Genomics Data 12 (2017) 49–51

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Data in Brief

Draft genome sequence of *Thermoanaerobacterium* sp. strain PSU-2 isolated from thermophilic hydrogen producing reactor

Sompong O-Thong [a,b,*], Peerawat Khongkliang [b], Chonticha Mamimin [b], Apinya Singkhala [b], Poonsuk Prasertsan [c], Nils-Kåre Birkeland [d]

[a] Research Center in Energy and Environment, Faculty of Science, Thaksin University, Phatthalung 93210, Thailand
[b] Biotechnology Program, Department of Biology, Faculty of Science, Thaksin University, Phatthalung 93210, Thailand
[c] Department of Industrial Biotechnology, Faculty of Agro-Industry, Prince of Songkla University, Songkhla 90112, Thailand
[d] Department of Biology and Centre for Geobiology, University of Bergen, P.O. Box 7800, N-5020 Bergen, Norway

ARTICLE INFO

Article history:
Received 14 February 2017
Accepted 24 February 2017
Available online 28 February 2017

Keywords:
Whole genome sequencing
Thermoanaerobacterium sp.
Hydrogen producing bacteria
Thermophile

ABSTRACT

Thermoanaerobacterium sp. strain PSU-2 was isolated from thermophilic hydrogen producing reactor and subjected to draft genome sequencing on 454 pyrosequencing and annotated on RAST. The draft genome sequence of strain PSU-2 contains 2,552,497 bases with an estimated G + C content of 35.2%, 2555 CDS, 8 rRNAs and 57 tRNAs. The strain had a number of genes responsible for carbohydrates metabolic, amino acids and derivatives, and protein metabolism of 17.7%, 14.39% and 9.81%, respectively. Strain PSU-2 also had gene responsible for hydrogen biosynthesis as well as the genes related to Ni-Fe hydrogenase. Comparative genomic analysis indicates strain PSU-2 shares about 94% genome sequence similarity with *Thermoanaerobacterium xylanolyticum* LX-11. The nucleotide sequence of this draft genome was deposited into DDBJ/ENA/GenBank under the accession MSQD00000000.

Specifications	
Organism/cell line/tissue	*Thermoanaerobacterium* sp.
Strain	PSU-2
Sequencer or array type	454 pyrosequencing
Data format	analyzed
Experimental factors	microbial strain
Experimental features	draft genome analysis and gene annotation of PSU-2
Consent	N/A
Sample source location	Songkhla, Thailand

1. Direct link to deposited data

The draft genome sequences could be found at the site http://www.ncbi.nlm.nih.gov/nuccore/MSQD00000000

2. Experimental design, materials and methods

The genus *Thermoanaerobacterium* is a group of anaerobic, gram-positive, rod shaped, reduce thiosulfate to elemental sulfur that belong to *Firmicutes* as previously described by Lee et al. [1]. Genus *Thermoanaerobacterium* are thermophilic that specialize in polysaccharide and carbohydrate fermentation, producing primarily L-lactic acid, acetic acid, ethanol, CO_2, and H_2 [2,3]. The majority of characterized *Thermoanaerobacterium* strains have been isolated from hot springs and other thermal environments [4]; however, they have also been isolated from leachate of a waste pile from a canning factory [5], thermophilic bioreactor for biohydrogen production [6,7] and deep subsurface environments [8]. This genus has been considered for biotechnological applications, such as conversion of lignocellulosic biomass to ethanol [3], biohydrogen and other chemicals [9]. *Thermoanaerobacterium* strain PSU-2 is a rod shaped, gram-positive, spore-forming and thermophilic hydrogen producing bacteria belonging to *Firmicutes* that was isolated from a biohydrogen reactor fed with palm oil mill effluent (POME). Phylogenetic analysis based on 16S rRNA genes indicated that strain PSU-2 belonged to the genus *Thermoanaerobacterium* [6]. This genus had been previously studied for hydrogen production from various carbohydrates, such as starch, sucrose and molasses [10]. Strain PSU-2 has a high hydrogen production capacity within a wide range of pH (4.5–8) and temperature (45–70 °C), with the optimal temperature 60 °C and optimal initial pH about 6.25. The strain performed ethanol–acetate type fermentation in inorganic nitrogen amended medium, while it performed butyrate–acetate type fermentation in organic nitrogen amended medium [6].

* Corresponding author at: Biotechnology Program, Department of Biology, Faculty of Science, Thaksin University, Phatthalung 93210, Thailand.
E-mail address: sompong.o@gmail.com (S. O-Thong).

http://dx.doi.org/10.1016/j.gdata.2017.02.012

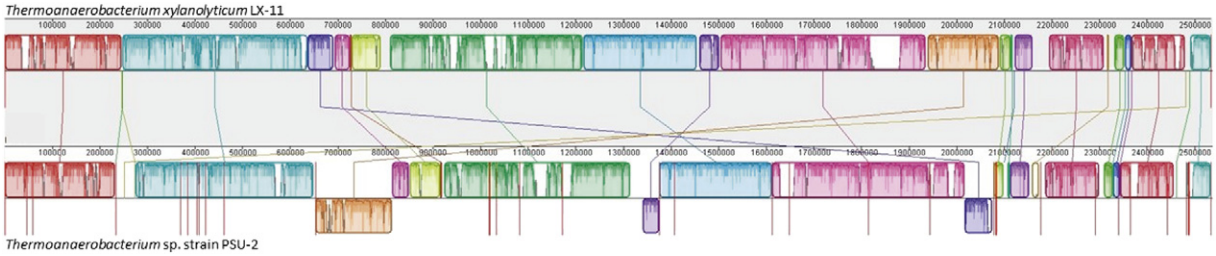

Fig. 1. Draft genome alignment of *Thermoanaerobacterium* sp. strain PSU-2 with *Thermoanaerobacterium xylanolyticum* LX-11.

The draft genome of strain PSU-2 was sequenced with 454 technology using a GS-FLX pyrosequencer at GATC Biotech, Germany (http://www.gatc-biotech.com). A total of 2,552,497 bases were obtained. Assembly into 44 contigs was done with Newbler version 2.9 accessed through the Lifeportal, University of Oslo (http://www.uio.no/english / services/it/research/hpc/lifeportal) and annotation was conducted on RAST [11]. SEED viewer was used for subsystem functional categorization of the predicted open reading frames (ORFs) and visualization [12]. An average nucleotide identity (ANI) was analysis using the online ANI calculator (http://enve-omics.ce.gatech.edu/ani/index). An *In Silico* genomic DNA:DNA hybridization was performed by genome-to-genome distance calculator (http://ggdc.dsmz.de).

3. Data description

The draft genome of *Thermoanaerobacterium* sp. strain PSU-2 consisted of single DNA chromosome of 2,552,497 bases, a G + C content of 35.2%. The draft genome was predicted to contain 2555 protein-coding sequence, 8 rRNAs and 57 tRNAs. These genes were annotated and classified into 337 subsystems. Most of the annotated genes were involved in carbohydrates metabolic (17.7%), amino acids (14.39%) and derivatives, and protein metabolism (9.81%) (Fig.1). Strain PSU-2 also had gene responsible for hydrogen biosynthesis as well as the genes related to Ni-Fe hydrogenase. Comparative genomic analysis indicates PSU-2 by an average nucleotide identity (ANI) analysis using the online ANI calculator (http://enve-omics.ce.gatech.edu/ani/index) revealed ANI values of 94, when the PSU-2 draft sequence was compared with complete sequences of *Thermoanaerobacterium xylanolyticum* LX-11 species, isolated from geothermal areas of Yellowstone National Park, Wyoming, USA (Fig. 2). This indicates that PSU-2 represents a separate species, as this value is lower than the threshold value of 95%, which corresponds to a genomic DNA:DNA hybridization value of 70% and is a common threshold value for distinction between species [13].

4. Nucleotide accession number

This whole genome project has been deposited at DDBJ/ENA/GenBank under accession no. MSQD00000000. The version described in this paper is version MSQD01000000.

Conflict of interest

The authors clarified that this work and writing has no conflict of interest.

Acknowledgements

This work was supported by the Core-to-Core Program, which was financially supported by Japan Society for the Promotion of Science (JSPS), National Research Council of Thailand (NRCT), Vietnam Ministry of Science and Technology (MOST), the National University of Laos, Beuth University of Applied Sciences and Brawijaya University, Research and Development Institute Thaksin University (RDITSU), Agricultural Research Development Agency (ARDA), Research Group for

Name	Count	Percent
Cofactors, Vitamins, Prosthetic Groups, Pigments	113	6.55
Cell Wall and Capsule	106	6.15
Virulence, Disease and Defense	43	2.49
Potassium metabolism	10	0.58
Photosynthesis	0	0
Miscellaneous	16	0.92
Phages, Prophages, Transposable elements, Plasmids	9	0.52
Membrane Transport	57	3.31
Iron acquisition and metabolism	1	0.06
RNA Metabolism	104	6.04
Nucleosides and Nucleotides	67	3.87
Protein Metabolism	169	9.81
Cell Division and Cell Cycle	28	163
Motility and Chemotaxis	80	4.64
Regulation and Cell signaling	23	1.33
Secondary Metabolism	4	0.23
DNA Metabolism	90	5.22
Fatty Acids, Lipids, and Isoprenoids	37	2.15
Nitrogen Metabolism	14	0.81
Dormancy and Sporulation	80	4.64
Respiration	31	1.80
Stress Response	52	3.02
Metabolism of Aromatic Compounds	1	0.06
Amino Acids and Derivatives	248	14.39
Sulfur Metabolism	8	0.46
Phosphorus Metabolism	27	1.57
Carbohydrates	305	17.70
Not in COG	1	0.06

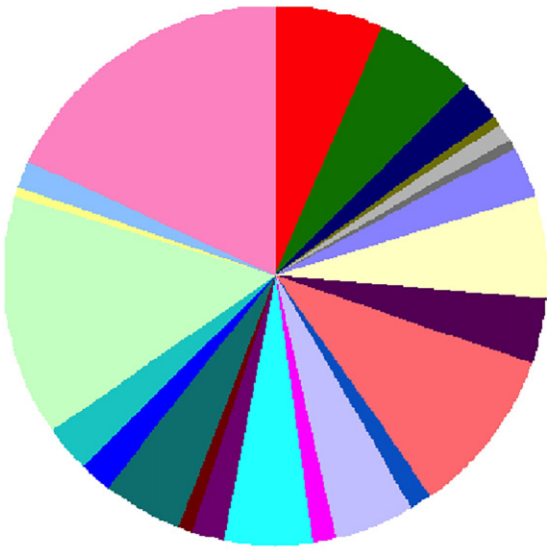

Fig. 2. Distribution and counts of genes in COG categories for draft genome of *Thermoanaerobacterium* sp. strain PSU-2.

Development of Microbial Hydrogen Production Process from Biomass, Khon Kaen University and Thailand Research Fund through grant number PHD57K0042 and RTA5780002.

References

[1] Y.E. Lee, M.K. Jain, C.Y. Lee, S.E. Lowe, J.G. Zeikus, Taxonomic distinction of saccharolytic thermophilic anaerobes description of *Thermoanaerobacterium xylanolyticum* gen nov, sp nov, and *Thermoanaerobacterium saccharolyticum* gen nov, sp nov reclassification of *Thermoanaerobium brockii*, *Clostridium thermosulfurogenes*, and *Clostridium thermohydrosulfuricum* E100-69 as *Thermoanaerobacter brockii* comb nov, *Thermoanaerobacterium thermosulfurigenes* comb nov, and *Thermoanaerobacter thermohydrosulfuricus* comb nov, respectively and transfer of *Clostridium thermohydrosulfuricum* to *Thermoanaerobacter ethanolicus*. Int. J. Syst. Bacteriol. 34 (1993) 41–51.

[2] J. Wiegel, C.P. Mothershed, J. Puls, Differences in xylan degradation by various noncellulolytic thermophilic anaerobes and *Clostridium thermocellum*. Appl. Environ. Microbiol. 49 (1985) 656–659.

[3] L.R. Lynd, P.J. Weimer, W.H. van Zyl, I.S. Pretorius, Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. 66 (2002) 506–577.

[4] W. Shao, S. DeBlois, J. Wiegel, A high-molecular-weight, cellassociated xylanase isolated from exponentially growing *Thermoanaerobacterium* sp. strain JW/SL-YS485. Appl. Environ. Microbiol. 61 (1995) 937–940.

[5] I.K.O. Cann, P.G. Stroot, K.R. Mackie, B.A. White, R.I. Mackie, Characterization of two novel saccharolytic, anaerobic thermophiles, *Thermoanaerobacterium polysaccharolyticum* sp nov and *Thermoanaerobacterium zeae* sp nov., and emendation of the genus *Thermoanaerobacterium*. Int. J. Syst. Evol. Microbiol. 51 (2001) 293–302.

[6] S. O-Thong, P. Prasertsan, D. Karakashev, I. Angelidaki, Thermophilic fermentative hydrogen production by the newly isolated *Thermoanerobacterium thermosaccharolyticum* PSU-2. Int. J. Hydrog. Energy 33 (2008) 1204–1214.

[7] N. Ren, G. Cao, A. Wang, D.J. Lee, W. Guo, Y. Zhu, Dark fermentation of xylose and glucose mix using isolated *Thermoanaerobacterium thermosaccharolyticum* W16. Int. J. Hydrog. Energy 33 (2008) 6124–6132.

[8] S.Y. Liu, F.A. Rainey, H.W. Morgan, F. Mayer, J. Wiegel, *Thermoanaerobacterium aotearoense* sp. nov., a slightly acidophilic, anaerobic thermophile isolated from various hot springs in New Zealand, and emendation of the genus *Thermoanaerobacterium*. Int. J. Syst. Bacteriol. 46 (1996) 388–396.

[9] N.E. Altaras, M.R. Etzel, D.C. Cameron, Conversion of sugars to 1,2-propanediol by *Thermoanaerobacterium thermosaccharolyticum* HG-8. Biotechnol. Prog. 17 (2001) 52–56.

[10] S. O-Thong, P. Prasertsan, D. Karakashev, I. Angelidaki, 16S rRNA-targeted probes for specific detection of *Thermoanaerobacterium* spp., *Thermoanaerobacterium thermosaccharolyticum*, and *Caldicellulosiruptor* spp. by fluorescent in situ hybridization in biohydrogen producing systems. Int. J. Hydrog. Energy 33 (2008) 6082–6091.

[11] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, The RAST server: rapid annotations using subsystems technology. BMC Genomics 9 (1) (2008) 75.

[12] R. Overbeek, R. Olson, G.D. Pusch, G.J. Olsen, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, B. Parrello, M. Shukla, The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res. 42 (D1) (2014) D206–D214.

[13] J. Goris, K.T. Konstantinidis, J.A. Klappenbach, T. Coenye, P. Vandamme, J.M. Tiedje, DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. 57 (2007) 81–91.