Lawrence Berkeley National Laboratory

Biological Systems & Engineering

Title

High-Quality Draft Genome Sequences of Four Lignocellulose-Degrading Bacteria Isolated from Puerto Rican Forest Soil: Gordonia sp., Paenibacillus sp., Variovorax sp., and Vogesella sp.

Permalink

https://escholarship.org/uc/item/4x70p4zs

Journal

Microbiology Resource Announcements, 5(18)

ISSN

2576-098X

Authors

Woo, Hannah L
DeAngelis, Kristen M
Teshima, Hazuki
et al.

Publication Date

2017-05-04

DOI

10.1128/genomea.00300-17

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

AMERICAN SOCIETY FOR MICROBIOLOGY genome Announcements™

PROKARYOTES

High-Quality Draft Genome Sequences of Four Lignocellulose-Degrading Bacteria Isolated from Puerto Rican Forest Soil: *Gordonia* sp., *Paenibacillus* sp., *Variovorax* sp., and *Vogesella* sp.

Hannah L. Woo,[a,b] Kristen M. DeAngelis,[c] Hazuki Teshima,[d] Karen Davenport,[d] Hajnalka Daligault,[d] Tracy Erkkila,[d] Lynne Goodwin,[d] Wei Gu,[d] Chien-Chi Lo,[d] Christine Munk,[d] Matthew Scholz,[d] Yan Xu,[d] Patrick Chain,[d] David Bruce,[d] Chris Detter,[d] Roxanne Tapia,[d] Cliff Han,[d] Blake A. Simmons,[a] Terry C. Hazen[a,b,e,f,g]

Joint Bioenergy Institute, Emeryville, California, USA[a]; Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, Tennessee, USA[b]; Microbiology Department, University of Massachusetts–Amherst, Amherst, Massachusetts, USA[c]; Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA[d]; Department of Microbiology, University of Tennessee, Knoxville, Tennessee, USA[e]; Department of Earth and Planetary Sciences, University of Tennessee, Knoxville, Tennessee, USA[f]; Institute for Secure and Sustainable Environment, University of Tennessee, Knoxville, Tennessee, USA[g]

ABSTRACT Here, we report the high-quality draft genome sequences of four phylogenetically diverse lignocellulose-degrading bacteria isolated from tropical soil (*Gordonia* sp., *Paenibacillus* sp., *Variovorax* sp., and *Vogesella* sp.) to elucidate the genetic basis of their ability to degrade lignocellulose. These isolates may provide novel enzymes for biofuel production.

Received 11 March 2017 **Accepted** 15 March 2017 **Published** 4 May 2017

Citation Woo HL, DeAngelis KM, Teshima H, Davenport K, Daligault H, Erkkila T, Goodwin L, Gu W, Lo C-C, Munk C, Scholz M, Xu Y, Chain P, Bruce D, Detter C, Tapia R, Han C, Simmons BA, Hazen TC. 2017. High-quality draft genome sequences of four lignocellulose-degrading bacteria isolated from Puerto Rican forest soil: *Gordonia* sp., *Paenibacillus* sp., *Variovorax* sp., and *Vogesella* sp. Genome Announc 5:e00300-17. https://doi.org/10.1128/genomeA.00300-17.

Copyright © 2017 Woo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Terry C. Hazen, tchazen@utk.edu.

Previous studies have shown that plant litter decomposition can occur rapidly in tropical forests (1). The strains presented here were cultivated under oxic conditions from the soil of the Luqillo Experimental Forest in Puerto Rico using lignin or cellulose in minimal media agar (2). These strains were selected for genome sequencing based on their ability to degrade model carbohydrates or phenolics (2).

The genomes were sequenced by the Joint Genome Institute using Illumina sequencing technology. The short and long paired-end library preparation and assembly followed the methodology previously described by Everroad et al. (3). To raise the quality of the final sequence, PCR PacBio consensus sequences were used to close gaps. The total amount of data from the Illumina reads ranged from 3.9 to 5.7 Mb per isolate. The average G+C content was between 50 and 67%. *Variovorax* sp. had the largest estimated genome at 7.7 Mb. The genome sizes of *Gordonia* sp. and *Paenibacillus* sp. were fairly similar at about 6.9 Mb and 6.3 Mb, respectively. *Vogesella* sp. had the smallest genome size at 4.2 Mb (Table 1).

The genomes possess genes related to lignocellulose degradation. Genomes of three bacterial strains isolated on alkali lignin (*Variovorax* sp., *Gordonia* sp., and *Vogesella* sp.) possess the beta-ketoadipate pathway for aromatic catabolism of lignin monomers and other phenolics into tricarboxylic acid cycle intermediates (4). *Variovorax* sp. and *Gordonia* sp. have multiple dioxygenases to metabolize two different aromatic catabolism intermediates (protocatechuate and catechol), while *Vogesella* sp. only has genes for protocatechuate degradation. *Paenibacillus* sp., the strain isolated on carboxymethyl cellulose in minimal media, possesses five different endo-1,4-betaxylanses. One or more of these xylanases could be highly active, as

TABLE 1 Metadata of Puerto Rican soil strains and sequencing run details

Species	Isolation carbon source	Strain	Phylum; order	Genome size (Mb)	G+C content (%)	Accession no.	Sequencing library paired-end reads (short/long)	Total reads (Mb)	Coverage
Gordonia sp.	Alkali lignin	HW436	*Actinomycetes*; *Actinomycetales*	6.3	67	ARVZ00000000	14,821,334/19,192,428	4,142	657×
Paenibacillus sp.	Carboxymethyl cellulose	HW567	*Firmicutes*; *Bacillales*	6.9	50	ARFI00000000	14,326,726/18,199,996	4,879	707×
Vogesella sp.[a]	Alkali lignin	LIG4	*Proteobacteria*; *Neisseriales*	4.2	64	LT607802	18,542,852/11,702,786	3,952	941×
Variovorax sp.	Alkali lignin	HW608	*Proteobacteria*; *Burkholderiales*	7.7	67	LT607803	14,536,906/35,495,686	5,730	744×

[a] Originally submitted as *Pseudogulbankiania* sp. LIG4.

Paenibacillus sp. grows well on cellulose and xylan agar and degrades beta-D-glucopyranoside at notable rates.

Variovorax sp. and *Vogesella* sp. may also contribute to nitrogen cycling. *Variovorax* sp. has nitrogenases (*nifK*, *nifD*, and *nifH*) that are related to nitrogen fixation. *Vogesella* sp. has genes encoding respiratory nitrate reductase (alpha, beta, and gamma subunits) that are related to dissimilatory nitrate reduction.

All four genomes are part of an ongoing investigation of the genetic basis of lignocellulose degradation in tropical soils. These genomes will be compared to other genomes of lignocellulose-degrading bacteria from tropical forest environments, such as *Enterobacter lignolyticus* SCF1 (5), *Klebsiella* sp. BRL6-2 (6), and *Burkholderia* sp. LIG30 (7). The discovery of genes encoding lignocellulose-degrading enzymes would benefit biofuel production, for which lignocellulosic biomasses must be rapidly deconstructed and saccharified using enzymes.

Accession number(s). The whole-genome sequences reported here were deposited in DDBJ/EMBL/GenBank under the accession numbers listed in Table 1. *Vogesella* sp. LIG4 was originally submitted in 2012 as *Pseudogulbankiania* sp. LIG4, another *Neisserales* species, but has since been determined to be more closely related to other *Vogesella* sp. strains by 16S rRNA gene analyses with BLASTn and average nucleotide identity by BLAST (ANIb).

ACKNOWLEDGMENTS

Work conducted by the Joint Bioenergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under contract number DE-AC02-05CH11231. This material is also based upon work supported by the National Science Foundation Graduate Research Fellowship Program under grant number DGE-1452154.

REFERENCES

1. Parton W, Silver WL, Burke IC, Grassens L, Harmon ME, Currie WS, King JY, Adair EC, Brandt LA, Hart SC, Fasth B. 2007. Global-scale similarities in nitrogen release patterns during long-term decomposition. Science 315:361–364. https://doi.org/10.1126/science.1134853.
2. Woo HL, Hazen TC, Simmons BA, DeAngelis KM. 2014. Enzyme activities of aerobic lignocellulolytic bacteria isolated from wet tropical forest soils. Syst Appl Microbiol 37:60–67. https://doi.org/10.1016/j.syapm.2013.10.001.
3. Everroad RC, Woebken D, Singer SW, Burow LC, Kyrpides N, Woyke T, Goodwin L, Detweiler A, Prufert-Bebout L, Pett-Ridge J. 2013. Draft genome sequence of an oscillatorian cyanobacterium, strain ESFC-1. Genome Announc 1(4):e00527-13. https://doi.org/10.1128/genomeA.00527-13.
4. Harwood CS, Parales RE. 1996. The β-ketoadipate pathway and the biology of self-identity. Annu Rev Microbiol 50:553–590. https://doi.org/10.1146/annurev.micro.50.1.553.
5. DeAngelis KM, D'Haeseleer P, Chivian D, Fortney JL, Khudyakov J, Simmons B, Woo H, Arkin AP, Davenport KW, Goodwin L, Chen A, Ivanova N, Kyrpides NC, Mavromatis K, Woyke T, Hazen TC. 2011. Complete genome sequence of "*Enterobacter lignolyticus*" SCF1. Stand Genomic Sci 5:69–85. https://doi.org/10.4056/sigs.2104875.
6. Woo HL, Ballor NR, Hazen TC, Fortney JL, Simmons B, Davenport KW, Goodwin L, Ivanova N, Kyrpides NC, Mavromatis K, Woyke T, Jansson J, Kimbrel J, DeAngelis KM. 2014. Complete genome sequence of the lignin-degrading bacterium *Klebsiella* sp. strain BRL6-2. Stand Genomic Sci 9:19. https://doi.org/10.1186/1944-3277-9-19.
7. Woo HL, Utturkar S, Klingeman D, Simmons BA, DeAngelis KM, Brown SD, Hazen TC. 2014. Draft genome sequence of the lignin-degrading *Burkholderia* sp. strain LIG30, isolated from wet tropical forest soil. Genome Announc 2(3):e00637-14. https://doi.org/10.1128/genomeA.00637-14.