

Lawrence Berkeley National Laboratory

Lawrence Berkeley National Laboratory

Title

Complete Genome Sequence of the Marine Cellulose-and Xylan-Degrading Bacterium
Glaciecola sp. Strain 4H-3-7+YE-5

Permalink

<https://escholarship.org/uc/item/0ng9w52t>

Author

Klippel, Barbara

Publication Date

2013-01-11

DOI

10.1128/JB.05468-11

Complete Genome Sequence of the Marine Cellulose- and Xylan-Degrading Bacterium *Glaciecola* sp. Strain 4H-3-7+YE-5

Barbara Klippel¹, Adriane Lochner^{1,2}, David C. Bruce³, Karen Walston Davenport³, Chris Detter³, Lynne A. Goodwin⁴, James Han⁴, Shunsheng Han³, Loren Hauser², Miriam L. Land², Natalia Mikhailova⁴, Matt Nolan⁴, Len Pennacchio⁴, Sam Pitluck⁴, Roxanne Tapia³, Tanja Woyke⁴, Sigrid Wiebusch¹, Alexander Basner¹, Fumiyoshi Abe⁵, Koki Horikoshi⁵, Martin Keller², Garabed Antranikian^{1*}

¹Institute of Technical Microbiology, Hamburg University of Technology, Hamburg, Germany

²Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

³Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

⁴DOE Joint Genome Institute, Walnut Creek, California 94598, USA

⁵Extremobiosphere Research Center (XBR), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan

*Corresponding author. Institute of Technical Microbiology, Hamburg University of Technology, Kasernenstr. 12, D-21073 Hamburg, Germany. Phone: +49 40 42878 3117. Fax: +49 40 42878 2582. Email: antranikian@tuhh.de

Acknowledgements:

This study was funded in part by the BioEnergy Science Center, a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science under contract DOE-AC05-00OR22725. Oak Ridge National Laboratory is managed by University of Tennessee-Battelle LLC for the Department of Energy. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231.

LLNL-JRNL-501352

Complete Genome Sequence of the Marine Cellulose- and Xylan-Degrading Bacterium *Glaciecola* sp. Strain 4H-3-7+YE-5

DISCLAIMER:

[LLNL] This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

[LBNL] This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

Genome Announcement

Complete genome sequence of the marine, cellulose and xylan degrading bacterium *Glaciecola* sp. 4H-3-7+YE-5

Barbara Klippel¹, Adriane Lochner^{1,2}, David C. Bruce³, Karen Walston Davenport³, Chris Detter³, Lynne A. Goodwin⁴, James Han⁴, Shunsheng Han³, Miriam L. Land², Natalia Mikhailova⁴, Matt Nolan⁴, Len Pennacchio⁴, Sam Pitluck⁴, Roxanne Tapia³, Tanja Woyke⁴, Martin Keller², Sigrid Wiebusch¹, Alexander Basner¹, Fumiyoshi Abe⁵, Koki Horikoshi⁵, Garabed Antranikian^{1*}

¹ Institute of Technical Microbiology, Hamburg University of Technology, Hamburg, Germany

² Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

³ Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

⁴ DOE Joint Genome Institute, Walnut Creek, California 94598, USA

⁵ Extremobiosphere Research Center (XBR), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan

*Corresponding author. Institute of Technical Microbiology, Hamburg University of Technology, Kasernenstr. 12, D-21073 Hamburg, Germany. Phone: +49 40 42878 3117. Fax: +49 40 42878 2582. Email: antranikian@tuhh.de

Abstract

Glaciecola sp. 4H-3-7+YE-5 was isolated from deep sea sediments at Suruga Bay in Japan and is capable to efficiently hydrolyze cellulose and xylan. The complete genome sequence of *Glaciecola* sp. 4H-3-7+YE-5 revealed several genes encoding putatively novel glycoside hydrolases involved in plant biomass degradation.

Members of the genus *Glaciecola* are Gram-negative, aerobic and slightly halophilic bacteria (3) that can be found in various marine habitats (1, 4, 11-15, 17, 18). *Glaciecola* sp. 4H-3-7+YE-5 was isolated from ocean sediments (43.1 m below seafloor) at a water depth of 755 m at Suruga Bay (Japan) after enrichment on

cellulose, xylan and chitin as sole carbon sources. Even though only little is known about the cellulolytic and hemicellulolytic enzyme system of *Glacielecola* spp., a previous study implied the presence of cold-active, salt-tolerant biocatalysts (8).

In order to gain insight into the complete gene repertoire of *Glacielecola* sp. 4H-3-7+YE-5, the genome was sequenced at the DOE Joint genome Institute (JGI) using a combination of Illumina (2) and 454 technologies (10). To this end, we constructed and sequenced an Illumina GAii shotgun library which generated 50,060,436 reads totaling 3,804 Mb, a 454 Titanium standard library which generated 233,681 reads and three paired end 454 libraries with an average insert sizes of 10.0 kb, 5.4 kb, and 5.9 kb which generated 272,557 reads totaling 164.4 Mb of 454 data. All general aspects of library construction and sequencing can be found at <http://www.jgi.doe.gov/>. The initial draft assembly contained 55 contigs in 2 scaffolds. The 454 Titanium standard data and the 454 paired end data were assembled together with Newbler, version 2.3 while the Illumina sequencing data was assembled with VELVET, version 0.7.63 (16). Newbler and Illumina VELVET consensus data as well as read pairs in the 454 paired end library were integrated using parallel phrap, version SPS - 4.24 (High Performance Software, LLC). The software Consed (5-7) was used in the following finishing process. Illumina data was used to increase consensus quality using the software Polisher (Alla Lapidus, unpublished). Mis-assemblies were corrected using gapResolution (Cliff Han, unpublished), Dupfinisher (9), or sequencing cloned bridging PCR fragments. Gaps between contigs were closed by editing in Consed, by PCR and by Bubble PCR (J-F Cheng, unpublished) primer walks. A total of 209 additional reactions were necessary to close gaps and to raise the quality of the finished sequence. The total size of the genome is 5,393,591 bp and the final assembly is based on 137.8 Mb of 454 draft data which provides 25.6x average genome coverage and 1,774 Mb of Illumina draft

data which provides 329x average genome coverage.

The genome of *Glaciecola* sp. 4H-3-7+YE-5 is contained within one large chromosome (5,052,309 bp) and one plasmid (pGLAAG01, 341,282 bp). The complete genome has a total G+C content of 44% and comprises 4,548 predicted protein-encoding genes.

This is the first complete genome sequence for a member of the *Glaciecola* genus and analysis revealed the presence of numerous genes encoding carbohydrate active enzymes like glycoside hydrolases, glycosyl transferases and carbohydrate esterases, making the organism a promising source for biocatalysts.

Nucleotide sequence accession number. The complete chromosome and plasmid sequences of *Glaciecola* sp. 4H-3-7+YE-5 have been deposited in GenBank under accession numbers CP002526 and CP002527.

Acknowledgments

This study was funded in part by the BioEnergy Science Center, a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. Oak Ridge National Laboratory is managed by University of Tennessee-Battelle LLC for the Department of Energy under contract DOE-AC05-00OR22725. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References

1. Baik, K. S., Y. D. Park, C. N. Seong, E. M. Kim, K. S. Bae, and J. Chun. 2006. *Glaciecola nitratireducens* sp. nov., isolated from seawater. Int J Syst Evol Microbiol **56**:2185-2188.

2. Bennett, S. 2004. Solexa Ltd. *Pharmacogenomics* **5**:433-438.
3. Bowman, J., S. McCammon, J. Brown, and T. McMeekin. 1998. *Glaciecola punicea* gen. nov., sp. nov. and *Glaciecola pallidula* gen. nov., sp. nov.: psychrophilic bacteria from Antarctic sea-ice habitats. *Int J Syst Bacteriol* **48**:1213-1222.
4. Chen, L. P., H. Y. Xu, S. Z. Fu, H. X. Fan, Y. H. Liu, S. J. Liu, and Z. P. Liu. 2009. *Glaciecola lipolytica* sp. nov., isolated from seawater near Tianjin city, China. *Int J Syst Evol Microbiol* **59**:73-76.
5. Ewing, B., L. Hillier, M. C. Wendl, and P. Green. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**:175-185.
6. Ewing, B., and P. Green. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**:186-194.
7. Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. *Genome Res* **8**:195-202.
8. Guo, B., X. L. Chen, C. Y. Sun, B. C. Zhou, and Y. Z. Zhang. 2009. Gene cloning, expression and characterization of a new cold-active and salt-tolerant endo-beta-1,4-xylanase from marine *Glaciecola mesophila* KMM 241. *Appl Microbiol Biotechnol* **84**:1107-1115.
9. Han, C. S., and P. Chain (2006). Finishing repetitive regions automatically with Dupfinisher. Proceedings of the 2006 International Conference on Bioinformatics and Computational Biology. Las Vegas, NV.
10. Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. Alenquer, T. P. Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley, and J. M. Rothberg. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**:376-380.
11. Matsuyama, H., T. Hirabayashi, H. Kasahara, H. Minami, T. Hoshino, and I. Yumoto. 2006. *Glaciecola chathamensis* sp. nov., a novel marine polysaccharide-producing bacterium. *Int J Syst Evol Microbiol* **56**:2883-2886.
12. Prabakaran, S. R., R. Manorama, D. Delille, and S. Shivaji. 2007. Predominance of *Roseobacter*, *Sulfitobacter*, *Glaciecola* and *Psychrobacter* in seawater collected off Ushuaia, Argentina, Sub-Antarctica. *FEMS Microbiol Ecol* **59**:342-355.

13. Romanenko, L. A., N. V. Zhukova, M. Rohde, A. M. Lysenko, V. V. Mikhailov, and E. Stackebrandt. 2003. *Glaciecola mesophila* sp. nov., a novel marine agar-digesting bacterium. *Int J Syst Evol Microbiol* **53**:647-651.
14. Van Trappen, S., T. L. Tan, J. Yang, J. Mergaert, and J. Swings. 2004. *Glaciecola polaris* sp. nov., a novel budding and prosthecate bacterium from the Arctic Ocean, and emended description of the genus *Glaciecola*. *Int J Syst Evol Microbiol* **54**:1765-1771.
15. Yong, J. J., S. J. Park, H. J. Kim, and S. K. Rhee. 2007. *Glaciecola agarilytica* sp. nov., an agar-digesting marine bacterium from the East Sea, Korea. *Int J Syst Evol Microbiol* **57**:951-953.
16. Zerbino, D. R., and E. Birney. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**:821-829.
17. Zhang, D. C., Y. Yu, B. Chen, H. X. Wang, H. C. Liu, X. Z. Dong, and P. J. Zhou. 2006. *Glaciecola psychrophila* sp. nov., a novel psychrophilic bacterium isolated from the Arctic. *Int J Syst Evol Microbiol* **56**:2867-2869.
18. Zhang, Y. J., X. Y. Zhang, Z. H. Mi, C. X. Chen, Z. M. Gao, X. L. Chen, Y. Yu, B. Chen, and Y. Z. Zhang. 2010. *Glaciecola arctica* sp. nov., isolated from Arctic marine sediment. *Int J Syst Evol Microbiol*, in press.