Draft Genome Sequence of the Industrially Significant Bacterium *Bacillus amyloliquefaciens* NRRL 942

Matthew J. Meier,[a] Annette Dodge,[a] Lee A. Beaudette[a]

aBiological Assessment and Standardization Section, Environment and Climate Change Canada, Ottawa, Ontario, Canada

ABSTRACT *Bacillus amyloliquefaciens* NRRL 942 is a Gram-positive bacterium with several potential industrial uses. We have sequenced the whole genome of this organism to assist in understanding the biological mechanisms that might modulate human health or environmental risk in the event of its release into the environment.

*B*acillus amyloliquefaciens NRRL 942 is a bacterial strain on Canada's domestic substances list (DSL; a list of substances that were manufactured in, imported into, or used in Canada on a commercial scale prior to 1986 that is available at https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/substances-list/domestic.html). This strain (also known as L-A1, originally isolated by G. A. Jeffrey, and known as 13563-0 on the DSL) normally inhabits the rhizosphere in soil, and it is known for its production of the enzyme α-amylase, which is capable of liquefying starch (1, 2). It has applications in bioaugmentation for waste treatment, septic tanks, plant disease management, and a variety of bioremediation and/or composting applications (3). Here, we present the whole-genome sequence of *Bacillus amyloliquefaciens* NRRL 942.

A glycerol stock of *B. amyloliquefaciens* strain NRRL 942 was used to inoculate a nutrient agar plate. A single colony was used to inoculate an overnight culture in nutrient broth (37°C with shaking at 220 rpm), and total genomic DNA was isolated using the Wizard genomic DNA purification kit (Promega). Genomic DNA (1 μg) was enzymatically fragmented (10 min) and size selected using the E-Gel Size Select agarose gel to select 300-bp fragments. We constructed a genomic library using the Ion Xpress Plus DNA fragment library kit (Thermo Fisher Scientific) and sequenced it on the Ion Torrent Personal Genome Machine (PGM) platform (Life Technologies).

We obtained 3,105,875 raw reads with an average length of 251 bp (782,212,297 total bases). Reads were processed using the EDGE bioinformatics platform version 2.3.0 (4). Default parameters were used for quality trimming (with the addition of adapter sequence removal of the P1 and A adapter sequences), after which 3,093,720 reads with an average length of 252 bp remained (781,672,016 total bases). Assembly was performed using SPAdes version 3.11.1 with default parameters (5), which produced 77 contigs, with an N_{50} of 875,453 bp and a total assembly size of 3,792,988 bp. The mean GC content was 40%. Mapping quality-controlled reads to the contigs using BWA-MEM with default parameters (6) in EDGE revealed an average genome coverage of 205×. Contigs were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (7), resulting in a total of 3,967 genes (3,676 coding).

Bacillus amyloliquefaciens NRRL 942 is closely related to *Bacillus amyloliquefaciens* LL3, with 98.2% of NRRL 942 reads mapping to the genome of strain LL3 (GenBank accession no. NC_017190), covering 93.6% of the bases with an average coverage of 192× and differing by 376 single-nucleotide polymorphisms (SNPs) and 79 indels. This was determined using read mapping with BWA-MEM (as described above) and SNP/indel calling with the EDGE bioinformatics platform. We identified eight anti-

microbial resistance genes in this strain using the Antibiotic Resistance Genes database (http://ardb.cbcb.umd.edu/) through EDGE. This includes five efflux pumps, namely, the ABC transporter system macrolide-lincosamide-streptogramin B efflux (Gen-Pept accession no. NP_388149), MFS antibiotic efflux (GenPept accession no. P96712), tetracycline efflux (GenPept accession no. YP_001966009), and two SMR antibiotic efflux pumps (GenPept accession no. YP_001420884 and YP_001420885). We also identified genes conferring resistance to fosfomycin (GenPept accession no. YP_001420711), penicillins, cephalosporins and carbapenems (β-lactamase Subclass B1 136), and a undecaprenyl pyrophosphate phosphatase (GenPept accession no. YP_001422387). We also found two virulence factors, *capC* (capsule biosynthesis protein, YP_080919) and *tuf* (elongation factor Tu, NP_326236). This genome sequence can be used to guide industrial users of this microorganism in establishing safe practices and to inform regulatory agencies about potential genes of interest that may be used in risk assessment.

Data availability. This whole-genome shotgun project has been deposited in GenBank under the accession no. QVEJ00000000 (BioProject accession no. PRJNA486183 and BioSample accession no. SAMN09839534). This announcement describes the first version, QVEJ01000000.

REFERENCES

1. Welker NE, Campbell LL. 1967. Comparison of the alpha-amylase of *Bacillus subtilis* and *Bacillus amyloliquefaciens*. J Bacteriol 94:1131–1135.
2. Providenti MA, Begin M, Hynes S, Lamarche C, Chitty D, Hahn J, Beaudette LA, Scroggins R, Smith ML. 2009. Identification and application of AFLP-derived genetic markers for quantitative PCR-based tracking of *Bacillus* and *Paenibacillus* spp. released in soil. Can J Microbiol 55:1166–1175. https://doi.org/10.1139/W09-071.
3. Gupta R, Beg QK, Lorenz P. 2002. Bacterial alkaline proteases: molecular approaches and industrial applications. Appl Microbiol Biotechnol 59:15–32. https://doi.org/10.1007/s00253-002-0975-y.
4. Li P-E, Lo C-C, Anderson JJ, Davenport KW, Bishop-Lilly KA, Xu Y, Ahmed S, Feng S, Mokashi VP, Chain PSG. 2017. Enabling the democratization of the genomics revolution with a fully integrated Web-based bioinformatics platform. Nucleic Acids Res 45:67–80. https://doi.org/10.1093/nar/gkw1027.
5. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20:714–737. https://doi.org/10.1089/cmb.2013.0084.
6. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
7. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.