



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Draft genome sequence of *Brevibacterium epidermidis* EZ-K02 isolated from nitrocellulose-contaminated wastewater environments



Elvira E. Ziganshina^a, Waleed S. Mohammed^{a,b},
 Swapnil P. Doijad^c, Elena I. Shagimardanova^d,
 Natalia E. Gogoleva^d, Ayrat M. Ziganshin^{a,*}

^a Department of Microbiology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kremlyovskaya str. 18, Kazan 420008, Russia

^b Department of Biotechnology, Faculty of Agriculture, Al-Azhar University, Cairo 11651, Egypt

^c Institute of Medical Microbiology, Justus-Liebig University, Giessen 35392, Germany

^d Laboratory of Extreme Biology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kazan 420021, Russia

ARTICLE INFO

Article history:

Received 11 October 2017

Received in revised form

19 October 2017

Accepted 28 December 2017

Available online 3 January 2018

Keywords:

Brevibacterium epidermidis

Draft genome

Wastewater

ABSTRACT

Brevibacterium spp. are aerobic, nonbranched, asporogenous, gram-positive, rod-shaped bacteria which may exhibit a rod-coccus cycle when cells get older and can be found in various environments. Several *Brevibacterium* species have industrial importance and are capable of biotransformation of various contaminants. Here we describe the draft genome sequence of *Brevibacterium epidermidis* EZ-K02 isolated from nitrocellulose-contaminated wastewater environments. The genome comprises 3,885,924 bp, with a G + C content of 64.2%. This whole genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession PDHL00000000.

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

* Corresponding author.

E-mail address: a.ziganshin06@fulbrightmail.org (A.M. Ziganshin).

Specifications Table

Subject area	Biology
More specific subject area	Genome analysis
Type of data	Table, figures
How data was acquired	Illumina Miseq
Data format	Analyzed
Experimental factors	Genomic DNA from pure culture
Experimental features	Isolation of bacteria, genome sequencing, draft genome assembly and annotation
Data source location	Nitrocellulose-contaminated wastewater environments, Kazan, Russia
Data accessibility	Data are in public repository. This whole genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession PDHL00000000 (https://www.ncbi.nlm.nih.gov/nucore/PDHL00000000). The 16S rRNA gene sequence has been deposited at GenBank under the accession number MG050737 (https://www.ncbi.nlm.nih.gov/nucore/MG050737).

Value of the data

- Several *Brevibacterium* species have industrial importance and are capable of biotransformation of various contaminants; therefore, more investigations at the genomic level are necessary to improve our understanding of their ecology, genetics, as well as potential biotechnological applications.
- Data shown here can be useful for other groups working or studying in the field of application of brevibacteria in bioremediation processes.
- Data demonstrated here can be used by other researchers working or studying in the field of genome analysis.

1. Data

The draft genome sequence of *B. epidermidis* strain EZ-K02 constituted a total of 65 contigs (> 500 bp) with 3,885,924 bp, and a G + C content of 64.2% (Table 1). The RAST server predicted 3443 coding sequences where 1436 coding sequences (42%) were annotated as seed subsystem features and 2007 coding sequences (58%) annotated as outside of the seed subsystem. In total 2457 and 986 coding sequences were assigned as non-hypothetical and hypothetical, respectively. The genome was shown to encode at least 3 rRNAs and 47 tRNAs. The strain EZ-K02 possesses a substantial number of genes which are responsible for nitrate/nitrite ammonification (for example, in case with nitrate released during nitrocellulose denitration) as well as for metabolism of aromatic compounds, including genes involved in benzoate, *p*-hydroxybenzoate, acetophenone, catechol, gentisate and several other compounds biodegradation. Numerous genes responsible for resistance to toxic compounds, including arsenic, cobalt and cadmium, were additionally detected. Hence, *B. epidermidis* EZ-K02 may have high importance in the field of development of several effective environmental biotechnologies, such as environmental bioremediation and wastewater treatment.

Table 1

Comparison of the genomic feature of *Brevibacterium epidermidis* EZ-K02 strain with various *Brevibacterium* strains. The information regarding the reference genomes was received from the EzBioCloud database [11].

Organism	DB accession number	Isolation source	Contigs	Genome size (bp)	G + C (%)	CDSs	r + tRNA genes
<i>B. epidermidis</i> EZ-K02	GCA_002573745.1	Wastewater	65	3,885,924	64.2	3443	3 + 47
<i>B. album</i> DSM 18261	GCA_000426445.1	Saline alkali soil	15	4,094,970	70.9	3559	9 + 59
<i>B. casei</i> S18	GCA_000314575.1	Human skin	43	3,664,641	68.1	3233	6 + 46
<i>B. epidermidis</i> NBRC 14811	GCA_001570805.1	Human skin	25	3,703,261	64.3	3261	3 + 48
<i>B. linens</i> SMQ-1335	GCA_001729525.1	Cheese	1	4,209,935	62.6	3863	12 + 49
<i>B. ravenespurgense</i> CCUG 56047	GCA_001584615.1	Human specimens	14	2,297,397	62.4	2022	6 + 43
<i>B. senegalense</i> JC43	GCA_000285835.2	Human stool	17	3,427,329	69.4	3058	3 + 46

2. Experimental design, materials and methods

B. epidermidis EZ-K02 was isolated from nitrocellulose-contaminated wastewater environments, Kazan, the Republic of Tatarstan, Russia. Such industrial wastes are represented by large amounts of wastewaters contaminated with different dissolved chemical compounds and nitrocellulose powder particles. Also, these microbes are of high importance for the development of effective bioremediation strategies of various polluted environments [1,2,3]. The bacterium optimally grown on LB agar at +30 °C had been cultivated for 24–48 hours. Genomic DNA of the bacterial strain was then extracted and purified with a FastDNA spin kit (MP Biomedicals) and a FastPrep-24 homogenizer (MP Biomedicals) according to the manufacturer's protocol. Concentration and purity (A260/A280) of the extracted genomic DNA were measured with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) and stored at –20 °C until further processing. The bacterial strain EZ-K02 was

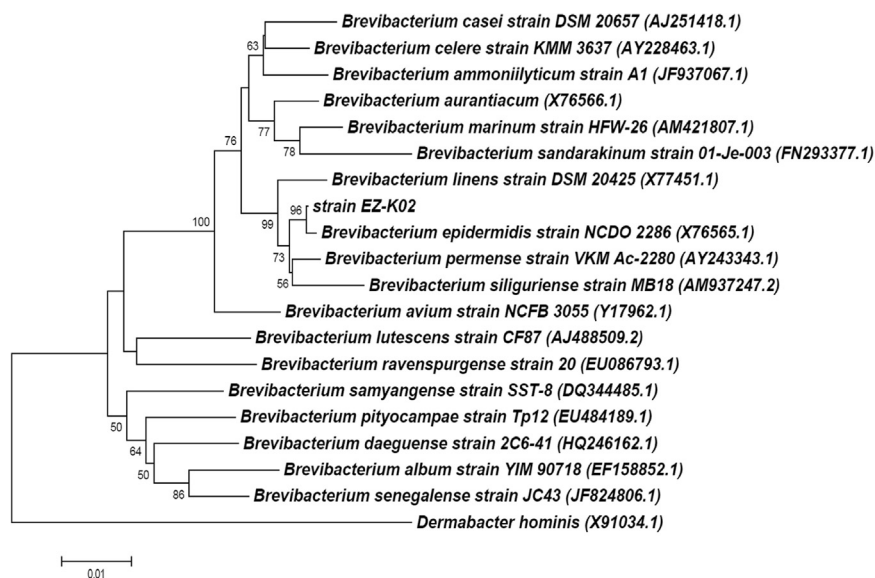


Fig. 1. A phylogenetic tree based on 16S rRNA gene sequences demonstrating the relationship between *Brevibacterium epidermidis* EZ-K02 (NCBI accession number of 16S rRNA gene: MG050737) and the type strains from the LPSN site (www.bacterio.net). Analysis was conducted in MEGA7 [10] using the neighbor-joining method based on Jukes-Cantor evolutionary distances. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. *Dermabacter hominis* was used as outgroup.

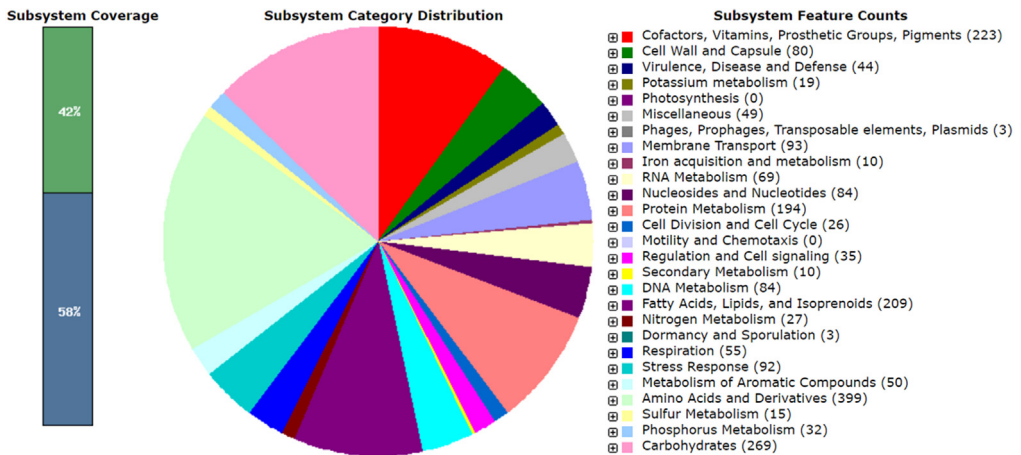


Fig. 2. An overview of the subsystem categories assigned to the genome of *Brevibacterium epidermidis* EZ-K02. The whole genome sequence of the strain EZ-K02 was annotated using the Rapid Annotation System Technology (RAST) server [7]. The pie chart demonstrates the count of each subsystem feature and the subsystem coverage.

morphologically identified and confirmed by PCR amplification using the primers UniBac27f, Bakt_805R and Univ1492r, followed by sequencing using an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems) and phylogenetic analysis (16S rRNA gene sequence; 1413 bp; Fig. 1). In order to perform whole genome analysis, DNA was fragmented using a Q800R2 Sonicator (Qsonica), and DNA library was then prepared with a NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs) according to the manufacturers' protocols. Both efficiency of DNA fragmentation and DNA library preparation were monitored using a 2100 Bioanalyzer (Agilent) and a High Sensitivity DNA kit (Agilent). Further sequencing was performed with a high-throughput Illumina MiSeq platform (Illumina) at Joint KFU-Riken Laboratory, Kazan Federal University (Kazan, Russia) by a MiSeq Reagent Kit v2PE 500 cycles (Illumina). Briefly, sequence read quality was assessed using PRINSEQ lite version 0.20.4 [4], the genome was assembled using Velvet version 1.2.10 [5], and the ordering of contigs was achieved using Mauve version 2.4.0 [6]. The whole genome sequence of *B. epidermidis* was annotated using the Rapid Annotation System Technology (RAST) server [7]. The pie chart demonstrated the counts for each subsystem feature and the subsystem coverage (Fig. 2). The rRNA and tRNA genes were identified using RNAmmer 1.2 [8] and tRNA scan-SE 1.23 [9], respectively.

Acknowledgements

The work was supported by the grant of the President of the Russian Federation for the young scientists [Grant no. МД-100.2017.4].

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at <https://doi.org/10.1016/j.dib.2017.12.053>.

References

- [1] E.E. Ziganshina, E.M. Ibragimov, O.N. Ilinskaya, A.M. Ziganshin, Bacterial communities inhabiting toxic industrial wastewater generated during nitrocellulose production, *Biologia* 71 (2016) 70–78.

- [2] A.M. Ziganshin, R. Gerlach, E.A. Naumenko, R.P. Naumova, Aerobic degradation of 2,4,6-trinitrotoluene by the yeast strain *Geotrichum candidum* AN-Z4, *Microbiology* 79 (2010) 178–183.
- [3] A.M. Ziganshin, E.E. Ziganshina, J. Byrne, R. Gerlach, E. Struve, T. Biktagirov, A. Rodionov, A. Kappler, Fe(III) mineral reduction followed by partial dissolution and reactive oxygen species generation during 2,4,6-trinitrotoluene transformation by the aerobic yeast *Yarrowia lipolytica*, *AMB Express* 5 (2015) 1–12.
- [4] R. Schmieder, R. Edwards, Quality control and preprocessing of metagenomic datasets, *Bioinformatics* 27 (2011) 863–864.
- [5] D.R. Zerbino, Using the Velvet *de novo* assembler for short-read sequencing technologies, *Curr. Protoc. Bioinforma.* 11.5 (2010) 1–13.
- [6] A.I. Rissman, B. Mau, B.S. Biehl, A.E. Darling, J.D. Glasner, N.T. Perna, Reordering contigs of draft genomes using the Mauve aligner, *Bioinformatics* 25 (2009) 2071–2073.
- [7] 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 Genom.* 9 (2008) 75.
- [8] K. Lagesen, P. Hallin, E.A. Rødland, H.H. Staerfeldt, T. Rognes, D.W. Ussery, RNAmmer: consistent and rapid annotation of ribosomal RNA genes, *Nucleic Acids Res.* 35 (2007) 3100–3108.
- [9] T.M. Lowe, S.R. Eddy, tRNA scan-SE: a program for improved detection of transfer RNA genes in genomic sequence, *Nucleic Acids Res.* 25 (1997) 955–964.
- [10] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, *Mol. Biol. Evol.* 33 (2016) 1870–1874.
- [11] S.H. Yoon, S.M. Ha, S. Kwon, J. Lim, Y. Kim, H. Seo, J. Chun, Introducing EzBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies, *Int. J. Syst. Evol. Microbiol.* 67 (2017) 1613–1617.