

Draft Genome Sequence of Exopolysaccharide-Producing Thermophilic Bacterium *Brevibacillus thermoruber* Strain 423

Songul Yasar Yildiz,^a Margarita Kambourova,^b Kazim Yalcin Arga,^a Ebru Toksoy Oner^a

Department of Bioengineering, Marmara University, Goztepe, Istanbul, Turkey^a; Department of Extremophilic Bacteria, Institute of Microbiology, BAS, Sofia, Bulgaria^b

***Brevibacillus thermoruber* strain 423 is a Gram-positive, spore-forming, aerobic, and thermophilic bacterium that produces mannogalactoglucan exopolysaccharide (EPS). We report the draft genome sequence of *B. thermoruber* 423, which will accelerate research on the cellular organization of thermophilic bacteria, as well as the rational design and optimization of EPS production.**

Received 28 August 2013 Accepted 30 August 2013 Published 26 September 2013

Citation Yasar Yildiz S, Kambourova M, Arga KY, Toksoy Oner E. 2013. Draft genome sequence of exopolysaccharide-producing thermophilic bacterium *Brevibacillus thermoruber* strain 423. *Genome Announc.* 1(5):e00774-13. doi:10.1128/genomeA.00774-13.

Copyright © 2013 Yasar Yildiz et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Ebru Toksoy Oner, ebru.toksoy@marmara.edu.tr.

The last decade witnessed significant improvements in the discovery and development of new microbial exopolysaccharides (EPSs) from extremophiles that possess novel industrial significance (1). Among them, thermophiles constitute the least explored group, with very limited literature on them despite their fast metabolism enabling high production rates (2). Given the need for a good EPS-producing thermophile that could be used as a model organism to study the biological mechanisms and whole-genome organization of EPS-producing thermophilic bacteria, *Brevibacillus thermoruber* strain 423, isolated from a hot spring (temperature, 59°C; pH 6.5) in southwest Bulgaria and found to excrete high titers of EPS outperforming reported yields (3), was sequenced via next-generation sequencing (NGS) technology. *B. thermoruber* has been identified as a Gram-positive, motile, red-pigmented, aerobic, and thermophilic bacterium (4).

In this study, whole-genome sequencing of *B. thermoruber* 423 was performed via Illumina HiSeq 2000 technology. Paired-end sequencing was conducted with two different insert sizes of 500 and 2,000 bp. Filtering of the raw data (669 Mb) for low-quality bases, adapter contamination, and duplicates resulted in clean sequence data (603 Mb) consisting of 958,150 reads (mean length of 454 bp) and 95,666 reads (mean length of 2,446 bp), providing approximately 150-fold coverage. The reads were assembled via the SOAPdenovo software version 1.05 (5), using a *k*-mer size of 15 bp and the key parameter *K* setting at 71. The assembly process resulted in 15 scaffolds, with a total length of 4.43 Mbp, a maximum length of 783,797 bp, and an N_{50} value of 596,854 bp. BLAST (6) alignments of the assembled scaffolds indicated statistically significant genomic similarity to other *Brevibacillus* species. The RAST server version 4.0 (7) was used to predict and annotate the genes of the draft genome sequence. The SEED viewer version 2.0 (8) was used to categorize predicted genes into functional subsystems.

The resultant draft genome sequence had a mean G+C content of 58.46%, 4,446 coding sequences, and 112 RNA genes. Of the coding sequences, 42% were assigned to subsystems. Putative functions were assigned to 3,020 protein-coding genes, whereas

1,426 hypothetical proteins had no match to any known proteins. Phylogenetic analysis of the draft genome sequence revealed a rather close relationship to *B. thermoruber* strain JAM-FM2501, with 99% similarity based on 16S rRNA sequence (GenBank accession no. KF192950) comparison. However, whole-genome comparison with the genome sequences available at the RAST server indicated closest similarity to *Brevibacillus brevis* NBRC 100599 (score, 500).

The whole-genome sequencing of *B. thermoruber* 423 marks an important step toward the rational genetic and metabolic optimization of EPS production. Furthermore, the information provided by the whole-genome sequence will accelerate research on EPS production from thermophilic microorganisms via elucidation of the biological mechanisms related to EPS production, and it serves as a framework for metabolic reconstructions of thermophilic microorganisms and as an analysis platform for the study of microbial polymer production.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession no. [ATNE000000000](https://www.ncbi.nlm.nih.gov/nuccore/ATNE000000000). The version described in this paper is the first version, ATNE01000000.

ACKNOWLEDGMENT

This research has been conducted under infrastructure built via financial support by the Scientific and Technological Research Council of Turkey (TUBITAK) through grant no. MAG/110M613.

REFERENCES

1. Nicolaus B, Kambourova M, Toksoy Oner E. 2010. Exopolysaccharides from extremophiles: from fundamentals to biotechnology. *Environ. Technol.* 31:1145–1158.
2. Kambourova M, Mandeva R, Dimova D, Poli A, Nicolaus B, Tommonaro G. 2009. Production and characterization of a microbial glucan, synthesized by *Geobacillus tepidamans* V264 isolated from Bulgarian hot spring. *Carbohydr. Polym.* 77:338–343.
3. Yasar Yildiz S. 2011. Investigation of exopolysaccharide production capacity of thermophilic microorganisms isolated from Turkey. M.Sc. thesis. Marmara University, Istanbul, Turkey.
4. Manachini PL, Fortina MG, Parini C, Craveri R. 1985. *Bacillus thermo-*

- ruber* sp. nov., nom. rev., a red-pigmented thermophilic bacterium. *Int. J. Syst. Bacteriol.* 35:493–496.
5. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J. 2010. *De novo* assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20:265–272.
 6. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.
 7. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil L, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi:10.1186/1471-2164-9-75.
 8. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang H, Cohoon M, Crecy-Legard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Rückert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33:5691–570.