AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequences of Gram-Positive Opportunistic Pathogens Isolated from Hospitals in Almaty, Kazakhstan

Ilya S. Korotetskiy,[a] Ardak B. Jumagaziyeva,[a] Bahkytzhan Kerimzhanova,[a] Oleg N. Reva,[b] Sergey V. Shilov,[a] Tatyana Kuznetsova,[a] Ludmila Ivanova,[a] Nadezhda Korotetskaya,[a] Aimana Bekmukhamedova,[c] Ganiya Satylgankyzy,[c] Nailya G. Klivleyeva[d]

[a]Scientific Center for Anti-Infectious Drugs, Almaty, Kazakhstan
[b]Centre for Bioinformatics and Computational Biology, Department of Biochemistry, Genetics, and Microbiology, University of Pretoria, Pretoria, South Africa
[c]Department of Vascular Surgery, National Scientific Center of Surgery named after Syzganov, Almaty, Kazakhstan
[d]The Research and Production Center for Microbiology and Virology, Almaty, Kazakhstan

ABSTRACT The appearance of drug-resistant pathogens reduces the therapeutic applicability of antibiotics and increases the rate of hospital infections among patients. Complete genome sequences of four Gram-positive clinical isolates of *Streptococcus* and *Staphylococcus* were obtained and analyzed to serve as model microorganisms for further studies on drug-induced antibiotic resistance reversion.

Antimicrobial resistance (AMR) is one of the main threats to human health. Monitoring the spread of genetic determinants and keeping under control the etiological and taxonomical composition of nosocomial pathogens will allow timely identification of threats to human health (1).

Streptococcus pneumoniae, *Staphylococcus epidermidis*, and two strains of *Staphylococcus aureus* were isolated in 2021 in the Department of Vascular Surgery of Syzganov's National Scientific Center of Surgery in Almaty, Kazakhstan. Isolates were obtained by plating from biological material on selective and differential-diagnostic media (Table 1). This study was approved by the Committee of Institutional Animal Care and Use of the Scientific Center for Anti-Infectious Drugs (SCAID), Almaty, Kazakhstan (2). For more details on the isolates, see NCBI BioProject number PRJNA754843. For DNA extraction, cultures were grown on nutrient agar (HiMedia) for 24 h at 37°C. DNA was extracted using a PureLink genomic DNA minikit (Invitrogen, USA). DNA was sheared using the Megaruptor shearing kit 3. The DNA library was prepared using a PacBio SMRTbell Express template prep kit 2.0. SMRTbell templates were annealed using a Sequel binding and internal control kit 3.0. The Sequel sequencing kit 3.0 and SMRT cell 1M v3 tray were used for sequencing. For each SMRT cell (Pacific Biosciences), 600-min movies were captured using the PacBio Sequel-I sequencing platform by Macrogen (Seoul, South Korea). Statistical parameters of the generated reads are shown in Table 1. Default settings were used for all software unless otherwise noted. DNA reads were quality controlled and checked for remaining adapters using LongQC 1.2.0c (3). Assembly and circularization of contigs were performed using Canu 2.0 (4). Plasmid contigs were identified using Platon 1.6 (5). Contigs were joined using MeDuSa at http://combo.dbe.unifi.it/medusa (6) using the most similar reference genomes identified in GenBank using BLASTN (Table 1). The consensus sequences were annotated on the RAST server (https://rast.nmpdr.org/) using the RASTtk algorithm (7) with the fix frameshifts setting. Chromosomal sequences were edited with Artemis 14.0.0 (8) to start with *dnaA* on the positive strand, and the plasmid sequences were shifted for 50 kb to check circularization by mapping the initial PacBio reads against the obtained contigs using pbmm2 (SMRT-Link 10.10.119588). The final consensus sequences were

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

Address correspondence to Ilya S. Korotetskiy, laeda81@gmail.com.

The authors declare no conflict of interest.

Received 1 February 2022
Accepted 26 February 2022
Published 14 March 2022

TABLE 1 Deposited sequences of complete genomes of Gram-positive isolates[a]

Strain name	Sample type	Media	Total reads/N_{50} (bp)	Coverage (×)	Replicon	Length (bp)	GC content (%)	Resistance[b]	MLST/ serotype[c]	Reference genome (GenBank accession no.)	GenBank accession no.	SRA no.
Streptococcus pneumoniae SCAID PHRX1-2021	Swab from pharynx	CHROMagar Orientation, Endo agar	149,806/4,594	386	Chromosome	2,098,200	39.72	OX, AMP, E	ST377	LR216047	CP082820	SRR15674694
Staphylococcus epidermidis SCAID OTT1-2021 (597)	Swab from ear	CHROMagar Orientation, Endo agar	209,855/10,195	405	Chromosome	2,099,244	32.35	AMX, E, AMP, AZM	ND	CP028282	CP082816	SRS9980277
				405	Plasmid 1	24,456	27.9			MW364979	CP082817	
				405	Plasmid 2	24,520	29.15			CP030247	CP082818	
				405	Plasmid 3	13,203	28.99			LR735442	CP082819	
Staphylococcus aureus SCAID OTT1-2021 (597/2)	Swab from ear	CHROMagar Orientation, Endo agar	202,151/10,148	804	Chromosome	2,737,085	32.97	AMX, OX, AMP	ST508	CP047795	CP082813	SRS9980279
				804	Plasmid	33,923	28.19			CP029650	CP082814	
Staphylococcus aureus SCAID WND1-2021 (598)	Swab from wound	CHROMagar Orientation, Cetrimide agar	200,444/10,267	654	Chromosome	2,889,511	32.83	AMX, AMP	ST30	CP047791	CP082815	SRS9980278

[a]The susceptibility was evaluated by the disk diffusion method in Mueller-Hinton agar (HiMedia, India). The results of the threshold inhibition zones were evaluated according to the CLSI.
[b]AMP, ampicillin; AMX, amoxicillin; E, erythromycin; OX, oxacillin, intermediate resistance. The resistance to antibiotics was determined experimentally.
[c]ND, not detected.

generated from the alignments of mapped reads, annotated using GenBank PGAP, and deposited at the NCBI (Table 1). Multilocus sequence typing (MLST) was performed using the Bacterial Isolate Genome Sequence Database (BIGSdb) (https://bigsdb.pasteur.fr/) and Copenhagen Business School (CBS) (www.cbs.dtu.dk/services/MLST) databases (9, 10).

The MLST of *Staphylococcus epidermidis* strain SCAID OTT1-2021 (597) could not be identified because the *apoE* gene variant of this isolate did not match any known variants. *S. aureus* isolates belong to sequence type 508 (ST508) (strain 597/2) and ST30 (strain 598). The *S. pneumoniae* strain belongs to ST377.

Data availability. The genomes are available from NCBI BioProject number PRJNA754843 under the accession numbers shown in Table 1.

ACKNOWLEDGMENT

Sequencing was funded by grant BR09458960 of the program "Study of Reversion of Antibiotic Resistance of Pathogenic Microorganisms" provided by the Industrial Development Committee of the Ministry of Industry and Infrastructure Development of the Republic of Kazakhstan.

REFERENCES

1. Joubert M, Reva ON, Korotetskiy IS, Shvidko SV, Shilov SV, Jumagaziyeva AB, Kenesheva ST, Suldina NA, Ilin AI. 2019. Assembly of complete genome sequences of negative-control and experimental strain variants of *Staphylococcus aureus* ATCC BAA-39 selected under the effect of the drug FS-1, which induces antibiotic resistance reversion. Microbiol Resour Announc 8:e00579-19. https://doi.org/10.1128/MRA.00579-19.
2. Korotetskiy IS, Jumagaziyeva AB, Kerimzhanova B, Reva ON, Shilov SV, Kuznetsova T, Zubenko N, Korotetskaya N, Bekmukhamedova A, Satylgankyzy G, Klivleyeva NG. 2021. Complete genome sequences of Gram-negative opportunistic pathogens isolated in hospitals in Almaty, Kazakhstan. Microbiol Resour Announc 10:e00974-21. https://doi.org/10.1128/MRA.00974-21.
3. Fukasawa Y, Ermini L, Wang H, Carty K, Cheung MS. 2020. LongQC: a quality control tool for third generation sequencing long read data. G3 (Bethesda) 10:1193–1196. https://doi.org/10.1534/g3.119.400864.
4. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.
5. Schwengers O, Barth P, Falgenhauer L, Hain T, Chakraborty T, Goesmann A. 2020. Platon: identification and characterization of bacterial plasmid contigs in short-read draft assemblies exploiting protein sequence-based replicon distribution scores. Microb Genom 6:mgen000398. https://doi.org/10.1099/mgen.0.000398.
6. Bosi E, Donati B, Galardini M, Brunetti S, Sagot MF, Lió P, Crescenzi P, Fani R, Fondi M. 2015. MeDuSa: a multi-draft based scaffolder. Bioinformatics 31:2443–2451. https://doi.org/10.1093/bioinformatics/btv171.
7. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.
8. Carver T, Berriman M, Tivey A, Patel C, Böhme U, Barrell BG, Parkhill J, Rajandream MA. 2008. Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. Bioinformatics 24:2672–2676. https://doi.org/10.1093/bioinformatics/btn529.
9. Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595. https://doi.org/10.1186/1471-2105-11-595.
10. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, Lund O. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol 50:1355–1361. https://doi.org/10.1128/JCM.06094-11.