Draft Genome Sequence of *Streptococcus thermophilus* Strain CBC-S77, Isolated from Homemade Dairy Foods in Bulgaria

Hanan R. Shehata,[a,b] Richmond A. Chandler,[c] Steven G. Newmaster[a]

aNHP Research Alliance, College of Biological Sciences, University of Guelph, Guelph, Ontario, Canada
bDepartment of Microbiology, Faculty of Pharmacy, Mansoura University, Mansoura, Egypt
cChandler Biopharmaceutical Corp., Milton, Ontario, Canada

ABSTRACT Here, we report the draft genome sequence of *Streptococcus thermophilus* strain CBC-S77. The strain was originally isolated from naturally processed, homemade dairy foods in West Rhode Mountain, Bulgaria. The genome was assembled in 148 contigs with a total length of 1,707,130 bp, with 1,563 coding genes and a GC content of 39.11%.

Streptococcus thermophilus is a species of lactic acid bacteria that is commonly used as a starter culture in the dairy industry, especially for yogurt and cheese (1–3). In addition to its role in milk fermentation, it plays an important role in enhancing the flavor and texture of yogurt (4–6). *Streptococcus thermophilus* was also reported to have various health benefits, such as antimicrobial activity (7), anti-inflammatory potential (8), a role in the alleviation of lactose intolerance (3), gastroprotective effects (9), and antioxidant and immunomodulation activities (3).

Streptococcus thermophilus strain CBC-S77 was isolated from naturally processed, homemade dairy foods from an ecologically pure and chemically unaffected geographical area (West Rhode Mountain, Bulgaria) in May 1977. *Streptococcus thermophilus* strain CBC-S77 was isolated on M17 agar (Oxoid) plates that were incubated anaerobically at 42°C for 72 h (10). Here, we present the genome sequence of *Streptococcus thermophilus* strain CBC-S77. The purpose of sequencing the genome of strain CBC-S77 is to better understand the probiotic properties and the underlying mechanisms of this strain and to confirm its safety for human consumption.

Strain CBC-S77 was obtained from Chandler Biopharmaceutical Corp. as a lyophilized powder. The NucleoSpin food kit (product number 740945.50; Macherey-Nagel, Germany) was used for genomic DNA extraction directly from 50 mg of the lyophilized powder. The extraction was performed according to the manufacturer's instructions. A Qubit v4.0 fluorometer was used for DNA quantification, and the DNA was diluted to 2 ng/μl. The DNA was then submitted to the Advanced Analysis Centre Genomics Facility at the University of Guelph (Guelph, ON, Canada) for library preparation and sequencing on an Illumina MiSeq system using an Illumina Nextera XT kit and Illumina MiSeq reagent v3 kit (600 cycles; 2 $\times$ 300-bp reads).

CLC Genomics Workbench v20.0.3 (Qiagen Bioinformatics) was used to analyze the sequencing output. Default parameters were used except where otherwise noted. First, the reads were subjected to quality trimming, in which low-quality sequences (limit of 0.05 base-calling error probability) were removed and a maximum of 2 ambiguous nucleotides were allowed. The total numbers of paired reads before and after quality trimming were 2,285,426 and 2,279,381 reads, respectively. High-quality reads with a total of 331,600,799 bases were used in *de novo* assembly, yielding ~190$\times$ coverage. The genome was assembled in 148 contigs, with a total length of 1,707,130 bp, a GC content of 39.11%, and an N_{50} value of 9,389 bp. The ContEst16S

tool (11) was used to extract the full-length 16S rRNA gene sequence from the genome sequence, and the gene sequence was used for a BLAST search with GenBank to confirm species identity (12). The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.11 (https://www.ncbi.nlm.nih.gov/genome/annotation_prok) was used for genome annotation, which predicted 1,563 protein-coding sequences and 39 tRNA genes.

Data availability. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession number JABWOT000000000. The version described in this paper is version JABWOT000000000.1. The raw files were deposited in the SRA database under accession number SRR12037890.

REFERENCES

1. Zhang J, Liu M, Xu J, Qi Y, Zhao N, Fan M. 2020. First insight into the probiotic properties of ten *Streptococcus thermophilus* strains based on in vitro conditions. Curr Microbiol 77:343–352. https://doi.org/10.1007/s00284-019-01840-3.
2. Wu Q, Tun HM, Leung FC-C, Shah NP. 2014. Genomic insights into high exopolysaccharide-producing dairy starter bacterium *Streptococcus thermophilus* ASCC 1275. Sci Rep 4:4974. https://doi.org/10.1038/srep04974.
3. Uriot O, Denis S, Junjua M, Roussel Y, Dary-Mourot A, Blanquet-Diot S. 2017. *Streptococcus thermophilus*: from yogurt starter to a new promising probiotic candidate? J Funct Foods 37:74–89. https://doi.org/10.1016/j.jff.2017.07.038.
4. Sun Z, Chen X, Wang J, Zhao W, Shao Y, Wu L, Zhou Z, Sun T, Wang L, Meng H, Zhang H, Chen W. 2011. Complete genome sequence of *Streptococcus thermophilus* strain ND03. J Bacteriol 193:793–794. https://doi.org/10.1128/JB.01374-10.
5. Linares DM, O'Callaghan TF, O'Connor PM, Ross RP, Stanton C. 2016. *Streptococcus thermophilus* APC151 strain is suitable for the manufacture of naturally GABA-enriched bioactive yogurt. Front Microbiol 7:1876. https://doi.org/10.3389/fmicb.2016.01876.
6. Cui Y, Xu T, Qu X, Hu T, Jiang X, Zhao C. 2016. New insights into various production characteristics of *Streptococcus thermophilus* strains. Int J Mol Sci 17:1701. https://doi.org/10.3390/ijms17101701.
7. Rossi F, Marzotto M, Cremonese S, Rizzotti L, Torriani S. 2013. Diversity of *Streptococcus thermophilus* in bacteriocin production: inhibitory spectrum and occurrence of thermophilin genes. Food Microbiol 35:27–33. https://doi.org/10.1016/j.fm.2013.02.006.
8. Junjua M, Kechaou N, Chain F, Awussi AA, Roussel Y, Perrin C, Roux E, Langella P, Bermúdez-Humarán LG, Le Roux Y, Chatel J-M, Dary-Mourot A. 2016. A large scale *in vitro* screening of *Streptococcus thermophilus* strains revealed strains with a high anti-inflammatory potential. LWT 70:78–87. https://doi.org/10.1016/j.lwt.2016.02.006.
9. Rodriguez C, Medici M, Rodriguez AV, Mozzi F, de Valdez GF. 2009. Prevention of chronic gastritis by fermented milks made with exopolysaccharide-producing *Streptococcus thermophilus* strains. J Dairy Sci 92:2423–2434. https://doi.org/10.3168/jds.2008-1724.
10. Michaylova M, Minkova S, Kimura K, Sasaki T, Isawa K. 2007. Isolation and characterization of *Lactobacillus delbrueckii* ssp. *bulgaricus* and *Streptococcus thermophilus* from plants in Bulgaria. FEMS Microbiol Lett 269:160–169. https://doi.org/10.1111/j.1574-6968.2007.00631.x.
11. Lee I, Chalita M, Ha SM, Na SI, Yoon SH, Chun J. 2017. ContEst16S: an algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. Int J Syst Evol Microbiol 67:2053–2057. https://doi.org/10.1099/ijsem.0.001872.
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.