AMERICAN SOCIETY FOR MICROBIOLOGY | **Microbiology**® Resource Announcements

Complete Genome Sequences of Three Uropathogenic *Klebsiella quasipneumoniae* Strains Isolated from Postmenopausal Women with Recurrent Urinary Tract Infection

Sundharamani Venkitapathi,[a] Belle M. Sharon,[a] Tahira A. Ratna,[a] Amanda P. Arute,[a] Philippe E. Zimmern,[b] ⓘ Nicole J. De Nisco[a]

[a]Department of Biological Sciences, University of Texas at Dallas, Richardson, Texas, USA
[b]Department of Urology, University of Texas Southwestern Medical Center, Dallas, Texas, USA

Sundharamani Venkitapathi and Belle M. Sharon contributed equally to this article. Author order was determined by seniority.

ABSTRACT Recurrent urinary tract infection (rUTI) poses a major health issue, especially among postmenopausal women. We report complete genome sequences of three *Klebsiella quasipneumoniae* strains isolated from the urine of postmenopausal women with rUTI. *K. quasipneumoniae* is a recently identified *Klebsiella* species with clinical and virulence characteristics distinct from those of *K. pneumoniae*.

Recurrent urinary tract infection (rUTI), defined as two UTI episodes within 6 months or three within 12 months, poses a major health issue (1, 2). The genus *Klebsiella* is a leading cause of rUTI, accounting for 15% to 17% of cases (3). *Klebsiella quasipneumoniae* was originally associated with environmental niches; however, recent reports suggest a role in human infection (4). Lack of tests distinguishing *K. quasipneumoniae* from *Klebsiella pneumoniae* in clinical settings has led to a likely underestimation of *K. quasipneumoniae* prevalence (5).

Complete genome assemblies of uropathogenic *K. quasipneumoniae* isolates allow investigation of virulence and metabolic traits specific to this species. We report the complete genome sequences of three *K. quasipneumoniae* strains (Table 1) isolated from the urine of postmenopausal women meeting the criteria for uncomplicated rUTI as part of an institutional review board-approved study (STU032016-006, MR17-120) (6).

Clean-catch midstream urine was obtained from three postmenopausal women, plated onto CHROMagar Orientation medium (BD), and incubated overnight at 37°C. Isolated single colonies were chosen for genus identification by PCR amplification and Sanger sequencing of the 16S rRNA gene, followed by a MegaBLAST query (BLAST v2.10.0) (6, 7) against the nonredundant/nucleotide (nr/nt) database. Species identification was performed by PCR using primers specific for the *K. quasipneumoniae* beta-lactamase (*bla*) and deoxyribose regulator (*deoR*) genes (8). *K. quasipneumoniae* genomic DNA (gDNA) was extracted from overnight cultures grown at 37°C in Luria broth (LB) using the gDNA extraction kit (BioBasic), followed by quality assessment using a 260/280-nm absorbance ratio and agarose gel electrophoresis. gDNA was sequenced using the Illumina NextSeq 500 system and the Oxford Nanopore Technologies (ONT) MinION platform. For Illumina sequencing, libraries were prepared using the Nextera DNA Flex library prep kit and sequenced to generate 2×150-bp paired-end reads. Illumina reads were quality assessed and trimmed using CLC Genomics Workbench v12.0.3 with cutoffs set at a minimum Phred score of 20 and a read length of 15 bp. Default parameters were used for all software unless otherwise specified. ONT libraries were prepared using a ligation sequencing kit (SQK-LSK109) and barcode expansion kit 13-24 (EXP-NBD114) and sequenced on R9 FLO-MIN106 flow cells. ONT MinKNOW

TABLE 1 Characteristics, assembly parameters, and accession numbers of three complete uropathogenic *Klebsiella quasipneumoniae* genome sequences

Strain	BioSample accession no.	SRA accession no.[a]	No. of raw reads	No. of trimmed reads	Read N_{50} (bp)	Read depth (×)	MLST[b]	GenBank accession no.	Type of contig (circular)	Total length (bp)	GC content (%)	No. of CDSs[c]	Plasmid replicon(s)
KqPF9	SAMN17016746	SRX9774966 (O)	150,239	148,741	11,189	153	UN[d]	CP065841	Chromosome	5,272,842	58.1	5,137	NA[e]
		SRX9779643 (I)	16,937,480	16,068,400		359		CP065842	Plasmid	399,394	48.3	435	IncFIB
								CP065843	Plasmid	4,730	42.6	6	Col(pHAD28)
								CP065844	Plasmid	4,096	55.5	5	Col440I
								CP065845	Plasmid	4,000	46.3	4	Col440I
KqPF26	SAMN17016750	SRX9774961 (O)	26,486	26,344	13,446	39	3387	CP065838	Chromosome	5,242,686	58	5,039	NA
		SRX9779644 (I)	12,077,320	11,492,042		271		CP065839	Plasmid	144,959	52.5	152	IncFIB(K), IncFII(K)
								CP065840	Plasmid	3,478	45.7	6	UN
KqPF42	SAMN17016751	SRX9774962 (O)	125,656	124,662	11,819	147	1535	CP065846	Chromosome	5,278,208	58	5,092	NA
		SRX9779645 (I)	18,822,828	17,989,506		420		CP065847	Plasmid	223,863	50.9	252	IncFIB(K)

[a] O, ONT; I, Illumina.
[b] MLST, multilocus sequence type.
[c] CDSs, coding sequences.
[d] UN, unknown.
[e] NA, not applicable.

v3.6.16 was used for base calling, demultiplexing, and barcode trimming. ONT read quality was assessed with NanoStats v1.2.0 (9), and reads were trimmed with NanoFilt v2.6.0 (9). Reads with a Phred score of >7 and a length of >200 bp were retained (Table 1).

The Illumina and ONT reads were used to construct hybrid assemblies of each strain, and the circular genomic sequences were rotated to the starting base of *dnaA* or *repA*, if present, using Unicycler v0.4.8 (SPAdes v3.13.0, Racon v1.4.10, and Pilon v1.23) (10–13). The quality of the hybrid assembly was evaluated using QUAST v5.0.2 (14), and the genome completeness was assessed with Bandage v0.8.1 (15) and BUSCO v1 (16) using the bacterial ortholog set on the gVolante server v1.2.1 (17). The NCBI Prokaryotic Genome Annotation Pipeline v4.11 was used for genome annotation (18, 19). The GC content and coding sequences were evaluated using Geneious Prime v2020.0.5. The sequence type was determined using MLST v2.0 (http://www.genomicepidemiology.org/) with the *K. pneumoniae* configuration (20). The plasmid replicons were identified with PlasmidFinder v2.1 (21, 22), using the *Enterobacteriaceae* database and default cutoffs (Table 1).

Data availability. The sequencing data were submitted to GenBank under BioProject accession number PRJNA683049. The BioSample and SRA accession numbers are reported in Table 1.

REFERENCES

1. Scholes D, Hooton TM, Roberts PL, Stapleton AE, Gupta K, Stamm WE. 2000. Risk factors for recurrent urinary tract infection in young women. J Infect Dis 182:1177–1182. https://doi.org/10.1086/315827.
2. Foxman B. 1990. Recurring urinary tract infection: incidence and risk factors. Am J Public Health 80:331–333. https://doi.org/10.2105/ajph.80.3.331.
3. Cristea OM, Avrămescu CS, Bălăşoiu M, Popescu FD, Popescu F, Amzoiu MO. 2017. Urinary tract infection with Klebsiella pneumoniae in patients with chronic kidney disease. Curr Health Sci J 43:137–148. https://doi.org/10.12865/CHSJ.43.02.06.
4. Brisse S, Passet V, Grimont PAD. 2014. Description of Klebsiella quasipneumoniae sp. nov., isolated from human infections, with two subspecies, Klebsiella quasipneumoniae subsp. quasipneumoniae subsp. nov. and Klebsiella quasipneumoniae subsp. similipneumoniae subsp. nov., and demonstration that Klebsiella singaporensis is a junior heterotypic synonym of Klebsiella variicola. Int J Syst Evol Microbiol 64:3146–3152. https://doi.org/10.1099/ijs.0.062737-0.
5. Rodrigues C, Passet V, Rakotondrasoa A, Brisse S. 2018. Identification of Klebsiella pneumoniae, Klebsiella quasipneumoniae, Klebsiella variicola and related phylogroups by MALDI-TOF mass spectrometry. Front Microbiol 9:3000. https://doi.org/10.3389/fmicb.2018.03000.
6. De Nisco NJ, Neugent M, Mull J, Chen L, Kuprasertkul A, de Souza Santos M, Palmer KL, Zimmern P, Orth K. 2019. Direct detection of tissue-resident bacteria and chronic inflammation in the bladder wall of postmenopausal women with recurrent urinary tract infection. J Mol Biol 431:4368–4379. https://doi.org/10.1016/j.jmb.2019.04.008.
7. Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. J Comput Biol 7:203–214. https://doi.org/10.1089/10665270050081478.
8. Fonseca EL, Ramos NDV, Andrade BGN, Morais LLCS, Marin MFA, Vicente ACP. 2017. A one-step multiplex PCR to identify Klebsiella pneumoniae, Klebsiella variicola, and Klebsiella quasipneumoniae in the clinical routine. Diagn Microbiol Infect Dis 87:315–317. https://doi.org/10.1016/j.diagmicrobio.2017.01.005.
9. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149.
10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
11. Vaser R, Sovic I, Nagarajan N, Sikic M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.
12. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
13. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
14. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
15. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31:3350–3352. https://doi.org/10.1093/bioinformatics/btv383.
16. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.
17. Nishimura O, Hara Y, Kuraku S. 2017. gVolante for standardizing completeness assessment of genome and transcriptome assemblies. Bioinformatics 33:3635–3637. https://doi.org/10.1093/bioinformatics/btx445.
18. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851–D860. https://doi.org/10.1093/nar/gkx1068.
19. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.

20. 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.

21. Carattoli A, Hasman H. 2020. PlasmidFinder and in silico pMLST: identification and typing of plasmid replicons in whole-genome sequencing (WGS). Methods Mol Biol 2075:285–294. https://doi.org/10.1007/978-1-4939-9877-7_20.

22. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903. https://doi.org/10.1128/AAC.02412-14.