Finished Genome Sequence of *Bacillus cereus* Strain 03BB87, a Clinical Isolate with *B. anthracis* Virulence Genes

Shannon L. Johnson,[a] Timothy D. Minogue,[b] Hazuki Teshima,[a] Karen W. Davenport,[a] April A. Shea,[b] Haven L. Miner,[b] Mark J. Wolcott,[b] Patrick S. G. Chain[a]

Los Alamos National Laboratory (LANL), BioScience Division, Los Alamos, New Mexico, USA[a]; U.S. Army Medical Research Institute of Infectious Diseases (USAMRIID), Diagnostic Systems Division (DSD), Frederick, Maryland, USA[b]

Bacillus cereus strain 03BB87, a blood culture isolate, originated in a 56-year-old male muller operator with a fatal case of pneumonia in 2003. Here we present the finished genome sequence of that pathogen, including a 5.46-Mb chromosome and two plasmids (209 and 52 Kb, respectively).

Address correspondence to Shannon L. Johnson, shannonj@lanl.gov.

Bacillus cereus strain 03BB87 was isolated in 2003 from a 56-year-old male muller operator in Lubbock, Texas, USA, who perished from bacterial pneumonia (1). The strain carries a pX01 analogue plasmid (called pBCX01) that contains anthrax-associated virulence genes as well as a second, smaller plasmid. A copy of the organism was obtained from the Department of Defense Unified Culture Collection maintained at the U.S. Army Medical Institute for Infectious Diseases (USAMRIID) as BACI233.

Genomic sequencing included long-insert paired-end Roche 454 data (11× genome coverage), 100-bp Illumina standard data (283× genome coverage), and PacBio data (29× genome coverage). Each data set was assembled separately in a native assembler (Newbler, Velvet [2], or AllPaths [3]), and then the assemblies were combined using PhredPhrap (4, 5). Misassemblies were manually curated and addressed using in-house scripts and final manual corrections were made in Consed (6). The final assembly was annotated in-house using an Ergatis-based workflow manager (7). Data files are available in NCBI under accession number SRP049106.

The finished genome of *B. cereus* 03BB87 is 5,718,948 bp with a G+C content of 35.2%. The genome contains 5,788 protein-coding sequences, 107 tRNA genes, and 39 rRNA genes. A total of 74.3% of open reading frames (4,409) are protein-coding genes with function prediction. Preliminary review of the annotated genome finds resistance to tetracycline, vancomycin, and other antimicrobials.

Nucleotide sequence accession numbers. The genome is available in GenBank, with the chromosome listed as CP009941, pBCX01 as CP009940, and pBCN as CP009939.

ACKNOWLEDGMENTS

REFERENCES

1. **Hoffmaster AR, Hill KK, Gee JE, Marston CK, De BK, Popovic T, Sue D, Wilkins PP, Avashia SB, Drumgoole R, Helma CH, Ticknor LO, Okinaka RT, Jackson PJ.** 2006. Characterization of *Bacillus cereus* isolates associated with fatal pneumonias: strains are closely related to *Bacillus anthracis* and harbor *B. anthracis* virulence genes. J Clin Microbiol 44:3352–3360. http://dx.doi.org/10.1128/JCM.00561-06.
2. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. http://dx.doi.org/10.1101/gr.074492.107.
3. **Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, Nusbaum C, Jaffe DB.** 2008. ALLPATHS: de novo assembly of whole-genome shotgun microreads. Genome Res 18:810–820. http://dx.doi.org/10.1101/gr.7337908.
4. **Ewing B, Green P.** 1998. Base-calling of automated sequencer traces using Phred. II: error probabilities. Genome Res 8:186–194.
5. **Ewing B, Hillier L, Wendl MC, Green P.** 1998. Base-calling of automated sequencer traces using Phred. I: accuracy assessment. Genome Res 8:175–185. http://dx.doi.org/10.1101/gr.8.3.175.
6. **Gordon D, Green P.** 2013. *Consed*: a graphical editor for next-generation sequencing. BioInformatics 29:2936–2937. http://dx.doi.org/10.1093/bioinformatics/btt515.
7. **Hemmerich C, Buechlein A, Podicheti R, Revanna KV, Dong Q.** 2010. An Ergatis-based prokaryotic genome annotation web server. BioInformatics 26:1122–1124. http://dx.doi.org/10.1093/bioinformatics/btq090.