View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Genomics Data 5 (2015) 109–111

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

Isolation of chitinolytic Clostridium sp. NCR from Mehsani buffalo rumen, its genomic analysis and potential role in rumen

Neelam M. Nathani [a,c], Srinivas M. Duggirala [b], Chandra Shekar M. [a], Ramesh K. Kothari [c,*], Chaitanya G. Joshi [a]

[a] Department of Animal Biotechnology, Anand Agricultural University, Anand, Gujarat, India
[b] Department of Microbiology, Gujarat Vidyapeeth, Sadra, Gujarat, India
[c] Department of Biosciences, Saurashtra University, Rajkot, Gujarat, India

ARTICLE INFO

Article history:
Received 11 May 2015
Accepted 21 May 2015
Available online 30 May 2015

Keywords:
Clostridium sp. NCR
Chitinolytic
Genome sequence
Rumen microflora

ABSTRACT

Genomic analysis of Clostridium sp. NCR, an anaerobic Gram positive bacterium which was isolated from rumen fluid of Mehsani breed of buffalo revealed presence of various environmental gene tags (EGTs) involved in pathways for utilizing a wide range of substrates. Here we report the sequence of this rumen isolate, its whole genome sequence has been deposited in DDBJ/EMBL/GenBank under the accession number JQHY00000000. The genome comprises of a 3.62-Mb draft genome with a G + C content of 28.10%, which encodes a total of 3126 proteins. Functional analysis provides information about the microbe's role in maintaining host homeostasis and its fiber degradation potential.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Specifications	
Organism	*Clostridium* sp. NCR
Sex	N/A
Sequencer	Ion Torrent PGM
Data format	Analyzed
Experimental factors	N/A
Experimental features	Genome assembly from whole genome sequencing using Ion torrent
Consent	Allowed for reuse citing original authors
Sample source location	Anand, Gujarat, India

1. Direct link to deposited data

Deposited data can be found here: http://www.ncbi.nlm.nih.gov/genome/13521?genome_assembly_id=210853.

Rumen harbors one of the most explored complex microbial communities. Specifically, the rumen bacteria play a vital role in several functions that include digestion of host dietary feed substrates, inhibition of pathogenic inhabitants, host immune development, and also improvement of animal product quality leading to a simultaneous increase in the interest of researchers to study the rumen bacterial population [1–4]. The existence of diverse bacterial species provide the edge to ruminants for consumption of lignocellulosic biomass to obtain their energy requirements through volatile fatty acids such as acetate, propionate and butyrate generated from the degradation and anaerobic fermentation of fibrous materials. The phylotypes of rumen bacteria have been continuously increasing by the addition of 16S rDNA gene sequences. Nevertheless, a good number of these rumen bacteria classified into various taxa based on 16S rRNA sequences are still unknown and their biochemical properties need to be studied [5].

Clostridium strains have been reported to inhabit the human as well as animal gut [6,7]. They are known to possess the potential to synthesize various organic solvents such as butanol, acetate, butyrate, lactate, propionate, and ethanol as by-products via various fermentation pathways [2]. In our study, to elucidate the microflora of rumen, we isolated anaerobic Gram positive bacteria able to grow on polyphenols as the major substrate source in the culture medium. To obtain the genome sequence, shotgun sequencing of Clostridium sp. NCR, was performed using the ion 316-chip and 400-bp chemistry on Ion Torrent PGM platform as per the manufacturer's instructions.

2. Data processing and de novo assembly

The sequence reads were quality checked and filtered with a minimum quality score of 20 using FastQC and Prinseq. The reads were further subjected to De Novo assembly using GS Assembler V 2.6 resulting in the draft genome of 3,618,410 bp comprising 56 contigs of >200 bp in size. The average coverage of the assembled contigs was 87.85 fold. The gene annotation and screening for RNAs were performed using Rapid Annotation using Subsystem Technology (RAST) server [8], results enlisted in Table 1.

* Corresponding author at: Department of Biosciences, Saurashtra University, Rajkot 360005, Gujarat, India.
E-mail address: kothari1971@gmail.com (R.K. Kothari).

http://dx.doi.org/10.1016/j.gdata.2015.05.017
2213-5960/© 2015 The Authors. Published by Elsevier Inc. 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 Clostridium sp. NCR.

Features	Chromosome
Length [bp]	3,618,410
G + C content [%]	28.10
Coding sequences	3305
rRNA genes	16
tRNA genes	63

3. Taxonomic classification

For its taxonomic placement, the contigs were subjected to BLAST against the 16S rDNA database from NCBI. The results showed best hit at >98% identity with 16S rDNA sequence of Clostridium difficile and *Clostridium glycolicum*, but on performing reference guided assembly with the available reference genomes it showed negligible reference coverage (<10% average coverage). Ribosomal Database Project (RDP) classifier placed the species under the genus Clostridium Class X1 and Unclassified Peptostreptococcaceae. Further, in silico DNA–DNA hybridization analysis estimated the Average Nucleotide Identity (ANI) to be about ~80% between Clostridium sp. NCR and various strains of the nearest published genomes, which is much lower than their expected standard ANI values between genomes of the same species viz., 95% [9]. Genomic and phylogenetic analyses designated the isolate as a novel Clostridium species, indicating its phylogenetic placement within the Clostridium genus diverged from *C. difficile* and *C. glycolicum*.

4. Genome annotation and analysis

Annotation showed the presence of CDS (Coding Sequence) involved in degrading chitin and also complete pathways for production of volatile fatty acids acetate and butyrate. Clostridium has also earlier been found as an important genera present in the rumen content of different animals which plays an important role in chitin degradation and produced mainly acetate, butyrate and lactate as the end products [10].

Biochemical characterization using Biolog MicroPlates (AN), the strain showed a positive reaction for the assimilation of D-mannitol, raffinose, trehalose; L-rhamnose and maltose. The isolate is able to hydrolyze starch and is negative for the reduction of nitrate to nitrite. Carbohydrate active enzymes (CAZymes) were studied and compared with the nearest *C. difficile* and *C. glycolicum* using the sequence based annotation from Cazyme Analysis Toolkit (CAT) [11]. The results showed that the isolate varied in the proportions of EGTs for various Cazy families as compared to the nearing species (Fig. 1). Glycoside hydrolase was the most abundant family, followed by Carbohydrate Binding Modules (CBM), while the glycosyl transferase (GT) and carbohydrate esterase (CE) were poorly represented. There was a single Auxillary activity (AA) EGT sequence observed in the genome related to the AA6 family that codes for the nitrophenol reducing enzyme 1,4-benzoquinone reductase (EC. 1.6.5.6). Among the GH family, further classification included the maximum EGTs for GH18, followed by GH25, GH73, GH4, GH2, GH3, GH38 and GH39. Chitinase enzyme falls in the GH18 family which was the most abundant and the CBM also involve the CBM5 family that helps in binding of the chitin moieties. Other families included the EGTs for various enzymes utilizing sugar moieties like mannose, maltose, xylan and various other oligosaccharides.

5. Summary

Genomic analysis of the isolate gives a genetic confirmation to its chitinolytic potential. Thus, the information obtained from the whole genome sequence about the metabolic pathways of the strain helps reveal the genes coding for enzymes involved in microbe's role in maintaining host homeostasis and its fiber degradation potential. The sequence of has been deposited in DDBJ/EMBL/GenBank under the accession number JQHY00000000 and the version described in this paper is version JYOG01000000. After the primary analysis of the assembly described above, the sequence will be taken up for further analysis including comparison with the genome of other Clostridium species to yield better insight into the isolates' adaptation to the highly complex rumen environment and its interactions with other rumen microbes.

Competing interests

The authors have declared that no competing interest exists.

Acknowledgments

The work was supported by the Niche area of excellence program funded by Indian Council of Agricultural Research, New Delhi, India. The authors are thankful to Department of Science and Technology, Government of India for providing grant in the form of INSPIRE fellowship to NMN.

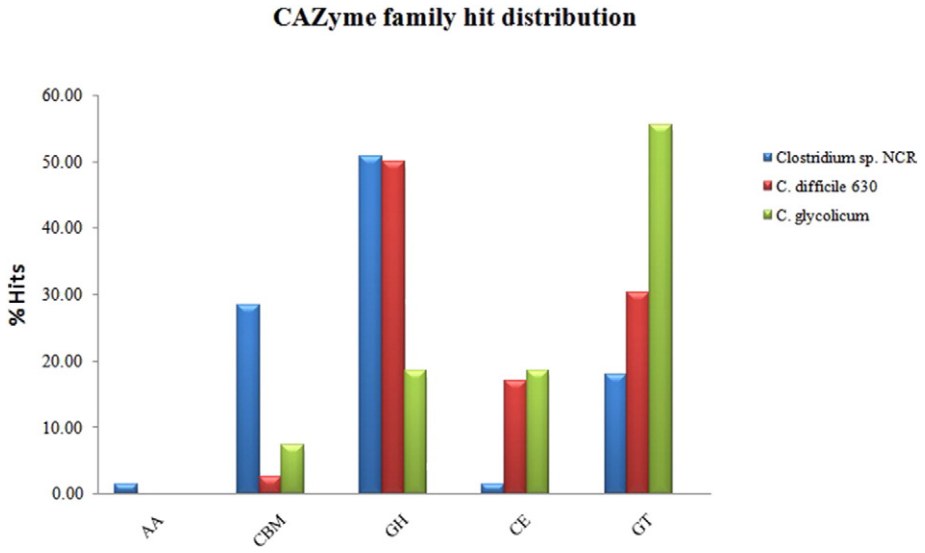

Fig. 1. Comparative distribution of CAZyme families between the studied *Clostridium* isolate and two related reference species: *C. difficile* and *C. glycolicum*.

References

[1] M. Castillo, S.M. Martin-Orue, E.G. Manzanilla, I. Badiola, M. Martin, J. Gasa, Quantification of total bacteria, enterobacteria and lactobacilli populations in pig digesta by real-time PCR. Vet. Microbiol. 114 (1–2) (2006) 165–170.

[2] S. Dabrowski, E. Zablotna, D. Pietrewicz-Kubicz, A. Dlugolecka, Screening of environmental samples for bacteria producing 1,3-propanediol from glycerol. Acta Biochim. Pol. 59 (3) (2012) 353–356.

[3] A.N. Hristov, T.R. Callaway, C. Lee, S.E. Dowd, Rumen bacterial, archaeal, and fungal diversity of dairy cows in response to ingestion of lauric or myristic acid. J. Anim. Sci 90 (12) (2012) 4449–4457.

[4] K.M. Singh, P.R. Pandya, A.K. Tripathi, G.R. Patel, S. Parnerkarb, R.K. Kothari, et al., Study of rumen metagenome community using qPCR under different diets. Meta Gene 2 (2014) 191–199.

[5] T. Nyonyo, T. Shinkai, M. Mitsumori, Improved culturability of cellulolytic rumen bacteria and phylogenetic diversity of culturable cellulolytic and xylanolytic bacteria newly isolated from the bovine rumen. FEMS Microbiol. Ecol. 88 (3) (2014) 528–537.

[6] R.E. Hungate, Microorganisms in the rumen of cattle fed a constant ration. Can. J. Microbiol. 3 (2) (1957) 289–311.

[7] B.J. Paster, J.B. Russell, C.M. Yang, J.M. Chow, C.R. Woese, R. Tanner, Phylogeny of the ammonia-producing ruminal bacteria *Peptostreptococcus anaerobius*, Clostridium *sticklandii*, and Clostridium *aminophilum* sp. nov. Int. J. Syst. Bacteriol. 43 (1) (1993) 107–110.

[8] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, et al., The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9 (2008) 75.

[9] 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.

[10] J. Kopecny, B. Hodrova, C.S. Stewart, The isolation and characterization of a rumen chitinolytic bacterium. Lett. Appl. Microbiol. 23 (3) (1996) 195–198.

[11] B.H. Park, T.V. Karpinets, M.H. Syed, M.R. Leuze, E.C. Uberbacher, CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. Glycobiology 20 (12) (2010) 1574–1584.