

Draft Genome Sequences of *Geobacillus* sp. Strains CAMR5420 and CAMR12739

Pieter De Maayer,^a Carolyn E. Williamson,^b Christopher T. Vennard,^b Michael J. Danson,^b Don A. Cowan^a

Centre for Microbial Ecology and Genomics, Genomics Research Institute, Department of Genetics, University of Pretoria, Pretoria, South Africa^a; Centre for Extremophile Research, Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom^b

Thermophilic *Geobacillus* spp. can efficiently hydrolyze hemicellulose polymers and are therefore of interest in biotechnological applications. Here we report the genome sequences of two hemicellulolytic strains, *Geobacillus* sp. CAMR12739 and CAMR5420.

Received 19 May 2014 Accepted 20 May 2014 Published 5 June 2014

Citation De Maayer P, Williamson CE, Vennard CT, Danson MJ, Cowan DA. 2014. Draft genome sequences of *Geobacillus* sp. strains CAMR5420 and CAMR12739. *Genome Announc.* 2(3):e00567-14. doi:10.1128/genomeA.00567-14.

Copyright © 2014 De Maayer et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Don A. Cowan, don.cowan@up.ac.za.

Members of the genus *Geobacillus* are Gram-positive, thermophilic, spore-forming, aerobic bacteria (1). They have an extensive capacity to degrade plant cell wall hemicellulose polymers into their component pentose sugars, a capacity that has been linked to a single chromosomal locus (2, 3). Coupled with this is their production of thermostable hydrolytic enzymes, resulting in considerable interest in the use of *Geobacillus* in a range of biotechnological applications (4). *Geobacillus* sp. strains CAMR5420 and CAMR12739 were acquired from the CAMR (Porton Down, United Kingdom) thermophile culture collection, which is now held by the Centre for Extremophile Research, Department of Biology and Biochemistry, at the University of Bath, United Kingdom. CAMR5420 was added to the collection in 1989; CAMR12739 was collected in Skaltholt, Iceland, and deposited in the collection in 1987.

Genome sequencing was performed using the 454 GS-FLX (CAMR12739) and Illumina GAIx (CAMR5420) platforms, yielding 2,533,592 reads (average read length of 138 nucleotides [nt]) for CAMR5420 and 119,419 reads (average read length of 365 nt) for CAMR12739. Assembly was performed using a combination of the CLC Genomics Workbench v6, DNASTar Ngen and Velvet v1.1 (5) assemblers. Finally, reference assembly was undertaken using the complete genomes of related *Geobacillus* spp. using Mauve (6). The genome of CAMR5420 was assembled into 96 contigs, yielding a total genome size of 3.50 megabases, with a mean G+C content of 51.89%. The genome of CAMR12739 was assembled into 74 contigs, with a total genome size of 3.41 Mb and a mean G+C content of 52.21%. Further *in silico* assembly was hampered by the presence of highly conserved transposase genes. The genomes were annotated using the Rapid Annotations using Subsystems Technology (RAST) server (7). The genomes of CAMR5420 and CAMR12739 code for 3,816 and 5,836 proteins, respectively. A single, but variable, locus underlying the capacity of these strains to hydrolyze and utilize hemicellulose polymers was identified. The genomes of *Geobacillus* sp. CAMR5420 and CAMR12739 will be employed for comparative genomic analyses, with the goal of developing recombinant *Geobacillus* strains with optimal hemicellulose degrading capacities.

Nucleotide sequence accession numbers. These whole-genome shotgun projects have been deposited at DDBJ/ENA/GenBank under the

accession no. JHUS00000000 (CAMR5420) and JHUR00000000 (CAMR12739). The versions described in this paper are the first versions, JHUS01000000 (CAMR5420) and JHUR01000000 (CAMR12739).

ACKNOWLEDGMENTS

We gratefully acknowledge financial and technical support from the Genomics Research Institute, University of Pretoria, and TMO Renewables Ltd., United Kingdom, and sequencing support from the University of the Western Cape sequencing facility.

REFERENCES

- Coorevits A, Dinsdale AE, Halket G, Lebbe L, De Vos P, Van Landschoot A, Logan NA. 2012. Taxonomic revision of the genus *Geobacillus*: emendation of *Geobacillus*, *G. stearothersophilus*, *G. jurassicus*, *G. toebii*, *G. thermodenitrificans* and *G. thermoglucosidans* (nom. Corrig., formerly “*thermoglucosidarius*”); transfer of *Bacillus thermantarcticus* to the genus as *G. thermantarcticus* comb. nov.; proposal of *Caldibacillus debilis* gen. nov., comb. nov.; transfer of *G. tepidamans* to *Anoxybacillus* as *A. tepidamans* comb. nov.; and proposal of *Anoxybacillus caldiproteolyticus* sp. nov. *Int. J. Syst. Evol. Microbiol.* 62:1470–1485. <http://dx.doi.org/10.1099/ijs.0.030346-0>.
- Shulami S, Gat O, Sonenshein AL, Shoham Y. 1999. The glucuronic acid utilization gene cluster from *Bacillus stearothersophilus* T-6. *J. Bacteriol.* 181:3695–3704.
- Shulami S, Zaide G, Zolotnitsky G, Langut Y, Feld G, Sonenshein AL, Shoham Y. 2007. A two-component system regulates the expression of an ABC transporter for xylo-oligosaccharides in *Geobacillus stearothersophilus*. *Appl. Environ. Microbiol.* 73:874–884. <http://dx.doi.org/10.1128/AEM.02367-06>.
- Taylor MP, Eley KL, Martin S, Tuffin MI, Burton SG, Cowan DA. 2009. Thermophilic ethanologeneses: future prospects for second-generation bioethanol production. *Trends Biotechnol.* 27:398–405. <http://dx.doi.org/10.1016/j.tibtech.2009.03.006>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
- Darling AE, Tritt A, Eisen JA, Facciotti MT. 2011. Mauve assembly metrics. *Bioinformatics* 27:2756–2757. <http://dx.doi.org/10.1093/bioinformatics/btr451>.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formisano K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.