BMC Research Notes

DATA NOTE

Open Access

Draft genome sequence of *Clostridium jeddahense* EE-R19 isolated from an anaerobic digester

Elvira E. Ziganshina[1], Waleed S. Mohammed[2] and Ayrat M. Ziganshin[1*]

Abstract

Objectives: *Clostridium* species of the order *Clostridiales* are mostly strictly anaerobic rod-shaped bacteria. They can be detected in a variety of environments, including the intestines of humans and animals, soil, water, and biogas reactors. Species of the genus *Clostridium* are widely used in various biotechnological processes, but several of them have been identified as significant human pathogens. Therefore, investigation at the genome level is necessary to provide valuable information about the ecology, genetics, and phylogenetic diversity of various *Clostridium* species.

Data description: In the present study, we report the whole genome sequence of *Clostridium jeddahense* strain EE-R19, which was isolated from a mesophilic anaerobic digester. The draft genome of *C. jeddahense* EE-R19 consisted of 59 contigs (> 500 bp), which amounted to 3,562,974 with an overall G + C content of 51.79%. The whole genome shotgun project of *C. jeddahense* EE-R19 has been deposited at DDBJ/ENA/GenBank under the accession number JAAVNF000000000.

Keywords: Draft genome, *Firmicutes*, *Clostridium jeddahense*, Anaerobic digester

Objective

Members of the order *Clostridiales* can be found in a variety of environments, including the intestines of humans and animals, soil, water, marine environments, and biogas reactors [1–4]. The clostridia are basically strict anaerobes with gram-positive staining and the ability to sporulate. Their metabolism varies greatly because they can metabolize various compounds, including carbohydrates, proteins, alcohols, amino acids, and purines. Organic acids and alcohols can be obtained during the metabolism of carbohydrates or proteins. However, some of them have been identified as human pathogens [1]. The genome data will expand understanding of the genus *Clostridium* and provide valuable information about ecology and genetics of the species *Clostridium jeddahense*.

Data description

Clostridium jeddahense strain EE-R19 was originally isolated from anaerobically digested distillers grains and cow manure at mesophilic temperatures on nutrient agar plates in a BACTRON anaerobic chamber (Shel Lab) in Kazan, Republic of Tatarstan, Russia. The strain *C. jeddahense* EE-R19 was then cultivated on nutrient agar plates at + 38 °C for 72 h in a BACTRON anaerobic chamber. Biomass from a culture of *C. jeddahense* EE-R19 was harvested, washed twice with sterile K-Na-phosphate buffer (pH 7.0), and genomic DNA was extracted by using FastDNA spin kit (MP Biomedicals) according to the manufacture's protocol. The identity of the strain EE-R19 was confirmed based on morphological, biochemical, and growth characteristics, and finally by sequencing

*Correspondence: ayrat.ziganshin@kpfu.ru
[1] Department of Microbiology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kremlyovskaya Str. 18, Kazan 420008, Russia
Full list of author information is available at the end of the article

Table 1 Overview of data files/data sets

Label	Name of data file/data set	File types (file extension)	Data repository and identifier (DOI or accession number)
Data set 1	*C. jeddahense* EE-R19, whole genome shotgun sequencing project	FASTA	Genbank (https://www.ncbi.nlm.nih.gov/nuccore/JAAVNF000000000.1)
Data set 2	*C. jeddahense* EE-R19, 16S ribosomal RNA gene, partial sequence	FASTA	Genbank (https://identifiers.org/ncbi/insdc:MT253587.1)

its 16S rRNA gene (Table 1) (16S rRNA gene sequence, 1385 bp, BLAST identity of 99.33% to *Clostridium jeddahense* strain JCD (NR_144697.1)). In addition, the average Nucleotide Identity (ANI) values were estimated using the JSpeciesWS online service for the comparison of the whole genome sequences [5], which confirmed that the strain EE-R19 belongs to the *C. jeddahense* species.

DNA libraries were created as previously reported [6, 7] and according to the Illumina protocol. The genome was then sequenced using an Illumina MiSeq system on a paired-end library with MiSeq Reagent Kit v3 (600-cycle). Quality assessment of the FASTQ sequence files was based on the FastQC software (version 0.11.8) [8]. Velvet assembler (version 1.2.10) was used to assemble reads into contigs [9]. Mauve program (version 2.4.0) was used as a contig ordering tool [10]. The bacterial genome annotation was achieved by uploading the genome assembly of the *C. jeddahense* strain EE-R19 to Rapid Annotation using Subsystem Technology (RAST) server [11]. The number of rRNA and tRNA genes was further recognized by using Barrnap (version 0.9) [12] and Aragorn (version 1.2) [13], respectively.

After filtering and quality assessment, reads were then assembled into 59 contigs (>500 bp), finally creating a genome with a total size of 3,562,974, having an average G+C content of 51.79%. The RAST server predicted 3972 protein-coding genes. Most of the annotated genes determined the synthesis of amino acids and derivatives (258), carbohydrates metabolism (253), protein metabolism (164), synthesis of cofactors, vitamins, prosthetic groups and pigments (122), and fatty acids, lipids and isoprenoids metabolism (32). Barrnap and Aragorn predicted 5 rRNA and 47 tRNA genes, respectively. The strain *C. jeddahense* strain EE-R19 has several genes involved in the biodegradation of carbohydrates and proteins, mixed acid and lactate fermentation, butanol biosynthesis, as well as in the metabolism of acetoin and butanediol. Moreover, several genes have been identified that are responsible for resistance to toxic compounds, including copper, cobalt, zinc, and cadmium. The genome data presented here should contribute to further research on this organism.

Limitations

The exact genome length, synteny, number of rRNA genes, and repetitive elements cannot be certainly reported since the data obtained is based on the draft level genome sequence.

Acknowledgements
Not applicable.

Authors' contributions
EEZ and WSM conducted experiments, fulfilled genome analysis and interpretation of the data. AMZ designed the study and prepared the manuscript. All authors read and approved the final manuscript.

Funding
The reported study was funded by the Russian Foundation for Basic Research [Grant No. 18-29-25058]. RFBR did not take part in the design or realization of the study.

Availability of data materials
The data described in this Data note can be freely and openly accessed at DDBJ/ENA/GenBank. Accession Numbers–https://www.ncbi.nlm.nih.gov/nuccore/JAAVNF000000000.1 (whole genome project) [14] and https://identifiers.org/ncbi/insdc:MT253587.1 (16S rRNA gene sequence) [15].

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors have declared that no competing interest exists.

Author details
[1] Department of Microbiology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kremlyovskaya Str. 18, Kazan 420008, Russia. [2] Department of Biotechnology, Faculty of Agriculture, Al-Azhar University, Cairo 11651, Egypt.

Received: 29 April 2020 Accepted: 21 August 2020
Published online: 03 September 2020

References

1. Wells CL, Wilkins TD. *Clostridia*: spore forming anaerobic bacilli. In: Baron S, editor. Medical Microbiology. 4th ed. Galveston: University of Texas Medical Branch; 1996.
2. Lagier JC, Bibi F, Ramasamy D, Azhar EI, Robert C, Yasir M, et al. Non contiguous-finished genome sequence and description of *Clostridium jeddahense* sp. nov. Stand Genomic Sci. 2014;9:1003–19.
3. Ziganshina EE, Belostotskiy DE, Shushlyaev RV, Miluykov VA, Vankov PY, Ziganshin AM. Microbial community diversity in anaerobic reactors

digesting turkey, chicken, and swine wastes. J Microbiol Biotechnol. 2014;24:1464–72.
4. Ziganshin AM, Wintsche B, Seifert J, Carstensen M, Born J, Kleinsteuber S. Spatial separation of metabolic stages in a tube anaerobic baffled reactor: reactor performance and microbial community dynamics. Appl Microbiol Biotechnol. 2019;103:3915–29.
5. Richter M, Rossello-Mora R, Glockner FO, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics. 2016;32:929–31.
6. Ziganshina EE, Mohammed WS, Doijad SP, Shagimardanova EI, Gogoleva NE, Ziganshin AM. Draft genome sequence of *Brevibacterium epidermidis* EZ-K02 isolated from nitrocellulose-contaminated wastewater environments. Data Brief. 2018;17:119–23.
7. Mohammed WS, Ziganshina EE, Shagimardanova EI, Gogoleva NE, Ziganshin AM. Draft genome sequence of *Paenibacillus* sp. EZ-K15 isolated from wastewater systems. BMC Res Notes. 2017;10:734.
8. Babraham Bioinformatics. FastQC: a quality control tool for high throughput sequence data. Cambridge: Babraham Institute; 2011.
9. Zerbino DR. Using the Velvet *de novo* assembler for short-read sequencing technologies. Curr Protoc Bioinform. 2010;11(11):5.
10. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics. 2009;25:2071–3.
11. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:75.
12. Barrnap version 0.9. https://github.com/tseemann/barrnap.
13. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. 2004;32:11–6.
14. Ziganshina EE, Mohammed WS, Ziganshin AM. *Clostridium jeddahense* EE-R19, whole genome shotgun sequencing project. Genbank *JAAVNF000000000.* 2020. https://www.ncbi.nlm.nih.gov/nuccore/JAAVNF000000000.1>.
15. Ziganshina EE, Mohammed WS, Ziganshin AM. *Clostridium jeddahense* EE-R19, 16S ribosomal RNA gene, partial sequence. Genbank *MH651712.* 2020. https://identifiers.org/ncbi/insdc:MT253587.1.

Publisher's Note