

Complete genome sequence of *Bacillus velezensis* L194 strain isolated from soil in Tunisia

Imen BEN SLIMENE^{1*}, Manel CHAOUACHI¹, Takwa MARZOUK¹, Bilel KHIARI¹,
Amani BEN ALAYA¹, Joe WIN², Sophien KAMOUN², Naceur DJEBALI^{1**}

¹ Laboratory of Bioactive Substances, Centre of Biotechnology of Borj Cedria, BP 901,
Hammam-Lif, 2050, Tunisia.

² The Sainsbury Laboratory, University of East Anglia, Norwich Research Park, Norwich,
United Kingdom

Corresponding authors: *ibenslimen@yahoo.fr ; **dnaceur2014@gmail.com

Abstract

The whole genome shotgun sequence of a soil-borne *Bacillus velezensis* strain L194 known as a biological control agent is presented in this study. The genome sequence of this strain enhances current genomics resources of *Bacillus* species and gives insights to genes required for their mechanisms of action.

Introduction

Many microorganisms have the ability to boost plant development, and microbial products that improve plant health and growth have been developed and marketed for public consumption. Marketed microbial products are derived from one of the most described bacteria belonging to *Bacillus* genus. *Bacillus* is a vast genus of Gram-positive bacteria that belongs to the Firmicutes phylum. They are rod-shaped, endospore-forming bacteria with aerobic or facultative anaerobic metabolism. One hundred and forty-two species of *Bacillus* have been discovered, with the number continually expanding (Logan and De Vos, 2009). Bacteria in this category are quite common because they can live in a wide range of ecological niches; they may be found in soil, water, and air, as well as on plant surfaces and rhizospheres, as well as in many extreme environments (Pignatelli et al., 2009; Connor et al., 2010). Some *Bacillus* strains promote plant growth through a variety of mechanisms, including biofertilization, which increases the plant's accessibility to primary nutrients such as nitrogen, phosphate, and potassium, phytostimulation via the production of phytohormones such as indole acetic acid (IAA), auxin, and ethylene, and biocontrol via the production of antimicrobial metabolites such as lipopeptides (Yang et al.,

2009 ; Ben Slimene et al., 2012) and volatiles organic compounds (Chaouachi et al., 2021; Marzouk et al., 2021). Furthermore, *Bacillus* species may generate spores, which helps this bacterial group to thrive under harsh environments (Hashem et al., 2019).

Bacillus subtilis, *Bacillus licheniformis*, and *Bacillus pumilus* have been recognized as the “original members” of the genus *Bacillus* (Gordon et al., 1973). Fukumoto identified *B. amyloliquefaciens*, a soil-derived bacterium that generates liquefying amylase, for the first time in 1943 (Priest et al., 1987). Based on phylogenetic and phenetic data, *B. amyloliquefaciens* was later merged with the closely related *B. licheniformis*, *B. pumilus*, and *B. subtilis* as the “*B. subtilis* species complex” (Fritze, 2004). The extremely conserved character of the protein-encoding sequences in the *B. subtilis* species complex was used to make this categorization (Chen et al., 2007). For many years, it was difficult to define these closely related species using traditional taxonomic markers like as morphology, physiological features, guanine-cytosine concentration, and phylogenetic analyses utilizing 16S rRNA gene sequencing. *B. velezensis* (strains CR-502T and CR-14b) was initially isolated from environmental samples in Spain (Ruiz-García et al., 2005). The bacteria were identified as *Bacillus* species and were closely related to *Bacillus subtilis* and *Bacillus amyloliquefaciens* based on phenotypic testing and phylogenetic analysis. Further DNA–DNA hybridization tests demonstrated that the new strains had less than 20% similarity with existing *Bacillus* species, indicating that they were a separate *Bacillus* species (Ruiz-García et al., 2005). Since the sequencing of 16S rRNA gene is not sufficient to distinguish between *Bacillus* species, the use of high-throughput genomic technologies can help to simplify this difficult procedure. In addition, the incorporation of genomic data analysis will aid in the prediction of essential genes involved in metabolic pathways produced by different *Bacillus* species. In this prospect, this study presents the draft genome sequence of *B. velezensis* strain L194 isolated from soil in Tunisia. This bacterial strain exhibits different type of lipopeptides mainly iturins, surfactins and fengycins with long-chain fatty acids with a high antagonistic effect on *Phoma medicaginis* (Ben Slimene et al., 2012).

Material and Methods

Bacterial isolation and whole genome sequencing

The L194 strain was isolated from local soil at Borj Cedria situated at 25 km south of Tunis, Tunisia.

Bacterial culturing, DNA extraction, genome sequencing and assembly were conducted by microbesNG according to their protocols <https://microbesng.com/>. The reads were trimmed using Trimmomatic version 0.39 (Bolger et al., 2014). Sequence reads were assembled into contigs using Quast software version 5.0.2 (Gurevich et al., 2013). The genomes were annotated with Prokka 1.14.3 (<https://github.com/tseemann/prokka>), protein coding features and tRNA were predicted using Prodigal version 2.6 (Hyatt et al., 2010) and rRNA was predicted by ARAGORN version 1.2 (Laslett and Canback, 2004).

Results

Genome properties

The draft genome contains 3, 979,549 bp and a GC content of 46.39 %. The total number of genes identified for this strain includes 3796 protein-coding genes, 84 tRNA and a tmRNA. The results based on the ANI test and current taxonomic nomenclature revealed an identity over 99 % of the submitted genome sequence to *Bacillus* with low identity to *B. velezensis* species (Table 1).

Data availability

This whole-genome shotgun sequence assembly was deposited at DDBJ/ENA/GenBank under accession number GCA_020685825.1. Raw sequence reads were deposited in the NCBI Sequence Read Archive under BioProject number [PRJNA761700](#) and run number [SRR16156570](#).

Statement on continuing work

The datasets are being shared prior to formal publication and should be considered preliminary. We encourage and welcome feedback from the community. Please get in touch with Dr. Naceur DJEBALI for more information.

Acknowledgements

The sequencing of the bacterial strains was supported by the GetGenome initiative of The Sainsbury Laboratory, Norwich, UK, with support from the Gatsby Charitable Foundation and the Biotechnology and Biological Sciences Research Council (BBBSRC).

Table 1. Summary statistics for *Bacillus velenzensis* strain L194 genome assembled from Illumina reads.

Bacterial strain	<i>Bacillus velenzensis</i> strain L194
Source	Soil
Biosample	SAMN21972334
# contig s \geq1000 bp	27
Largest contig	869582
Total length	3974200
GC (%)	46.39
Mean coverage	49.7477
N50	409545
CDS	3796
tRNA	84
tmRNA	1
GenBank Accession (Assembly)	GCA_020685825.1
GenBank Accession (Raw reads)	SRR16156570

References

- Ben Slimene I, Tabbene O, Djebali N, Cosette P, Schmitter JM, Jouenne T, Urdaci MC, Limam F. Putative use of a *Bacillus subtilis* L194 strain for biocontrol of *Phoma medicaginis* in *Medicago truncatula* seedlings. *Res Microbiol.* 2012 Jun; 163(5):388-97.
- Bolger, A. M., Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Chaouachi, M., Marzouk, T., Jallouli, S., Elkahoui, S., Gentzbittel, L., Ben, C., Djéali, N. (2021). Activity assessment of tomato endophytic bacteria bioactive compounds for the postharvest biocontrol of *Botrytis cinerea*. *Postharvest Biology and Technology*, 172, 111389. <https://doi.org/10.1016/j.postharvbio.2020.111389>.
- Chen, X.H.; Koumoutsis, A.; Scholz, R.; Eisenreich, A.; Schneider, K.; Heinemeyer, I.; Morgenstern, B.; Voss, B.; Hess, W.R.; Reva, O.; et al. Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *Nat. Biotechnol.* 2007, 25, 1007–1014.
- Connor, N., Sikorski, J., Rooney, A.P., Kopac, S., Koeppel, A.F., Burger, A., Cole, S.G., Perry, E.B., Krizanc, D., Field, N.C., Slaton, M., Cohan, F.M., 2010. Ecology of speciation in the genus *Bacillus*. *Appl. Environ. Microbiol.* 76, 1349-1358.
- Dongjing Yang, Bo Wang, Jianxin Wang, Yu Chen, Mingguo Zhou. Activity and efficacy of *Bacillus subtilis* strain NJ-18 against rice sheath blight and *Sclerotinia* stem rot of rape, *Biological Control*. 2009, 51, 61-65
- Fritze, D. Taxonomy of the genus *Bacillus* and related genera: The aerobic endospore-forming bacteria. *Phytopathology* 2004, 94, 1245–1248.
- Gordon, R.E.; Haynes, W.C.; Pang, C.H.-N. The genus *Bacillus*; Agriculture Handbook No. 427; U.S. Department of Agriculture: Washington, DC, USA, 1973.
- Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072-1075. <https://doi.org/10.1007/978-3-030-44368-9>

Hashem, A., Tabassum, B., & Abd_Allah, E. F. (2019). *Bacillus subtilis*: A plant-growth promoting rhizobacterium that also impacts biotic stress. *Saudi journal of biological sciences*, 26(6), 1291-1297.

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, 11(1), 1-11. <https://doi.org/10.1186/1471-2105-11-119>

Laslett, D., Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic acids research*, 32(1), 11-16. <https://doi.org/10.1093/nar/gkh152>

Logan, N.A., De Vos, P., 2009. Genus *Bacillus*, in: De Vos, P., Garrity, G.M., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.H., Whitman, W.B. (Eds.), Volume Three: The Firmicutes. *Bergey's Manual of Systematic Bacteriology*, 2nd Edition. Springer, Dordrecht, Heidelberg, London, New York, pp. 21–128.

Marzouk, T., Chaouachi, M., Sharma, A., Jallouli, S., Mhamdi, R., Kaushik, N., & Djébali, N. (2021). Biocontrol of *Rhizoctonia solani* using volatile organic compounds of Solanaceae seed-borne endophytic bacteria. *Postharvest Biology and Technology*, 181, 111655. <https://doi.org/10.1016/j.postharvbio.2021.111655>.

Pignatelli, M., Moya, A., Tamames, J., 2009. EnvDB, a database for describing the environmental distribution of prokaryotic taxa. *Environ. Microbiol. Rep.* 1, 191-197.

Priest, F.G.; Goodfellow, M.; Shute, L.A.; Berkeley, R.C.W. *Bacillus amyloliquefaciens* sp. nov., nom. rev. *Int. J. Syst. Bacteriol.* 1987, 37, 69–71.

Ruiz-García, C.; Béjar, V.; Martínez-Checa, F.; Llamas, I.; Quesada, E. *Bacillus velezensis* sp. nov., a surfactant producing bacterium isolated from the river Vélez in Málaga, southern Spain. *Int. J. Syst. Evol. Microbiol.* 2005, 55, 191–195.