Brief Report

Genome Sequence of *Bacillus altitudinis* HH-03 Isolated from Dairy Pipeline

Yan-Hong Lu , Xi-Bei Fan , Yang Liu , Zhan-Hui Guo , Rui-Long Xie , Zhe-Xue Quan *

Posted Date: 3 February 2025

doi: 10.20944/preprints202502.0063.v1

Keywords: Genome sequence; Bacillus altitudinis; dairy pipeline

Brief Report

Genome Sequence of *Bacillus altitudinis* HH-03 Isolated From Dairy Pipeline

Yan-Hong Lu [1], Xi-Bei Fan [1], Ning Lv [1], Yang Liu [2], Zhan-Hui Guo [3], Rui-Long Xie [3] and Zhe-Xue Quan [1,*]

[1] School of Life Sciences, Fudan University, Shanghai, China

[2] Key Laboratory of Milk and Dairy Products Detection and Monitoring Technology, State Administration for Market Regulation, Shanghai, China

[3] Inner Mongolia Yili Industrial Group Co. Ltd., Hohhot, China

* Correspondence: quanzx@fudan.edu.cn

Abstract: In this study, *Bacillus altitudinis* HH-03 was isolated from a dairy pipeline and its genomic information was obtained by performing sequencing. The genome size of strain HH-03 was 3.85 Mbp with a GC content of 41.2%. Based on the annotation procedure, 4,010 protein-coding genes were identified. The genomic information of strain HH-03 will be helpful for the development of clean-in-place processes.

Keywords: genome sequence; *Bacillus altitudinis*; dairy pipeline

Announcement

Dairy products are susceptible to contamination by *Bacillus* during processing and production since *Bacillus* can form endospores and biofilms [1], possesses high heat resistance and dispersal abilities, and can grow and multiplies in milk pipelines [2]. A significantly elevated incidence of *Bacillus cereus* contamination was reported in unprocessed milk samples [3], and *Bacillus altitudinis* contamination was detected in pre-pasteurized milk [4].

We isolated heat-tolerant bacteria from the washed water of the milk pipeline after boiling the sample for 10 min and growing the bacteria on Nutrient Agar with manganese. Most colonies were identified as *Bacillus altitudinis* based on nearly complete 16S rRNA gene sequences. We selected strain HH-03 for further genome analysis. Genomic DNA from the pure culture of strain HH-03 was extracted using the MagAttract® HMW DNA Kit (Qiagen, USA). The sequencing library was constructed using the Transposase Enzyme Linked Long-read Sequencing (TELL-Seq™) WGS Library Prep Kit (Universal Sequencing Technology, USA) and the sequencing primers provided by the TELL-Seq™ Illumina Sequencing Primer Kit (Universal Sequencing Technology, USA), following the manufacturer's instructions [5]. The libraries were sequenced on an Illumina NovaSeq 6000 150 bp paired-end platform to obtain raw sequencing data. Quality control and splicing of the raw sequencing data of strain HH-03 were conducted using the TELL-Read and TELL-Link processes provided by Universal Sequencing Technology [5].

Following the removal of contigs exhibiting inadequate sequencing depth and contamination, the strain HH-03 yielded 12 contigs with more than 100 sequencing depth, with the largest one being 3.82 Mbp in length.

To gain insight into the genome structure, the coding sequences of the positive and negative strands of the genome, the location of non-coding RNAs, and the spliced genome were plotted in a genome feature circle diagram. The spliced FASTA files were uploaded to CGViewer [6], an online genome feature map website, for online interactive operations, and the genomes were annotated using Prokka [7]. The genomic circle diagram of strain HH-03 is shown in **Error! Reference source**

 doi:10.20944/preprints202502.0063.v1

not found.. The assembled genome size of the sample was 3,846,532 bp with a GC content of 41.2%. Additionally, the sample contained 4,010 protein-coding genes, 17 rRNA genes, and 77 tRNA genes.

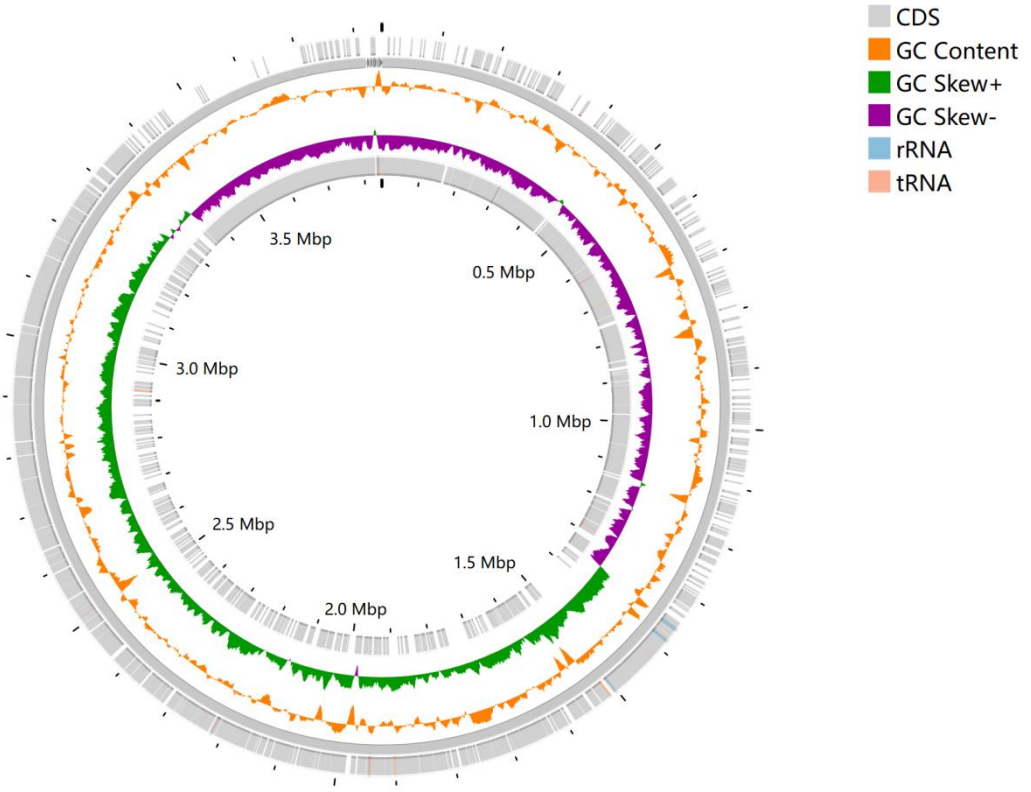

Figure 1. Circular genomic characteristics of *Bacillus altitudinis* HH-03.

Strain HH-03 was subjected to average nucleotide identity (ANI) analysis. The highest ANI value for strain HH-03 genome was 97.85% with *Bacillus altitudinis*, indicating that strain HH-03 should be identified as *Bacillus altitudinis*.

The genomic information of *Bacillus altitudinis* HH-03 will be helpful for the development of clean-in-place processes.

Bacillus altitudinis HH-03 was deposited in Guangdong Microbial Culture Collection Center (GDMCC) under accession number GDMCC 1.5556. The genome sequence of strain HH-03 from the current study have been deposited in eLMSG (an eLibrary of Microbial Systematics and Genomics, https://www.biosino.org/elmsg/index) under accession number LMSG_G000053875.1 and in NCBI under accession number JBLADY000000000.

Acknowledgments: This work was supported by Key Laboratory of Milk and Dairy Products Detection and Monitoring Technology for State Administration for Market Regulation (MDPDMT-2022-04).

References

1. Abee T, Kovács AT, and Kuipers OP, et al. Biofilm formation and dispersal in Gram-positive bacteria. 2011. *Curr Opin Biotechnol* 22(2): 172–179
2. Marchand S, De Block J, De Jonghe V, et al. Biofilm formation in milk production and processing environments; influence on milk quality and safety. 2012. *Compr Rev Food Sci Food Saf* 11: 133–147.
3. Owusu-Kwarteng J, Akabanda F, et al. Prevalence, virulence factor genes and antibiotic resistance of *Bacillus cereus* sensu lato isolated from dairy farms and traditional dairy products. 2017. *BMC Microbiol* 17(1): 65.

 doi:10.20944/preprints202502.0063.v1

4. Didouh N, Khadidja M, Campos C, et al. Assessment of biofilm, enzyme production and antibiotic susceptibility of bacteria from milk pre- and post-pasteurization pipelines in Algeria. 2023. *Int J Food Microbiol* 16(407): 110389.
5. Chen Z, Pham L, Wu TC, et al. Ultralow-input single-tube linked-read library method enables short-read second-generation sequencing systems to routinely generate highly accurate and economical long-range sequencing information. 2020. *Genome Res* 30(6): 898-909.
6. Stothard P, Grant JR, Van Domselaar G. Visualizing and comparing circular genomes using the CGView family of tools. 2019. *Briefings Bioinf* 20(4): 1576–1582.
7. Seemann T. Prokka: rapid prokaryotic genome annotation. 2014. *Bioinformatics* 30(14): 2068-2069.