Data in brief 24 (2019) 103853

Contents lists available at ScienceDirect

Data in brief

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

Data Article

Draft genome sequence and analysis of *Klebsiella oxytoca* strain NK-1 isolated from ureteral stent

Adeliia G. Giliazeva [a, *], Elena I. Shagimardanova [b], Leyla H. Shigapova [b], Daria S. Pudova [a], Margarita R. Sharipova [a], Ayslu M. Mardanova [a]

[a] *Department of Microbiology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kremlyovskaya Str. 18, Kazan 420008, Russia*
[b] *Laboratory of Extreme Biology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Parizhskoy Commuy Str. 9, Kazan 420021, Russia*

ARTICLE INFO

Article history:
Received 9 January 2019
Received in revised form 11 March 2019
Accepted 12 March 2019
Available online 3 April 2019

Keywords:
Klebsiella oxytoca
Catheter-associated urinary tract infection
Draft genome
Illumina MiSeq

ABSTRACT

Klebsiella oxytoca is a facultative aerobic, gram-negative, rod-shaped bacterium capable of causing nosocomial infections, in particular catheter-associated urinary tract infections (CAUTIs). Data on the possible roles of uncommon pathogens such as *K. oxytoca* in the pathogenesis of biofilm-associated infections such as CAUTIs have been already reported. Herein, we describe the draft genome sequence of *K. oxytoca* strain NK-1 isolated from the surface of ureteral stent retrieved from a Russian female. The genome comprises 6,232,464 bp, with a G + C content of 55.60% and an L_{50} of 7. A total of 6246 putative protein-encoding genes were predicted, including considerable number of genes responsible for adhesion, invasion, drug resistance, iron acquisition and other genes relevant for virulence. The NK-1 strain was ascribed a sequence type (ST) as ST 216 (4, 6, 19, 10, 46, 24, 31). Data comparison of the *rec*A gene sequences confirmed that the strain belongs to the species *K. oxytoca*. Minimal inhibitory concentration of different antibiotics have been determined. This whole genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number QPKC00000000.1.

* Corresponding author.
E-mail address: adeliyagilyazeva@gmail.com (A.G. Giliazeva).

https://doi.org/10.1016/j.dib.2019.103853

Specifications Table

Subject area	Biology
More specific subject area	Genome analysis
Type of data	Table, figures
How data was acquired	Genome sequencing: Illumina Miseq (Illumina, CA), Denovo sequence assembly: SPAdes v.3.10.0 software, Bioinformatics approaches: RAST (Rapid annotation sequence technology) server (http://rast.nmpdr.org/), RNAmmer 1.2 server (http://www.cbs.dtu.dk/services/RNAmmer/), tRNA scan-SE 2.0 server (http://trna.ucsc.edu/tRNAscan-SE/), *Klebsiella oxytoca* MLST website (https://pubmlst.org/koxytoca/), MEGA7 software (Multiple sequence alignment and phylogenetic analysis), JSpecies v1.2.1 (http://jspecies.ribohost.com/jspeciesws/)
Data format	Raw draft genome assembly and gene prediction
Experimental factors	Genomic DNA from pure culture
Experimental features	Isolation of bacteria, genome sequencing, draft genome assembly and annotation
Data source location	Ureteral stent of a 63-year-old female patient with ureter stones, Kazan, Russia
Data accessibility	Data are in public repository. This Whole Genome Shotgun (WGS) project has been deposited at DDBJ/ENA/GenBank under the accession QPKC00000000.1 (https://www.ncbi.nlm.nih.gov/nuccore/QPKC00000000.1).
Related research article	O. Tsaplina, E. Bozhokina, A. Mardanova, S. Khaitlina, Virulence factors contributing to invasive activities of *Serratia grimesii* and *Serratia proteamaculans*, Arch Microbiol. 197 (2015) 481−488. 10.1007/s00203-014-1079-7.

Value of the data

- *Klebsiella oxytoca* is recognized as a clinically significant bacterium related to healthcare-associated infections (HAIs). Therefore, understanding the genetics, physiology and virulence of *K. oxytoca* could have further applications in the control and treatment of CAUTIs.
- Data shown here can be useful for other groups working in the field of medical bacteriology. The complete genome sequence of *K. oxytoca* strain NK-1, which was isolated from the surface of ureteral stent, provides a genetic basis for understanding the epidemiology of catheter-associated pathogens.
- Data demonstrated here might be used by other researchers interested in the field of genome analysis.

1. Data

Device-associated HAIs (DA-HAIs), including catheter-associated urinary tract infections (CAUTIs), account for about 40% of all HAIs [1]. Despite *K. oxytoca* being not the most prevalent causative agent in CAUTIs, the potential role of uncommon bacteria in increasing of overall biofilm resistance to antibiotics during pathogenesis of CAUTI was previously reported [2]. Moreover, the genome of *K. oxytoca* encodes a gene homologous to the gene of grimelysin in *Serratia grimesii*, which is considered to contribute the invasion of bacteria into eukaryotic cells [3]. Therefore, the relevance in investigating the physiology and virulence factors of uncommon CAUTI-related bacteria such as *K. oxytoca* must be taken into account.

The genome coverage of *K. oxytoca* NK-1 was 113.8 $\times$. The draft genome sequence constituted 112 contigs (or 111 scaffolds), with an N_{50} (sequence length of the shortest contig at 50% of the total genome length) of 293,230 bp and L_{50} (the smallest number of contigs whose length sum makes up half of the genome size) of 7, while the largest contig spanned 618,959 bp. The resulted genome assembly of the *K. oxytoca* NK-1 contained 6,232,464 bp, with a G + C content of 55.60% (Table 1). The RAST server predicted 6246 coding sequences (CDSs). According to the approach realized within the SEED project curation of genomic data is accomplished by organizing predicted genes into subsystems based on related functionality. Such gene categorization is used in projecting onto other genome examples and allows for facilitated gene prediction and annotation. Out of all predicted in NK-1, 2156 CDSs (35%) were annotated as SEED subsystem features whereas 4090 CDSs (65%) were identified as

Table 1
Comparison of the genomic feature of *Klebsiella oxytoca* strain NK-1 with various *Klebsiella* strains isolated from human samples. The information regarding reference genomes was obtained from the ExBioCloud database [13] and NCBI database (https://www.ncbi.nlm.nih.gov).

Organism	DB accession number	Isolation source	Contigs	Genome size (bp)	G + C (%)	CDSs	r + tRNA genes
K. oxytoca strain NK-1	QPKC00000000	Ureteral stent	112	6,232,464	55.6	6246	10 + 84
K. oxytoca strain 11492-1	GCA_000252915.1	Peri-anal swab culture	213	6,176,536	54.9	5879	8 + 70
K. oxytoca strain UNM	GCA_002508265.1	Blood	108	5,980,111	54.9	5518	24 + 83
K. oxytoca strain FDAARGOS_432	GCA_002588345.1	Tracheal aspirate	10	6,434,810	54.3	6141	25 + 86
K. pneumoniae subsp. rhinoscleromatis strain ATCC 13884(TYPE)	GCA_000163455.1	Nose	51	5,450,034	55.4	5305	3 + 62
K. pneumoniae subsp. *ozaenae strain* RJF293	GCA_001530015.1	Blood	2	5,450,593	57.2	5006	25 + 79
K. pneumoniae subsp. *pneumoniae* strain blaNDM-1	GCA_000739495.1	Urine	3	5,510,332	57.3	5166	25 + 86
K. variicola strain BIDMC 61	GCA_000692995.1	Urine	7	5,580,642	57.1	5169	17 + 73
K. variicola strain MGH 76	GCA_000694735.1	Bile	16	5,960,376	51.4	5345	7 + 56
K. michiganensis strain MGH 41	GCA_000567705.1	Urine	6	6,440,856	55.5	6025	26 + 75
K. michiganensis strain 97_38	GCA_001945455.1	Blood	183	6,514,001	55.6	6137	14 + 74
K. grimontii strain 10-5250	GCA_000247915.1	Urogenital tract	3	6,121,536	55.1	5618	17 + 75
K. quasipneumoniae subsp. *similipneumoniae* strain MGH 44	GCA_000492795.1	Respiratory tract	2	5,241,816	57.7	4830	25 + 78
K. quasivariicola strain 10982	GCA_000523395.1	Peri-anal swab culture	218	6,082,330	56.6	5967	13 + 79
K. aerogenes strain CAV1320	NZ_CP011574.1	Perirectal	2	5,124,987	55.0	4939	25 + 87

outside of the SEED subsystem (Fig. 1). In general 4573 and 1673 CDSs were respectively assigned as non-hypothetical and hypothetical. There were 10 rRNAs and 84 tRNAs encoded by the genome. The strain NK-1 possessed 64 genes responsible for virulence (such as, genes necessary for adhesion, invasion and intracellular resistance), antibiotic resistance and host-mediated defense mechanisms. In addition, the RAST server identified multiple genes accountable for the formation of extracellular polysaccharides and polysaccharide capsule, which provide resistance against host defense systems. 69 genes were found to be essential for the acquisition and metabolism of iron. Out of these, 24 genes

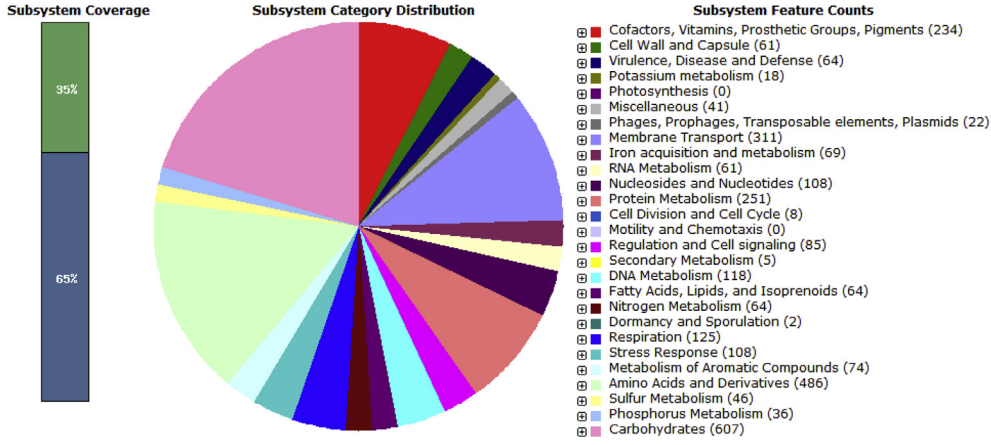

Fig. 1. Subsystem distribution of *Klebsiella oxytoca* strain NK-1 genome based on RAST annotation server [7]. The bar chart displays the ratio of genes organized in the SEED subsystems and not organized. The pie chart denotes the count of each subsystem feature.

were involved in siderophore biosynthesis, universally considered vital in conditions of iron deficiency in urine. *In silico* multilocus sequence typing (MLST) data were used to assign NK-1 strain as sequence type (ST) 216 (4, 6, 19, 10, 46, 24, 31). Comparison of the *recA* of *K. oxytoca* strain NK-1 with its closely related homologs denotes the phylogenetic position of the isolate (Fig. 2). *K. michiganensis* species is closely related to *K. oxytoca* species, which results in re-categorization of strains previously described as *K. oxytoca to K. michiganensis.* This fact implies that the NK-1 strain holds an intermediate position currently and could belong to the either the *K. oxytoca* or *K. michiganensis* species. Comparing the whole-genome sequence of NK-1 to other strains of *K. oxytoca* and *K. michiganensis* revealed its closer similarity to the *K. michiganensis* lineage than to the *K. oxytoca* species (respective ANI scores of 93.09–93.18% as against 90.79–90.94%) with the exception of the *K. oxytoca* JKo3 strain, which demonstrated the highest score in the overall nucleotide identity to the strain NK-1 (ANI score of 99.34%). Hence, more details regarding the nomenclature of the NK-1 strain is undeniably a prospective aspect for future studies.

The data on antimicrobial susceptibility testing (AST) is presented in Table 2. The MICs of aztreonam, cefepime and ceftazidime (0.06, 0.06 and 0.125 mg/L, respectively) for NK-1 did not differ significantly from their combinations with β-lactamases inhibitors – avibactam and clavulanate (aztreonam-avibactam – 0.06 mg/L; cefepime-clavulanate – 0.06 mg/L; ceftazidime-clavulanate and ceftazidime-avibactam – 0.06 and 0.125 mg/L, respectively). Thus it can be concluded that the *K. oxytoca* strain NK-1 does not produce extended-spectrum β-lactamases (ESBLs).

Out of 141 other sequenced genomes of *K. oxytoca*, only 4 strains were isolated from the human urinary tract; moreover, only one of these four was isolated from the urinary catheter. An interesting aspect of NK-1 lies in the fact that it was isolated from the biofilm on the stent retrieved from a female with stones in the ureter. Comparative analysis of genomes of the uropathogenic isolates might assist in revealing genetic features important for colonization of the urinary tract by *K. oxytoca*.

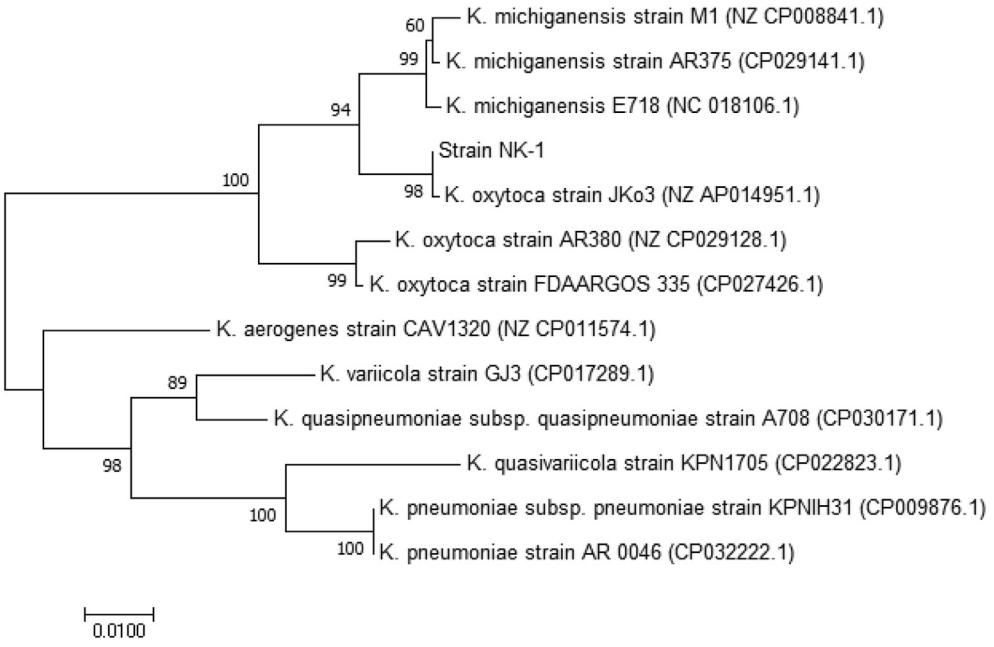

Fig. 2. A phylogenetic tree based on the *rec*A gene sequences demonstrating the relationship between *Klebsiella oxytoca* strain NK-1 and other *Klebsiella* species. Analysis was performed in MEGA7 [11] by using the Maximum Likelihood method based on the Tamura-Nei model [14]. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches.

Table 2
Antimicrobial sensitivity profile of *Klebsiella oxytoca* strain NK-1.

Antimicrobial substances	MIC, mg/L	Category
Ampicillin	16	R
Amoxicillin-clavulanic acid EUCAST	0.5	S
Piperacillin-tazobactam	4	S
Ticarcillin-clavulanic acid	2	S
Aztreonam	0.06	S
Aztreonam-Avibactam	0.06	
Cefepime	0.06	S
Cefepime-clavulanate	0.06	
Ceftazidime	0.125	S
Ceftazidime-clavulanate	0.06	
Ceftazidime-Avibactam	0.125	
Cefotaxime	0.06	S
Ertapenem	0.008	S
Imipenem	0.06	S
Meropenem	0.06	S
Doripenem	0.06	S
Ciprofloxacin	0.03	S
Chloramphenicol	2	S
Colistin	0.125	S
Amikacin	1	S
Gentamicin	0.5	S
Netilmicin	0.25	S
Tobramycin	0.5	S
Tigecycline	0.25	S
Trimethoprim-sulfamethoxazole	256	R

MIC - minimal inhibitory concentration; R - resistant; S - susceptible; combinations of third-generation cephalosporins with inhibitors of ESBL are underlined.

2. Experimental design, materials and methods

K. oxytoca NK-1 was isolated from the surface of ureteral stent obtained from the 63-year-old female diagnosed with ureter stones at the Urological Department of the University Clinic in Kazan, the Republic of Tatarstan, Russia. Identification of the bacterial strain NK-1 was carried out based on measurement of a unique molecular fingerprint using a MALDI Biotyper (Bruker Daltonik). The bacterium was cultivated in Lysogeny broth medium for 12 hours at 37 °C with aeration (200 rounds per minute). Genomic DNA of the bacterial strain was then isolated, purified and concentrated using phenol extraction and ethanol precipitation according to the protocol [4]. Concentration and purity (A260/A280) of the extracted genomic DNA were quantified with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), after which DNA was stored at −20 °C until further processing. Prior to carrying out whole-genome analysis, bacterial DNA was fragmented using Q800R2 Sonicator (Qsonica) followed by the preparation of DNA library using NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs) according to the manufacturers' instructions. Both efficiency of DNA fragmentation and DNA library preparation were controlled using a 2100 Bioanalyser (Agilent) and a High Sensitivity DNA Kit (Agilent). Using a MiSeq Reagent Kit v2PE 00 cycles (Illumina), sequencing was performed on a high-throughput Illumina MiSeq platform (Illumina) at the Joint KFU-Riken Laboratory, Kazan Federal University (Kazan, Russia). Obtained sequence read quality was assessed using FastQC version 0.11.3 software [5], the genome was assembled using SPAdes version 3.10.0 software [6]. Assembled genome feature was compared to the genomes of the other *Klebsiella* species (Table 1). The whole genome sequence was annotated with the Rapid Annotation System Technology (RAST) server [7]. The pie chart (Fig. 1) represents the subsystem distribution of the genome and the subsystem coverage. The rRNA and tRNA genes were recognized using RNAmmer-1.2 [8] and tRNA scan-SE 1.23 [9], respectively. The multilocus sequence typing of the isolate was performed using *K. oxytoca* MLST website sited at the University of Oxford [10] targeting seven housekeeping genes (*rpoB, gapA, mdh, pgi, phoE, infB*, and *tonB*). Since the 16S rRNA gene was not located within sequenced and assembled contigs, the *rec*A gene

sequence was chosen as the molecular-genetic instrument to establish the phylogenetic relationships of the isolate within the *Klebsiella* genus. Multiple sequence alignment and phylogenetic analysis based on *rec*A gene sequence were performed with the aid of the MEGA7 tool (Fig. 2) [11]. Average Nucleotide Identity (ANI) values during the comparison of the whole-genome sequences were computed with the JSpecies version 1.2.1 web server [12] based on MUMmer algorithm (ANIm). AST was performed using broth microdilution method according to ISO 20776–1:2006/GOST R ISO 20776-1-2010. Avibactam was tested in combinations with β-lactams at fixed concentration of 4 mg/l. AST results were interpreted according to EUCAST version 9.0 clinical breakpoints and Russian clinical recommendations "Susceptibility testing of microorganisms to antimicrobial agents", version 2015–02. The conclusion about the ESBL production was made on the basis of reduced sensitivity (MIC $\geq$1 mg/L) to ceftazidim, cefotaxime, cefepim or aztreonam and the presence of synergism with clavulanic acid (reduction of MIC by at least two dilutions in the presence of 4 mg/L clavulanic acid).

Acknowledgements

The work is performed in accordance with the Russian Government Program of Competitive Growth of Kazan Federal University and was funded by Russian Foundation for Basic Research according to the research project no. 18-34-00837. We are grateful to Shaidullina Elvira for the help with the preparation of antibiotic sensitivity testing.

Transparency document

Transparency document associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2019.103853.

References

[1] E. Gaid, A. Assiri, S. McNabb, W. Banjar, Device-associated nosocomial infection in general hospitals, Kingdom of Saudi Arabia, 2013-2016, J. Epidemiol. Glob. Health (1) (2018) S35–S40, https://doi.org/10.1016/j.jegh.2017.10.008.
[2] A.S. Azevedo, C. Almeida, L.F. Melo, N.F. Azevedo, Impact of polymicrobial biofilms in catheter-associated urinary tract infections, Crit. Rev. Microbiol. 43 (2017) 423–439, https://doi.org/10.1080/1040841X.2016.1240656.
[3] O. Tsaplina, E. Bozhokina, A. Mardanova, S. Khaitlina, Virulence factors contributing to invasive activities of *Serratia grimesii* and *Serratia proteamaculans*, Arch. Microbiol. 197 (2015) 481–488, https://doi.org/10.1007/s00203-014-1079-7.
[4] Y.K. Lee, H.W. Kim, C.L. Liu, H.K. Lee, A simple method for DNA extraction from marine bacteria that produce extracellular materials, J. Microbiol. Methods 52 (2003) 245–250.
[5] S. Andrews, FastQC: a Quality Control Tool for High Throughput Sequence Data, 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. (Accessed 26 April 2010).
[6] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, et al., SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (2012) 455–477, https://doi.org/10.1089/cmb.2012.0021.
[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 Genomics 9 (2008) 75, https://doi.org/10.1186/1471-2164-9-75.
[8] K. Lagesen, P. Hallin, E.A. Rødland, H.H. Staerfeldt, T. Rognes, D.W. Ussery, RNAmmer: consistent and rapid annotation of ribosomal RNAgenes, Nucleic Acids Res. 35 (2007) 3100–3108.
[9] T.M. Lowe, S.R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, Nucleic Acids Res. 25 (1997) 955–964.
[10] K.A. Jolley, J.E. Bray, M.C.J. Maiden, Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; referees: 2 approved], Wellcome Open Res 3 (2018) 124, https://doi.org/10.12688/wellcomeopenres.14826.1.
[11] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol. 33 (2016) 1870–1874, https://doi.org/10.1093/molbev/msw054.
[12] M. Richter, R. Rosselló-Móra, F.O. Glöckner, J. Peplies, JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison, Bioinformatics (2016), https://doi.org/10.1093/bioinformatics/btv681.
[13] 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, https://doi.org/10.1099/ijsem.0.001755.
[14] K. Tamura, M. Nei, Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees, Mol. Biol. Evol. 10 (1993) 512–526.