

Tennessee State University

## Digital Scholarship @ Tennessee State University

---

Agricultural and Environmental Sciences  
Faculty Research

Department of Agricultural and Environmental  
Sciences

---

3-30-2017

### Draft Genome Sequences of Three Cellulolytic *Bacillus licheniformis* Strains Isolated from Imperial Geyser, Amphitheater Springs, and Whiterock Springs inside Yellowstone National Park

Joshua A. O'Hair  
*Tennessee State University*

Hui Li  
*Tennessee State University*

Santosh Thapa  
*Tennessee State University*

Matthew Scholz  
*Vanderbilt University*

Suping Zhou  
*Tennessee State University*

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



Part of the [Ecology and Evolutionary Biology Commons](#), [Genetics Commons](#), and the [Microbiology Commons](#)

---

#### Recommended Citation

O' Hair JA, Li H, Thapa S, Scholz M, Zhou S. 2017. Draft genome sequences of three cellulolytic *Bacillus licheniformis* strains isolated from Imperial Geyser, Amphitheater Springs, and Whiterock Springs inside Yellowstone National Park. *Genome Announc* 5:e00065-17. <https://doi.org/10.1128/genomeA.00065-17>.

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](mailto:XGE@Tnstate.edu).



# Draft Genome Sequences of Three Cellulolytic *Bacillus licheniformis* Strains Isolated from Imperial Geyser, Amphitheater Springs, and Whiterock Springs inside Yellowstone National Park

Joshua A. O' Hair,<sup>a</sup> Hui Li,<sup>a</sup> Santosh Thapa,<sup>a</sup> Matthew Scholz,<sup>b</sup> Suping Zhou<sup>a</sup>

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

**ABSTRACT** Novel cellulolytic microorganisms are becoming more important for rapidly growing biofuel industries. This paper reports the draft genome sequences of *Bacillus licheniformis* strains YNP2-TSU, YNP3-TSU, and YNP5-TSU. These cellulolytic isolates were collected from several hydrothermal features inside Yellowstone National Park.

As the first national park, Yellowstone has had a long-documented history of preservation, making it an ideal location to study thermophilic specimens in their natural state (1). Three heat-tolerant cellulolytic *Bacillus licheniformis* isolates were identified through genome sequencing. The first, YNP2-TSU, was removed from Amphitheater Springs (latitude [lat] 44.8016, longitude [long] -110.7288), a hydrothermal feature southeast of the Solfatar Creek trailhead. Isolate YNP3-TSU was removed from Imperial Geyser runoff (lat 44.5316, long -110.8760) approximately 3 mi east along Fairy Falls trail, and the third isolate, YNP5-TSU, was collected near the confluence of two small creeks flowing through Whiterock Springs (lat 44.7803, long -110.6981). From each sampling site, 50 ml of water was vacuum-filtrated through 0.22- $\mu$ m-pore filters. Filters were then transferred to nutrient agar, and areas with substantial growth were restreaked to produce individual colonies (2). The bacterial strains reported here tested positive for extracellular endoglucanase activity on 10% carboxymethylcellulose (CMC) under the Congo Red assay (3). After positive cellulase testing, whole-genomic DNA was extracted using GenElute Sigma genomic DNA kit for Gram-positive strains (Sigma, CA).

Libraries were prepared with 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 HiSeq 3000). The 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). Raw reads were then trimmed to remove bases of average  $Q \leq 3$  using the BWA method (4). *De novo* assembly was performed using SPAdes version 3.7.1 (5), with default parameters and the -careful flag.

The draft genome of YNP2-TSU was assembled into 96 contigs, with a total genome size of 4,765,942 bp ( $N_{50}$ , 521,225 bp) and a G+C content of 45.0%. Automated annotation was performed by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP, version 3.3) and yielded 4,822 coding sequences (CDSs), 88 tRNAs, and 16 rRNAs. The draft genome of YNP3-TSU was assembled into 53 contigs, with a total genome size of 4,518,376 bp ( $N_{50}$ , 706,022 bp) and a G+C content of 45.3%. Through the NCBI PGAP (version 3.3) 4,822 CDSs, 88 tRNAs, and 16 rRNAs were predicted. The third genome,

Received 28 January 2017 Accepted 2 February 2017 Published 30 March 2017

**Citation** O' Hair JA, Li H, Thapa S, Scholz M, Zhou S. 2017. Draft genome sequences of three cellulolytic *Bacillus licheniformis* strains isolated from Imperial Geyser, Amphitheater Springs, and Whiterock Springs inside Yellowstone National Park. *Genome Announc* 5:e00065-17. <https://doi.org/10.1128/genomeA.00065-17>.

**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](mailto:zsuning@tnstate.edu).

YNP5-TSU, was assembled into 65 contigs, with a genome size of 4,536,725 bp ( $N_{50}$ , 747,706 bp) and G+C content of 45.5%. Annotation also predicted 4,523 CDSs, 83 tRNAs, and 16 rRNAs. With respect to all three genomes, several endoglucanase, beta-cellobiosidase, exo-1,4-beta-glucosidase, and beta-xylosidase genes were predicted. Enzymes were categorized into predicted glycoside hydrolase (GH) families 1, 3, 9, 43, and 48 (UniProt) (6). GH9, GH1, and GH3 also fell into carbohydrate-binding module families X2, 3, and 6, suggesting a cellulosome hierarchy in all three strains (7). Future work to purify and test enzymatic functions will be of great importance in evolving biofuel production.

**Accession number(s).** The whole-genome shotgun projects have been deposited at DDBJ/EMBL/GenBank under the accession numbers [MEDB00000000](#) (YNP2-TSU), [MEDC00000000](#) (YNP3-TSU), and [MEDD00000000](#) (YNP5-TSU) as the first versions.

## ACKNOWLEDGMENTS

We thank the Yellowstone National Park System for access to sensitive hydrothermal areas under permit YELL-2015-SCI-6074. Also, we give 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. We give special thanks to Olivia Koues for her assistance at VANTAGE and Sarabjit Bhatti in assisting lab analysis at Tennessee State University.

VANTAGE is supported in part by CTSA grant 5UL1 RR024975-03, the Vanderbilt Ingram Cancer Center (grant P30 CA68485), the Vanderbilt Vision Center (grant P30 EY08126), and NIH/NCRR (grant G20 RR030956).

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 United States Department of Agriculture nor the National Park Service.

## REFERENCES

1. Smith DW, Peterson RO, Houston DB. 2003. Yellowstone after wolves. *J Biosci* 53:330–340. [https://doi.org/10.1641/0006-3568\(2003\)053\[0330:YAWJ\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2003)053[0330:YAWJ]2.0.CO;2).
2. O'Hair JA, Li H, Thapa S, Scholz M, Zhou S. 2017. Draft Genome of *Bacillus licheniformis* strain YNP1-TSU isolated from Whiterock Springs inside Yellowstone National Park. *Genome Announc* 5:e01496-16. <https://doi.org/10.1128/genomeA.01496-16>.
3. Li H, Zhou S, Johnson T, Vercauteren 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>.
4. 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>.
5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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>.
6. UniProt Consortium. 2015. UniProt: a hub for protein information. *Nucleic Acids Res* 43:D204–D212. <https://doi.org/10.1093/nar/gku989>.
7. 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>.