

# Complete Genome Sequence of *Herbinix luporum* SD1D, a New Cellulose-Degrading Bacterium Isolated from a Thermophilic Biogas Reactor

Daniela E. Koeck,<sup>a</sup> Irena Maus,<sup>c</sup> Daniel Wibberg,<sup>c</sup> Anika Winkler,<sup>c</sup> Vladimir V. Zverlov,<sup>a,b</sup> Wolfgang Liebl,<sup>a</sup> Alfred Pühler,<sup>c</sup> Wolfgang H. Schwarz,<sup>a</sup> Andreas Schlüter<sup>c</sup>

Institute for Microbiology, Technische Universität München, Freising-Weihenstephan, Germany<sup>a</sup>; Institute of Molecular Genetics, Russian Academy of Science, Moscow, Russia<sup>b</sup>; Center for Biotechnology (CeBiTec), Genome Research of Industrial Microorganisms, Bielefeld University, Bielefeld, Germany<sup>c</sup>

D.E.K. and I.M. contributed equally to this article.

**A novel cellulolytic bacterial strain was isolated from an industrial-scale biogas plant. The 16S rRNA gene sequence of the strain SD1D showed 96.4% similarity to *Herbinix hemicellulosilytica* T3/55<sup>T</sup>, indicating a novel species within the genus *Herbinix* (family *Lachnospiraceae*). Here, the complete genome sequence of *Herbinix luporum* SD1D is reported.**

Received 25 May 2016 Accepted 30 May 2016 Published 21 July 2016

**Citation** Koeck DE, Maus I, Wibberg D, Winkler A, Zverlov VV, Liebl W, Pühler A, Schwarz WH, Schlüter A. 2016. Complete genome sequence of *Herbinix luporum* SD1D, a new cellulose-degrading bacterium isolated from a thermophilic biogas reactor. *Genome Announc* 4(4):e00687-16. doi:10.1128/genomeA.00687-16.

**Copyright** © 2016 Koeck 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 Andreas Schlüter, [aschluet@cebitec.uni-bielefeld.de](mailto:aschluet@cebitec.uni-bielefeld.de).

The isolation, characterization, and genome sequencing of new strains are pivotal to link their genetic potential to possible ecological roles in biomass digestion and biogas production (1). Cellulolytic isolates, like those of *Clostridium thermocellum*, *Clostridium cellulosi*, and *Herbinix hemicellulosilytica*, are a promising source for novel biomass-degrading enzymes that convert biomass into sugars, the basis for the production of biofuels, bulk chemicals, and other valuable products (2–4).

The new cellulose-degrading strain SD1D was isolated from a thermophilic biogas plant, as described previously (5). A nucleotide sequence comparison between the isolate SD1D and *H. hemicellulosilytica* T3/55<sup>T</sup> revealed a 16S rRNA sequence similarity of 96.4%, as well as an average nucleotide identity (ANI) of 85.73%, suggesting strain SD1D as a new species within the genus *Herbinix* (family *Lachnospiraceae*) (6, 7).

*Herbinix luporum* SD1D is able to digest cellulosic and hemicellulosic substrates (filter paper, phosphoric acid swollen cellulose, carboxymethyl cellulose, xylan, and barley). The highest activity was determined to occur on xylan, but in contrast to strain T3/55<sup>T</sup>, SD1D was not able to grow on xylan or xylose. Galactose, glucose, mannose, arabinose, cellobiose, and cellulose are utilized as sole sources of carbon. The major fermentation products are ethanol, acetic acid, and small amounts of propionic acid.

The DNA of *H. luporum* SD1D was extracted and used to construct an 8-kb mate-pair sequencing library (Nextera mate-pair sample preparation kit; Illumina, Inc.), which was sequenced on an Illumina MiSeq system. The sequencing run yielded 1,766,358 reads, accounting for 357,174,077 bases of total sequence information. The obtained sequences were *de novo* assembled (8) using the GS *De Novo* Assembler software (version 2.8; Roche) and resulted in one scaffold comprising 19 contigs. After an *in silico* gap closure approach, a circular chromosome of 2,609,352 bp in size, featuring a G + C content of 35.25%, was established. The software

platform GenDB (9) was applied to annotate the *H. luporum* SD1D genome, leading to the identification of 2,362 protein-coding sequences, 53 tRNA genes, and four *rrn* operons.

The SD1D genome was searched for the presence of carbohydrate-active enzymes by applying the Hidden-Markov-Model (HMM)-based carbohydrate-active enzyme annotation database dbCAN (10). In total, 174 genes encode enzymes or modules with predicted activity on carbohydrates that mainly belong to different families of glycoside hydrolases (GH). Among them, four putative cellulosomal genes were identified encoding GH family modules (2 × GH9, GH48, and GH11), each attached to a dockerin module. A corresponding scaffoldin (*cipA*) comprising five type 1 cohesin and four carbohydrate-binding modules (CBM\_X2) is encoded in the genome of SD1D. In addition, five potential cellulases of the GH family 5 were identified. Except for one endoglucanase gene, all other cellulosomal genes are arranged in one cluster. In contrast, *H. hemicellulosilytica* T3/55<sup>T</sup> does not possess any putative cellulosomal genes.

The availability of the *H. luporum* SD1D genome sequence provides the genetic basis for the biotechnological exploitation of genome features involved in the thermophilic degradation of lignocellulosic biomass.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited in the EMBL/GenBank database (EBI, NCBI) under the accession no. [LN879430](https://www.ncbi.nlm.nih.gov/nuccore/LN879430). The strain is available from the Leibniz Institute German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany) under the accession no. DSM 100831.

## ACKNOWLEDGMENTS

The bioinformatics support of the BMBF-funded project “Bielefeld-Gießen Center for Microbial Bioinformatics”-BiGi within the German Network for Bioinformatics Infrastructure (deNBI.de) is gratefully ac-

knowledge. I.M. and D.W. acknowledge the receipt of a scholarship from the CLIB Graduate Cluster Industrial Biotechnology, cofinanced by the Ministry of Innovation of North Rhine-Westphalia.

## FUNDING INFORMATION

This work, including the efforts of Alfred Pühler, was funded by Bundesministerium für Bildung und Forschung (BMBF) (031A533). This work, including the efforts of Wolfgang H. Schwarz, was funded by Bundesministerium für Ernährung und Landwirtschaft (BMEL) (22007012). This work, including the efforts of Andreas Schlüter, was funded by Bundesministerium für Ernährung und Landwirtschaft (BMEL) (22006712 and 22404015).

## REFERENCES

- Biddle A, Stewart L, Blanchard J, Leschine S. 2013. Untangling the genetic basis of fibrolytic specialization by *Lachnospiraceae* and *Ruminococcaceae* in diverse gut communities. *Diversity* 5:627–640. <http://dx.doi.org/10.3390/d5030627>.
- Koeck DE, Maus I, Wibberg D, Winkler A, Zverlov VV, Liebl W, Pühler A, Schwarz WH, Schlüter A. 2015. Draft genome sequence of *Herbinix hemicellulosilytica* T3/55, a new thermophilic cellulose degrading bacterium isolated from a thermophilic biogas reactor. *J Biotechnol* 214:59–60. <http://dx.doi.org/10.1016/j.jbiotec.2015.07.022>.
- Koeck DE, Wibberg D, Koellmeier T, Blom J, Jaenicke S, Winkler A, Albersmeier A, Zverlov VV, Pühler A, Schwarz WH, Schlüter A. 2013. Draft genome sequence of the cellulolytic *Clostridium thermocellum* wild-type strain DG5 playing a role in cellulosic biomass degradation. *J Biotechnol* 168:62–63. <http://dx.doi.org/10.1016/j.jbiotec.2013.08.011>.
- Koeck DE, Wibberg D, Maus I, Winkler A, Albersmeier A, Zverlov VV, Liebl W, Pühler A, Schwarz WH, Schlüter A. 2014. Complete genome sequence of the cellulolytic thermophile *Ruminoclostridium cellulosi* wild-type strain DG5 isolated from a thermophilic biogas plant. *J Biotechnol* 188:136–137. <http://dx.doi.org/10.1016/j.jbiotec.2014.08.024>.
- Koeck DE, Ludwig W, Wanner G, Zverlov VV, Liebl W, Schwarz WH. 2015. *Herbinix hemicellulosilytica*, gen. nov., sp. nov., a thermophilic cellulose-degrading bacterium isolated from a thermophilic biogas reactor. *Int J Syst Evol Microbiol* 65:2365–2371.
- Kittlmann S, Seedorf H, Walters WA, Clemente JC, Knight R, Gordon JI, Janssen PH. 2013. Simultaneous amplicon sequencing to explore co-occurrence patterns of bacterial, archaeal and eukaryotic microorganisms in rumen microbial communities. *PLoS One* 8:e47879. <http://dx.doi.org/10.1371/journal.pone.0047879>.
- Meehan CJ, Beiko RG. 2014. A phylogenomic view of ecological specialization in the *Lachnospiraceae*, a family of digestive tract-associated bacteria. *Genome Biol Evol* 6:703–713. <http://dx.doi.org/10.1093/gbe/evu050>.
- Wibberg D, Blom J, Jaenicke S, Kollin F, Rupp O, Scharf B, Schneiker-Bekel S, Sczcepanowski R, Goesmann A, Setubal JC, Schmitt R, Pühler A, Schlüter A. 2011. Complete genome sequencing of *Agrobacterium* sp. H13-3, the former *Rhizobium lupini* H13-3, reveals a tripartite genome consisting of a circular and a linear chromosome and an accessory plasmid but lacking a tumor-inducing Ti-plasmid. *J Biotechnol* 155:50–62. <http://dx.doi.org/10.1016/j.jbiotec.2011.01.010>.
- Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, Pühler A. 2003. GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res* 31:2187–2195. <http://dx.doi.org/10.1093/nar/gkg312>.
- Yin YB, Mao XZ, Yang JC, Chen X, Mao FL, Xu Y. 2012. dbCAN: a Web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40:W445–W451. <http://dx.doi.org/10.1093/nar/gks479>.