# Complete Genome Sequence of *Bacillus velezensis* Strain AL7, a Biocontrol Agent for Suppression of Cotton *Verticillium* Wilt

**Haiyang Liu,**[a] **Qingchao Zeng,**[b] **Wei Wang,**[a] **Renfu Zhang,**[a] **Ju Yao**[a]

[a]Institute of Plant Protection, Xinjiang Academy of Agricultural Sciences, Urumqi, China
[b]Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, Beijing, China

**ABSTRACT**   *Bacillus velezensis* AL7, isolated from cotton soil, had strong antagonistic activity to *Verticillium dahlia* Kleb. The AL7 genome consisted of one chromosome with 3,894,709 bp (46.64% G+C content). Genome annotation predicted 3,706 protein-coding genes, 86 tRNAs, and 27 rRNAs. We sequenced and annotated the complete AL7 genome to help us better understand use of this strain.

Vascular wilts caused by members of the genus *Verticillium* are among the most devastating fungal diseases worldwide. The genus *Verticillium* consists of a relatively small group of soilborne fungi, and several of them cause wilt disease in a variety of plant hosts, including tomato, potato, and cotton (1). *Verticillium* wilt, a devastating disease of cotton caused by the soilborne fungal pathogen *Verticillium dahliae* Kleb, is a serious disease that is responsible for severe economic losses in cotton production areas throughout the world (2). Control of *Verticillium* disease is difficult due to the long persistence of the resting structures in the field and the broad host ranges of some species (1). The management of this disease includes the use of resistant cultivars, the application of chemical fungicides, and crop rotation (3). Unfortunately, these methods have their specific concerns and limitations. Although several fungicides have been reported to be effective against *Verticillium* diseases, it is not easy to control *Verticillium* diseases because fungicides are not able to affect the pathogen once it enters the xylem (4). Moreover, extensive use of fungicides in agriculture systems has raised public concerns over the environment and human health. Therefore, there is a need for new environmentally friendly technologies and products to partly or fully replace chemical-based pesticides and contribute to safer crop disease control (5). Many papers have revealed that a number of biocontrol strains, including strains of *Pseudomonas*, *Bacillus*, and *Streptomyces*, were able to significantly reduce *Verticillium* disease (3, 6, 7). Biological preparations from spore-forming *Bacillus* spp. are preferred because their long-term viability facilitates the development of commercial products (8). In this study, we report the complete genome sequence of *Bacillus velezensis* AL7, an antagonist against *Verticillium* disease, isolated from cotton soil in the city of Xinjiang, China.

A single colony was incubated and cultured overnight in Luria-Bertani broth at 37°C prior to extraction. Genomic DNA of *B. velezensis* AL7 was extracted using the phenol-chloroform method (9). The quantity of DNA was measured with a NanoDrop 2000 instrument. Then, the DNA was sent to Beijing Novogene Bioinformatics Technology Co., Ltd., and sequenced using PacBio Sequel and Illumina HiSeq 4000 systems. Library preparation was performed with a SMRTbell template kit 1.0 (Pacific Biosciences), followed by single-molecule real-time (SMRT) sequencing on the Sequel platform. Additional Illumina sequence data were generated with a HiSeq (Illumina) 350-bp paired-end run using whole-genome shotgun libraries constructed with a NEBNext Ultra DNA library prep kit following the manufacturer's recommendations. SMRTlink v5.0 (minLength, 50; minReadScore, 0.8) and readfq v10 (default parameters) software

were used to filter the low-quality reads (10, 11). After quality filtering, the Illumina sequencing generated a total of 4,033,206 paired reads (150 bp). Meanwhile, the single-molecule sequencing produced 191,057 subreads, with a mean read length of 8,791 bp and an $N_{50}$ length of 9,638 bp. The high-quality reads were assembled using SPAdes software and a hybrid approach combining Illumina and PacBio reads (12). For the assembly software, the k-mer was set as 127. Then, the scaffolds shorter than 300 bp were removed. Finally, we got one scaffold for the strain AL7 genome sequence. Annotation was performed using Prokka software (v1.11) using default parameters (13). Putative proteins were searched against the Clusters of Orthologous Groups (COG), NCBI non-redundant (nr) protein, Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases.

The complete genome sequence of strain AL7 revealed a genome size of 3,894,709 bp and a G+C content of 46.64%. The genome sequences contained 3,706 coding sequences (CDSs), 27 rRNAs, and 86 tRNAs. The predicted protein coding genes represented 88.50% of the genome sequence and had a total length of 3,446,697 bp. A total of 2,646 protein-coding genes were assigned putative functions, and 2,872 genes were categorized into COG functional groups. The average nucleotide identity was calculated using online OAT tools (https://www.ezbiocloud.net/tools/orthoani). The genome of strain AL7 was found to be closely related to that of *B. velezensis* FZB42, with an average nucleotide identity of 98.39%. *B. velezensis* AL7 can also synthesize anti-fungal antibiotics, and gene clusters for antibiotic synthesis were found. AntiSMASH online prediction software (https://antismash.secondarymetabolites.org/) using default parameters was used to show that several genes, including surfactin, iturin, and fengycin, were represented in the genome sequence of *B. velezensis* AL7.

The genome sequence and annotation of *B. velezensis* AL7 contributed to revealing the molecular mechanism of its antimicrobial activity, which suggests that *B. velezensis* AL7 could be used as a biocontrol agent of various plant diseases.

**Data availability.** The genome sequence of *B. velezensis* AL7 obtained in this whole-genome shotgun project has been deposited at NCBI under the accession number CP045926 (assembly number GCA_009663035). The version described in this paper is the first version. The raw sequencing data have been deposited in the same database under the accession numbers SRR10420793 and SRR10420792.

## REFERENCES

1. Deketelaere S, Tyvaert L, Franca SC, Hofte M. 2017. Desirable traits of a good biocontrol agent against Verticillium wilt. Front Microbiol 8:1186. https://doi.org/10.3389/fmicb.2017.01186.
2. Huang J, Li H, Yuan H. 2006. Effect of organic amendments on Verticillium wilt of cotton. Crop Prot 25:1167–1173. https://doi.org/10.1016/j.cropro.2006.02.014.
3. Li S, Zhang N, Zhang Z, Luo J, Shen B, Zhang R, Shen Q. 2013. Antagonist Bacillus subtilis HJ5 controls Verticillium wilt of cotton by root colonization and biofilm formation. Biol Fertil Soils 49:295–303. https://doi.org/10.1007/s00374-012-0718-x.
4. Fradin EF, Thomma B. 2006. Physiology and molecular aspects of Verticillium wilt diseases caused by V. dahliae and V. albo-atrum. Mol Plant Pathol 7:71–86. https://doi.org/10.1111/j.1364-3703.2006.00323.x.
5. Perez-Garcia A, Romero D, de Vicente A. 2011. Plant protection and growth stimulation by microorganisms: biotechnological applications of bacilli in agriculture. Curr Opin Biotechnol 22:187–193. https://doi.org/10.1016/j.copbio.2010.12.003.
6. Xue L, Gu M, Xu W, Lu J, Xue Q. 2016. Antagonistic Streptomyces enhances defense-related responses in cotton for biocontrol of wilt caused by phytotoxin of Verticillium dahliae. Phytoparasitica 44:225–237. https://doi.org/10.1007/s12600-016-0517-2.
7. Gómez-Lama Cabanás C, Legarda G, Ruano-Rosa D, Pizarro-Tobías P, Valverde-Corredor A, Niqui JL, Triviño JC, Roca A, Mercado-Blanco J. 2018. Indigenous Pseudomonas spp. strains from the olive (Olea europaea L.) rhizosphere as effective biocontrol agents against Verticillium dahliae: from the host roots to the bacterial genomes. Front Microbiol 9:277. https://doi.org/10.3389/fmicb.2018.00277.
8. Qiao J, Wu H, Huo R, Gao X, Borriss R. 2014. Stimulation of plant growth and biocontrol by Bacillus amyloliquefaciens subsp. plantarum FZB42 engineered for improved action. Chem Biol Technol Agric 1:12. https://doi.org/10.1186/s40538-014-0012-2.
9. Blin N, Stafford DW. 1976. A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res 3:2303–2308. https://doi.org/10.1093/nar/3.9.2303.
10. Montella S, Ventorino V, Lombard V, Henrissat B, Pepe O, Faraco V. 2017. Discovery of genes coding for carbohydrate-active enzyme by meta-

genomic analysis of lignocellulosic biomasses. Sci Rep 7:42623. https://doi.org/10.1038/srep42623.

11. Ardui S, Ameur A, Vermeesch JR, Hestand MS. 2018. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. Nucleic Acids Res 46:2159–2168. https://doi.org/10.1093/nar/gky066.

12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Son P, 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. https://doi.org/10.1089/cmb.2012.0021.

13. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.