



Genome Announcement

Draft genome sequence of pectic polysaccharide-degrading moderate thermophilic bacterium *Geobacillus thermodenitrificans* DSM 101594

Raimonda Petkauskaitė^{a,*}, Jochen Blom^b, Alexander Goesmann^b, Nomeda Kuisiene^a

^a Vilnius University, Faculty of Natural Sciences, Department of Microbiology and Biotechnology, Vilnius, Lithuania

^b Justus-Liebig-Universität Giessen, Bioinformatik und Systembiologie, Giessen, Germany

ARTICLE INFO

Article history:

Received 29 March 2016

Accepted 9 June 2016

Available online 21 September 2016

Associate Editor: John Anthony McCulloch

Keywords:

Geobacilli

Polysaccharide degradation

Genome

ABSTRACT

Geobacillus thermodenitrificans DSM 101594 was isolated as a producer of extracellular thermostable pectic polysaccharide degrading enzymes. The completely sequenced genome was 3.6 Mb in length with GC content of 48.86%. A number of genes encoding enzymatic active against the high molecular weight polysaccharides of potential biotechnological importance were identified in the genome.

© 2016 Sociedade Brasileira de Microbiologia. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

In recent years, number of sequenced geobacilli genomes has increased substantially.¹ The annotation of genomes of newly isolated geobacilli strains is extremely important as these thermophiles are one of the major sources of thermoactive and/or thermostable enzymes of biotechnological importance.

This communication presents the draft genome sequence of *G. thermodenitrificans* PA-3 designated as DSM 101594, a highly active pectate lyase producer. This thermophilic bacterium was isolated from a soil sample collected from

the compost facility at Vilnius University Botanical Garden, Vingis Park, Vilnius, Lithuania. The strain was isolated via enrichment culture with polygalacturonic acid and subsequently apple pomace as the main source of carbon and energy. The potential of a recombinant thermostable pectate lyase (locus_tag GEPA3.0510) from *G. thermodenitrificans* DSM 101594 has already been demonstrated for the enzymatic production of long-chain pectic oligosaccharides, which serves as a valuable prebiotics obtained from the agro-industrial wastes.^{2,3}

* Corresponding author at: Sauletekio ave. 7, LT-10257 Vilnius, Lithuania.

E-mail: raimonda.petkauskaitė@gf.vu.lt (R. Petkauskaitė).

<http://dx.doi.org/10.1016/j.bjm.2016.06.013>

1517-8382/© 2016 Sociedade Brasileira de Microbiologia. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Table 1 – Genome features of *Geobacillus thermodenitrificans* DSM 101594.

Features	Chromosome
Length (bp)	3,646,477
G+C content (%)	48.86
ORFs	3638
rRNA regions	10
tRNA regions	92

16S rRNA gene phylogenetic analysis confirmed the strain PA-3 to be *G. thermodenitrificans*. 16S rRNA gene sequence comparison was performed according to UPGMA algorithm implemented in MEGA 6.⁴

A genome assembly was constructed in order to determine the potential biotechnological applicability of *G. thermodenitrificans* DSM 101594 and to identify the genes of interest. Total DNA was extracted from liquid growth culture aerobically cultivated in Difco™ Nutrient Broth (BD Diagnostics) at 60 °C overnight using GeneJET™ Genomic DNA Purification Kit (Thermo Fisher Scientific) following the manufacturer's instructions. A draft of whole-genome sequence was obtained using the next-generation sequencing (NGS); paired-end 100 cycles sequence reads were generated using Illumina HiSeq2500 system (BaseClear, Leiden, the Netherlands). FASTQ sequence reads were generated using Illumina Casava pipeline (version 1.8.3). 1,317,704 read pairs covering 213,468,048 bases were assembled using SPAdes Genome Assembler software (version 3.1.0) resulting in 208 contigs; of which 25 contigs shorter than 200 bp were discarded. The final assembly consisted of 3,646,477 bp with an average coverage of 58.5×. The genome has a GC content of 48.86%. Automated genome annotation was carried out by using GenDB software.⁵ The automated gene prediction identified 3,638 coding sequences (CDS), 10 rRNA regions, and 92 tRNA regions (Table 1). The direct manual annotation was performed using Pfam,⁶ InterProScan,⁷ and NCBI BLAST.

Determined genome sequence was compared using EDGAR⁸ software with the genome sequences of other geobacilli, including the phylogenetically nearest strains, *G. thermodenitrificans* NG80-2⁹ and *G. thermodenitrificans* DSM 465.¹⁰ The analysed *G. thermodenitrificans* DSM 101594 genome revealed an adaptational trait that enforces this thermophile to occupy the environmental niches rich in insoluble high molecular weight polysaccharides. A 60 kbp long region, previously not identified in any other geobacilli genomes encoding genes for the utilisation of a broad range of polysaccharides and genes of polysaccharide degradation product transport systems, is the key finding of this trait. The ability of *G. thermodenitrificans* DSM 101594 to use different polysaccharides for nutrition demonstrates flexible adaptation to available energy and/or carbon sources. Finally, a substantial number of genes encoding powerful biocatalysts was annotated in the genome of this strain.

Nucleotide sequence accession numbers

Draft genome sequence for *G. thermodenitrificans* DSM 101594 has been deposited at GenBank under accession no. LIDX00000000. The version described in this paper is version LIDX01000000.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

This study was funded by “TermozymOS” project of the National Research Programme “Healthy and Safe Food” by the Lithuanian Science Council (project no. SVE-08/2011). We acknowledge the access to resources within the de.NBI network financially supported by the BMBF grant FKZ 031A533.

REFERENCES

- Benson DA, Cavanaugh M, Clark K, et al. GenBank. *Nucleic Acids Res.* 2013;41 (Database issue):D36–D42.
- Petkauskaite R, Lukosius D, Dębski J, et al. Identification of proteins involved in starch and polygalacturonic acid degradation using LC–MS. *Cent Eur J Biol.* 2014;9(7):708–716.
- Kieraitė I, Petkauskaite R, Jasilionis A, Kuisiėnė N. Evaluation of potential of free and immobilized thermophilic bacterial enzymes in the degradation of agro-industrial wastes. *Arch Biol Sci.* 2015;67(1):161–172.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary genetics analysis Version 6.0. *Mol Biol Evol.* 2013;30(12):2725–2729.
- Meyer F, Goesmann A, McHardy AC, et al. GenDB – an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.* 2003;31(8):2187–2195.
- Punta M, Cogill PC, Eberhardt RY, et al. The Pfam protein families database. *Nucleic Acids Res.* 2012;40 (Database issue):D290–D301.
- Quevillon E, Silventoinen V, Pillai S, et al. InterProScan: protein domains identifier. *Nucleic Acids Res.* 2005;33(suppl 2):W116–W120.
- Blom J, Albaum SP, Doppmeier D, et al. EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinform.* 2009;10:154, <http://dx.doi.org/10.1186/1471-2105-10-154>.
- Feng L, Wang W, Cheng J, et al. Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep subsurface oil reservoir. *Proc Natl Acad Sci U S A.* 2007;104(13):5602–5607.
- Yao N, Ren Y, Wang W. Genome sequence of a thermophilic bacillus. *Geobacillus thermodenitrificans* DSM 465. *Genome Announc.* 2013;1(6):e01046–e1113, <http://dx.doi.org/10.1128/genomeA.01046-13>.