Complete Genome Sequence of *Bacilli bacterium* Strain VT-13-104 Isolated from the Intestine of a Patient with Duodenal Cancer

George Tetz, Victor Tetz

Institute of Human Microbiology, LLC, New York, New York, USA

We report the complete genome sequence of *Bacilli bacterium* strain VT-13-104 isolated from the intestine of a patient with duodenal cancer. The genome is composed of 3,573,421 bp, with a G+C content of 35.7%. It possesses 3,254 predicted protein-coding genes encoding multidrug resistance transporters, resistance to antibiotics, and virulence factors.

Bacteria belonging to the genus *Oceanobacillus* are motile, Gram positive, aerobic, and rod shaped. These bacteria have been isolated from deep-sea sediments, salt fields, and human feces (1–3). Sequencing of the complete 16S rRNA gene of *Bacilli bacterium* strain VT-13-104, which was isolated from the intestine of a patient with duodenal cancer, was found to possess 99% sequence identity with *Oceanobacillus caeni*, a bacterium first described in 2008 as an environmental microorganism found in wastewater (4). To date, *O. caeni* has not been found in humans.

Since biochemical and matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) analysis of *B. bacterium* strain VT-13-104 yielded divergent, low-discriminatory results, whole-genome sequencing of the bacterium was performed. The whole genome was determined using Illumina HiSeq 2500 sequencing technology (Illumina GA IIx, Illumina, CA). Library preparation, sequencing reactions, and runs were carried out according to the manufacturer's instructions. A total of 5,487,668 high-quality 125-bp paired-end reads were produced, resulting in an approximate coverage of 384×. *De novo* assembly was performed using the SPAdes assembler (version 3.5.0), which generated 80 contigs (5).

The assembled 80 contigs totaled 3,573,421 bp (max length, 393,299), with a G+C content of 35.7%. Sequence annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (6). There are 79 tRNA genes, 18 rRNA operons, 1 noncoding RNA (ncRNA) operon, and 3,254 protein coding sequences (CDSs) (934-bp average length) in the genome sequence. The analysis revealed multidrug resistance transporters of the ABC, MATE, MgtE, and MFS families, genes encoding resistance to antibiotics, including fosmidomycin, fosfomycin (*fosB*), quinolone, bacitracin (*bacA*), chloramphenicol, and the acriflavin resistance gene. The genome contains previously known virulence factors including adhesins and a RelE toxin in addition to capsular, flagellar, and sporulation proteins (7–9).

In comparison to the genome of *O. massiliensis*, the phylogenetically closest full genome sequence available, *B. bacterium* strain VT-13-104 is larger (3,573,421 bp versus 3,532,675 bp), has a lower G+C content (35.7% versus 40.35%), and fewer protein-coding genes (3,254 versus 3,519). An *in silico* DNA-DNA hybridization (DDH) analysis of the genomes of both *B. bacterium* VT-13-104 and *O. massiliensis* using the Genome-to-Genome Distance Calculator (GGDC 2.0) algorithm produced a DDH value of 42.10%, which is well below the threshold value of 70% set for genomes belonging to the same species (10, 11).

The complete genome sequence of *B. bacterium* VT-13-104 and its characterization will allow for a better understanding of its role in patients with duodenal cancer, as well as the host interaction and pathogenicity of this bacterium.

Nucleotide sequence accession number. The complete genome sequence has been deposited in the NCBI database under the accession no. LAZH00000000.

REFERENCES

1. **Lu J, Nogi Y, Takami H.** 2001. *Oceanobacillus iheyensis* gen. nov., sp. nov., a deep-sea extremely halotolerant and alkaliphilic species isolated from a depth of 1050 m on the Iheya Ridge. FEMS Microbiol Lett **205**:291–297. http://dx.doi.org/10.1111/j.1574-6968.2001.tb10963.x.
2. **Lee JS, Lim JM, Lee KC, Lee JC, Park YH, Kim CJ.** 2006. *Virgibacillus koreensis* sp. nov., a novel bacterium from a salt field, and transfer of *Virgibacillus picturae* to the genus *Oceanobacillus* as *Oceanobacillus picturae* comb. nov. with emended descriptions. Int J Syst Evol Microbiol **56**:251–257. http://dx.doi.org/10.1099/ijs.0.63734-0.
3. **Roux V, Million M, Robert C, Magne A, Raoult D.** 2013. Noncontiguous finished genome sequence and description of *Oceanobacillus massiliensis* sp. nov. Stand Genomic Sci **9**:370–384. http://dx.doi.org/10.4056/sigs.4267953.
4. **Nam JH, Bae W, Lee DH.** 2008. *Oceanobacillus caeni* sp. nov., isolated from a *Bacillus*-dominated wastewater treatment system in Korea. Int J Syst Evol Microbiol **58**:1109–1113. http://dx.doi.org/10.1099/ijs.0.65335-0.
5. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.** 2012. SPAdes: a new genome assembly algorithm and its applications to single-

cell sequencing. J Comput Biol **19:**455–477. http://dx.doi.org/10.1089/cmb.2012.0021.

6. **Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W.** 2013. Prokaryotic genome annotation pipeline. *In* Beck J, Benson D, Coleman J, Hoeppner M, Johnson M, Maglott D, Mizrachi I, Morris R, Ostell J, Pruitt K, Rubinstein W, Sayers E, Sirotkin K, Tatusova T (ed), The NCBI handbook, 2nd ed. National Center for Biotechnology Information, Bethesda, MD. http://www.ncbi.nlm.nih.gov/books/NBK143764/.

7. **Paulsen IT, Nguyen L, Sliwinski MK, Rabus R, Saier MH.** 2000. Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. J Mol Biol **301:**75–100. http://dx.doi.org/10.1006/jmbi.2000.3961.

8. **Konz D, Klens A, Schörgendorfer K, Marahiel MA.** 1997. The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. Chem Biol **4:**927–937. http://dx.doi.org/10.1016/S1074-5521(97)90301-X.

9. **Cui W, Chen L, Huang T, Gao Q, Jiang M, Zhang N, Zheng L, Feng K, Cai Y, Wang H.** 2013. Computationally identifying virulence factors based on KEGG pathways. Mol Biosyst **9:**1447–1452. http://dx.doi.org/10.1039/c3mb70024k.

10. **Auch AF, von Jan M, Klenk H-P, Göker M.** 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci **2:**117–134. http://dx.doi.org/10.4056/sigs.531120.

11. **Chun J, Rainey FA.** 2014. Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. Int J Syst Evol Microbiol **64:**316–324. http://dx.doi.org/10.1099/ijs.0.054171-0.