

Tennessee State University

Digital Scholarship @ Tennessee State University

Agricultural and Environmental Sciences
Faculty Research

Department of Agricultural and Environmental
Sciences

3-2-2017

Draft Genome Sequence of *Bacillus licheniformis* Strain YNP1-TSU Isolated from Whiterock Springs in Yellowstone National Park

Joshua A. O'Hair
Tennessee State University

Hui Li
Tennessee State University

Santosh Thapa
Tennessee State University

Matthew B. Scholz
Vanderbilt University

Suping Zhou
Tennessee State University

Follow this and additional works at: <https://digitalscholarship.tnstate.edu/agricultural-and-environmental-sciences-faculty>

Recommended Citation

O'Hair JA, Li H, Thapa S, Scholz MB, Zhou S. 2017. Draft genome sequence of *Bacillus licheniformis* strain YNP1-TSU isolated from Whiterock Springs in Yellowstone National Park. *Genome Announc* 5:e01496-16. <https://doi.org/10.1128/genomeA.01496-16>.

This Article is brought to you for free and open access by the Department of Agricultural and Environmental Sciences at Digital Scholarship @ Tennessee State University. It has been accepted for inclusion in Agricultural and Environmental Sciences Faculty Research by an authorized administrator of Digital Scholarship @ Tennessee State University. For more information, please contact XGE@Tnstate.edu.



Draft Genome Sequence of *Bacillus licheniformis* Strain YNP1-TSU Isolated from Whiterock Springs in Yellowstone National Park

Joshua A. O'Hair,^a Hui Li,^a Santosh Thapa,^a Matthew B. Scholz,^b Suping Zhou^a

Department of Agricultural and Environmental Sciences, Tennessee State University, Nashville, Tennessee, USA^a; Vanderbilt Technologies for Advanced Genetics (VANTAGE), Vanderbilt University Medical Center, Nashville, Tennessee, USA^b

ABSTRACT Novel cellulolytic microorganisms can potentially influence second-generation biofuel production. This paper reports the draft genome sequence of *Bacillus licheniformis* strain YNP1-TSU, isolated from hydrothermal-vegetative microbiomes inside Yellowstone National Park. The assembled sequence contigs predicted 4,230 coding genes, 66 tRNAs, and 10 rRNAs through automated annotation.

As greenhouse gas (GHG) emissions and global average temperatures rise, research for alternative renewable energy has been actively increasing. Bioethanol, a likely substitute, can be produced by lignocellulolytic microorganisms as they ferment waste cellulosic biomass. According to the USDA, 1 billion tons of cellulose in the form of straw, corn stover, and wood wastes could be collected annually, thus replacing approximately 30% of gasoline consumption in the United States (1). Cellulosic biomass when used as a feed source also reduces GHG emissions by 86% when compared to petroleum (2). However, during consolidated bioprocessing (CBP), excess heat can become a limiting factor to mesophilic fermenting bacteria. Thermophilic lignocellulolytic microbes, on the other hand, are able to withstand increased temperatures during CBP and carry out optimal hydrolysis at or below 47°C (3). The heated springs and hydrothermal features of Yellowstone National Park are ideal locations to find these potentially unclassified cellulolytic thermophiles.

Five individual water samples were removed from Whiterock Springs, a hydrothermal feature along Solfatarata Creek Trail (Yellowstone National Park, WY, USA) and vacuum-filtrated through 0.22- μ m filters. Strips from the five filter paper samples were cut out and placed onto Difco nutrient agar and incubated at 37°C for 24 h. Areas with substantial growth around filter papers were harvested and restreaked on additional nutrient agar at 37°C for 24 h for individual colonies. The bacterial strain reported here tested positive for extracellular endoglucanase activity on 10% carboxymethyl cellulose plates under the Congo red assay (4). After positive cellulase testing, whole-genomic DNA was extracted using the GenElute Sigma Genomic DNA kit for Gram-positive strains (Sigma, CA, USA) following the protocol described previously (4).

Libraries were prepared with Illumina TruSeq DNA Nano sample kits using indexed adaptors (Illumina). Pooled libraries were subjected to 150-bp paired-end sequencing according to the manufacturer's protocol (Illumina HiSeq3000). Bcl2fastq2 conversion software (Illumina) was used to generate demultiplexed FastQ files. This work was performed at the Vanderbilt Technologies for Advanced Genomics (VANTAGE) at Vanderbilt University (Nashville, TN, USA). Raw reads were then trimmed to remove bases of average below $Q \leq 3$ using the Burrows–Wheeler alignment method (5). *De novo* assembly was performed using SPAdes version 3.7.1 (6) with default parameters

Received 8 November 2016 Accepted 26 December 2016 Published 2 March 2017

Citation O'Hair JA, Li H, Thapa S, Scholz MB, Zhou S. 2017. Draft genome sequence of *Bacillus licheniformis* strain YNP1-TSU isolated from Whiterock Springs in Yellowstone National Park. *Genome Announc* 5:e01496-16. <https://doi.org/10.1128/genomeA.01496-16>.

Copyright © 2017 O'Hair et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Suping Zhou, zsuning@tnstate.edu.

and the “-careful” flag. The draft genome was assembled into 66 contigs with a total genome size of 4,280,972 bp (N_{50} , 253,807) and a G+C content of 45.9%. From the 66 contigs, 39 scaffolds (nine unlocalized) were constructed through a multiple-genome alignment of YNP1-TSU and its closest neighbor *Bacillus licheniformis* ATCC 14580 (obtained through MegaBLAST against the NT/NR database) with progressiveMAUVE (7). Scaffold annotation was performed with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 3.3 and yielded 4,383 genes, 4,230 of which were coding genes, followed by 66 tRNAs and 10 rRNAs. With respect to the YNP1-TSU genome, several predicted cellulolytic genes (e.g., endoglucanase, cellulobiase) contained the cellulose-binding domains CBM X2 and CBM 3 (pfam03442, pfam00942) (8), suggesting that this strain implements a cellulosome mechanism when degrading cellulose (9). Further research to examine the cellulolytic activity of individual enzymes could greatly benefit second-generation bioethanol producers.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [MIGE00000000](https://www.ncbi.nlm.nih.gov/nuclink/MIGE00000000). The version described in this paper is the first version, MIGE01000000.

ACKNOWLEDGMENTS

We thank the Yellowstone National Park System for access to sensitive hydrothermal areas under permit number YELL-2015-SCI-6074. Also a special thanks to Stacey Gunther and Sarah Haas for their help in permit-coordination and field-sampling guidance at Yellowstone National Park. The Vanderbilt VANTAGE Core provided technical assistance for this work. VANTAGE is supported in part by the CTSA (5UL1 RR024975-03), the Vanderbilt Ingram Cancer Center (P30 CA68485), the Vanderbilt Vision Center (P30 EY08126), and NIH/NCRR (G20 RR030956). A special thanks to Olivia Koues for her assistance at VANTAGE and Sarabjit Bhatti for assisting with the lab analysis at Tennessee State University.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the U.S. Department of Agriculture or the National Park Service.

REFERENCES

1. U.S. Department of Energy. 2011. U.S. billion-ton update: biomass supply for a bioenergy and bioproducts industry. ORNL/TM-2011/224. Oak Ridge National Laboratory, Oak Ridge, TN.
2. Wang M, Han J, Dunn JB, Cai H, Elgowainy A. 2012. Well-to-wheels energy use and greenhouse gas emissions of ethanol from corn, sugarcane and cellulosic biomass for US use. *Environ Res Lett* 7:045905. <https://doi.org/10.1088/1748-9326/7/4/045905>.
3. Palmqvist E, Hahn-Hägerdal B. 2000. Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresour Technol* 74:25–33. [https://doi.org/10.1016/S0960-8524\(99\)00161-3](https://doi.org/10.1016/S0960-8524(99)00161-3).
4. Li H, Zhou S, Johnson T, Vercruyse K, Ropelewski AJ, Thannhauser TW. 2014. Draft genome sequence of new *Bacillus cereus* strain tsu1. *Genome Announc* 2(6):e01294-14. <https://doi.org/10.1128/genomeA.01294-14>.
5. 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>.
6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
7. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.
8. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI’s conserved domain database. *Nucleic Acids Res* 43:D222–D226. <https://doi.org/10.1093/nar/gku1221>.
9. Gilbert HJ. 2007. Cellulosomes: microbial nanomachines that display plasticity in quaternary structure. *Mol Microbiol* 63:1568–1576. <https://doi.org/10.1111/j.1365-2958.2007.05640.x>.