song F. 2013. Complete genome sequence of *Bacillus thuringiensis* subsp. *kurstaki* strain HD73. *Genome Announc.* 1(2):e00080-13. doi:10.1128/genomeA.00080-13.

Copyright © 2013 Liu 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 Jie Zhang, jzhang@ippcaas.cn, or Fuping Song, fpsong@ippcaas.cn.

Bacillus thuringiensis is a Gram-positive bacterium that produces intracellular protein crystals (cry) toxic to a wide variety of insect larvae and is the most commonly used biological pesticide (1). Many *B. thuringiensis* strains contain multiple cry genes, harbored in plasmids (2). *Bacillus thuringiensis* subspecies *kurstaki* strain HD73 (*B. thuringiensis* HD73), toxic to lepidopteran larvae (serotype 3ab) (3), contains large self-transmissible plasmids, including pHT73 and pAW63 (4). This strain was obtained from the Centre OILB (Institut Pasteur, France). This strain was provided as the *B. thuringiensis* subspecies *kurstaki* type strain and was also designated the *B. thuringiensis* subspecies *kurstaki* KT0 strain (5). Only one endotoxin gene, the *cryIAc* gene, was found to be harbored in pHT73 (6–8), and this gene was considered to be a typical example of a sporulation-dependent crystal gene because the *cryIA*-like promoter is controlled by sigma E and sigma K during sporulation (9, 10). However, weak transcription of this promoter was detected in the nonsporulating cell (11).

Here, we determined the whole genome sequence to obtain the genomic information. Genomic DNA was isolated from *B. thuringiensis* HD73. Genome sequencing was performed with 454 GS-FLX titanium (Roche Applied Science) and Illumina GAI (Illumina, United States) platforms. A total of 167,017 high-quality Roche 454 reads with an average read length of 412 bp were produced, providing about 12-fold coverage of the genome, while the Illumina reads provide 370-fold coverage with 21.7 million reads of 100 bp (with insert size 8,000 bp). After preprocessing, Roche 454 reads were assembled into contigs with Newbler, version 2.6, and then scaffolded with Illumina mate-pair reads using SSPACE (12). A 40-kb fosmid library was constructed and subjected to bidirectional end sequencing of 2,261 clones. The gaps were closed by PCR amplification and primer walking. The open reading frames (ORFs) were identified by using Glimmer version 3.02 (13). The tRNAs and rRNAs were predicted using tRNAscan-SE (14) and RNAmmer (15), respectively. The functions of encoding

genes were annotated by using NCBI nr, COG (clusters of orthologous groups) (16), KEGG (17), and InterProScan (18).

The genome of *B. thuringiensis* HD-73 contains 8 replicons, including a circular chromosome and 7 plasmids. The chromosome was 5.6 Mb in length with a GC content of 31.4%, containing 5,892 protein-encoding genes, 104 tRNA, and 36 rRNA-encoding operons. Approximately 32.1% of all coding sequences (a total of 1,894) were assigned to COGs, and 1,282 CDSs can be annotated into the 1109 KEGG orthology system by using KAAS (19). The 7-plasmid length ranges from 8 kb to 77 kb, with GC content from 29.73% to 34.66%, carrying a total of 235 ORFs. Besides the self-transmissible plasmid pHT73, our data revealed another large plasmid, pHT77, in addition to those revealed in a previous study (4), because the size similarity makes it difficult to distinguish these plasmids by gel electrophoresis (pHT77 of 76,490 bp and pHT73 of 77,351 bp). The genome sequence provides insights into plasmid conjugation, spore formation, crystal formation, virulence factor interaction with insect host, and evolution of *B. thuringiensis*.

Nucleotide sequence accession numbers. The annotated chromosome and plasmids have been deposited in GenBank under the accession numbers: CP004069 (chromosome), CP004070 (pHT73), CP004071 (pHT77), CP004073 (pHT11), CP004074 (pHT8_1), CP004075 (pHT8_2), and CP004076 (pHT7).

ACKNOWLEDGMENTS

This research was supported by grants from the Major State Basic Research Development Program of China (973 Program) (no. 2009CB118902) and the National Natural Science Foundation (no. 31070083). D.L. was supported by a grant from the MICA department of INRA.

REFERENCES

- Sanahuja G, Banakar R, Twyman RM, Capell T, Christou P. 2011. *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant Biotechnol. J.* 9:283–300.
- Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J,

- Zeigler DR, Dean DH. 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* 62:775–806.
3. Krywienczyk J, Dulmage HT, Fast PG. 1978. Occurrence of two serologically distinct groups within *Bacillus thuringiensis* serotype 3 ab var. *Kurstaki*. *J. Invertebr. Pathol.* 31:372–375.
 4. Wilcks A, Jayaswal N, Lereclus D, Andrup L. 1998. Characterization of plasmid pAW63, a second self-transmissible plasmid in *Bacillus thuringiensis* subsp. *kurstaki* HD73. *Microbiology* 144:1263–1270.
 5. Lereclus D, Menou G, Lecadet MM. 1983. Isolation of a DNA sequence related to several plasmids from *Bacillus thuringiensis* after a mating involving the *Streptococcus faecalis* plasmid pAM beta 1. *Mol. Gen. Genet.* 191:307–313.
 6. Menou G, Mahillon J, Lecadet MM, Lereclus D. 1990. Structural and genetic organization of IS232, a new insertion sequence of *Bacillus thuringiensis*. *J. Bacteriol.* 172:6689–6696.
 7. Vilas-Boas GF, Vilas-Boas LA, Lereclus D, Arantes OM. 1998. *Bacillus thuringiensis* conjugation under environmental conditions. *FEMS Microbiol. Ecol.* 25:369–374.
 8. Kronstad JW, Whiteley HR. 1984. Inverted repeat sequences flank a *Bacillus thuringiensis* crystal protein gene. *J. Bacteriol.* 160:95–102.
 9. Wong HC, Schnepf HE, Whiteley HR. 1983. Transcriptional and translational start sites for the *Bacillus thuringiensis* crystal protein gene. *J. Biol. Chem.* 258:1960–1967.
 10. Bravo A, Agaisse H, Salamitou S, Lereclus D. 1996. Analysis of *cryIAa* expression in *sigE* and *sigK* mutants of *Bacillus thuringiensis*. *Mol. Gen. Genet.* 250:734–741.
 11. Yang H, Wang P, Peng Q, Rong R, Liu C, Lereclus D, Zhang J, Song F, Huang D. 2012. Weak transcription of the *cryIAc* gene in nonsporulating *Bacillus thuringiensis* cells. *Appl. Environ. Microbiol.* 78:6466–6474.
 12. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579.
 13. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23:673–679.
 14. 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.
 15. 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.
 16. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4:41.
 17. 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.
 18. Zdobnov EM, Apweiler R. 2001. InterProScan—an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17:847–848.
 19. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35:W182–W185.