




# Genome Sequence of *Klebsiella pneumoniae* YBQ, a Clinical Strain Isolated from the Sputum of a Patient with Severe Pneumonia

Jintao Zou,<sup>a</sup> Zhifu Chen,<sup>a</sup> Xiaoli Zhang,<sup>b</sup> Yue Yuan,<sup>a</sup> Haiming Jing,<sup>a</sup> Quanming Zou,<sup>a</sup>  Jinyong Zhang<sup>a</sup>

<sup>a</sup>National Engineering Research Center of Immunological Products, Department of Microbiology and Biochemical Pharmacy, College of Pharmacy, Third Military Medical University, Chongqing, People's Republic of China

<sup>b</sup>Department of Clinical Hematology, College of Pharmacy, Third Military Medical University, Chongqing, People's Republic of China

Jintao Zou and Zhifu Chen contributed equally to this work. Author order was determined in order of increasing seniority.

**ABSTRACT** We report here the genome of *Klebsiella pneumoniae* YBQ, a clinical strain isolated from the sputum of a patient with acute *Klebsiella pneumoniae* infection. The genome consists of a 5,119,471-bp circular chromosome and a 184,347-bp plasmid. Genome annotation predicted 5,028 coding DNA sequences (CDSs), 84 tRNAs, 25 rRNAs, and 47 small RNAs (sRNAs).

*Klebsiella pneumoniae* is an important opportunistic pathogen that frequently causes life-threatening infections, including pneumonia, bacteremia, and urinary tract infection (1). It accounts for up to 10% of all nosocomial bacterial infections (2) and is associated with high mortality (3, 4). Recently, due to the steadily increasing frequency of the subset of highly invasive and hypervirulent *K. pneumoniae* (hvKp) and carbapenem-resistant *K. pneumoniae* (CRKP) strains (5), infections caused by this bacterium have become a major threat to public health.

hvKp usually exhibits a hypermucoviscous phenotype, which is associated with high expression of extracellular polysaccharides (6). Recently, we isolated a clinical strain of *K. pneumoniae*, named YBQ, from the sputum of a pneumonia patient with acute and severe *K. pneumoniae* infection (ethics approval granted by the Ethical and Experimental Committee of the Army Medical University). A string test indicated that YBQ exhibits a hypermucoviscous phenotype (7). Animal challenge studies indicate that YBQ is a strain with high virulence (8). Here, we report the complete genome sequence of YBQ, which may be essential for further study of the virulence and pathogenic mechanism of this isolate.

A single colony of YBQ was picked and grown for 6 h in 20 ml LB medium at 37°C. Genomic DNA was isolated using the DNeasy blood and tissue kit (Qiagen), evaluated with gel electrophoresis, and quantified with a 2100 NanoDrop spectrophotometer. A SMRTbell template was used to establish a standard library of 20-kb fragments, and the genome was sequenced by single-molecule real-time (SMRT) sequencing using the PacBio RS II platform. PacBio subreads less than 1 kb long were removed. Draft genomic unitigs, which are uncontested groups of fragments, were assembled using Canu v1.5 (<https://github.com/marbl/canu/releases/tag/v1.5>). The total number of reads sequenced with the PacBio SMRT platform is 8,764,210, with 158,211 subreads and a subread  $N_{50}$  value of 12,359 bp. To improve the accuracy of the genome sequences, the GATK (<https://www.broadinstitute.org/gatk/>) and SOAP (SOAP2, SOAPsnp, SOAPindel) tool packages were used to make single-base corrections. To trace the presence of any plasmid, the filtered Illumina reads were mapped using SOAP to the bacterial plasmid database (<http://www.ebi.ac.uk/genomes/plasmid.html>).

**Citation** Zou J, Chen Z, Zhang X, Yuan Y, Jing H, Zou Q, Zhang J. 2020. Genome sequence of *Klebsiella pneumoniae* YBQ, a clinical strain isolated from the sputum of a patient with severe pneumonia. *Microbiol Resour Announc* 9:e01223-20. <https://doi.org/10.1128/MRA.01223-20>.

**Editor** John J. Dennehy, Queens College

**Copyright** © 2020 Zou et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Jinyong Zhang, [zhangjy198217@126.com](mailto:zhangjy198217@126.com).

**Received** 26 October 2020

**Accepted** 28 October 2020

**Published** 19 November 2020

**TABLE 1** Statistical results of the functional annotation of the genome of YBQ

Database	No. (%) of genes annotated
KEGG	3,658 (72.75)
InterPro	4,567 (90.83)
SwissProt	3,617 (71.93)
COG	4,271 (84.94)
GO	3,530 (70.2)
NR	5,021 (99.86)
Total	5,023 (99.9)

Gene prediction was performed on the YBQ genome assembly by GLIMMER v3.02 with hidden Markov models. tRNA, rRNA, and small RNA (sRNA) recognition was performed using tRNAscan-SE v1.3.1 (9), RNAmmer v1.2 (10), and Rfam v9.1 (11). CRISPRfinder v4.2.19 was used for CRISPR identification (12). The final assemblies of YBQ resulted in a 5,119,471-bp circular chromosome and a 184,347-bp circular plasmid, with mean G+C contents of 57.67% and 50.47%, respectively. The total genes predicted for YBQ contain 5,028 CDSs, 84 tRNAs, 25 rRNAs, and 47 sRNAs, along with 2 CRISPR arrays. In addition, 118 tandem repeats were obtained using the Tandem Repeat Finder v4.04 (13); 58 minisatellite DNAs and 9 microsatellite DNAs were selected based on the number and length of repeat units.

Seven databases, including KEGG (14), Clusters of Orthologous Groups (COG) (15), nonredundant protein database (NR), Swiss-Prot (16), Gene Ontology (GO) (17), TrEMBL (18), and eggNOG (19), were used for general function annotation. The summary of the functional annotation of the genome is listed in Table 1.

**Data availability.** The annotated complete genome assembly of *K. pneumoniae* YBQ was deposited in the SRA. The BioProject accession number is [PRJNA663465](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA663465), and the SRA accession number is [SRX9123846](https://www.ncbi.nlm.nih.gov/sra/SRX9123846).

## ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (number 31970138) and Chongqing Natural Science Foundation (number cstc2019jcyj-msxmX0377).

## REFERENCES

- Vading M, Naucler P, Kalin M, Giske CG. 2018. Invasive infection caused by *Klebsiella pneumoniae* is a disease affecting patients with high comorbidity and associated with high long-term mortality. *PLoS One* 13:e0195258. <https://doi.org/10.1371/journal.pone.0195258>.
- Struve C, Krogfelt KA. 2004. Pathogenic potential of environmental *Klebsiella pneumoniae* isolates. *Environ Microbiol* 6:584–590.
- Jung Y, Lee MJ, Sin H-Y, Kim N-H, Hwang J-H, Park J, Choe PG, Park WB, Kim ES, Park S-W, Park KU, Kim HB, Kim N-J, Kim E-C, Song K-H, Oh M-D. 2012. Differences in characteristics between healthcare-associated and community-acquired infection in community-onset *Klebsiella pneumoniae* bloodstream infection in Korea. *BMC Infect Dis* 12:239. <https://doi.org/10.1186/1471-2334-12-239>.
- Thom KA, Schweizer ML, Osih RB, McGregor JC, Furuno JP, Perencevich EN, Harris AD. 2008. Impact of empiric antimicrobial therapy on outcomes in patients with *Escherichia coli* and *Klebsiella pneumoniae* bacteremia: a cohort study. *BMC Infect Dis* 8:116. <https://doi.org/10.1186/1471-2334-8-116>.
- Shon AS, Bajwa RPS, Russo TA. 2013. Hypervirulent (hypermucoviscous) *Klebsiella pneumoniae*: a new and dangerous breed. *Virulence* 4:107–118. <https://doi.org/10.4161/viru.22718>.
- Choby JE, Howard-Anderson J, Weiss DS. 2020. Hypervirulent *Klebsiella pneumoniae*—clinical and molecular perspectives. *J Intern Med* 287:283–300. <https://doi.org/10.1111/joim.13007>.
- Feldman MF, Mayer Bridwell AE, Scott NE, Vinogradov E, McKee SR, Chavez SM, Twentyman J, Stallings CL, Rosen DA, Harding CM. 2019. A promising bioconjugate vaccine against hypervirulent *Klebsiella pneumoniae*. *Proc Natl Acad Sci U S A* 116:18655–18663. <https://doi.org/10.1073/pnas.1907833116>.
- Liu D, Chen Z, Yuan Y, Jing H, Zou J, Zhang X, Zeng X, Zhang W, Zou Q, Zhang J. 2020. Innate immune effectors play essential roles in acute respiratory infection caused by *Klebsiella pneumoniae*. *J Immunol Res* 2020:1–11. <https://doi.org/10.1155/2020/5291714>.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964. <https://doi.org/10.1093/nar/25.5.955>.
- Lagesen K, Hallin P, Rodland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <https://doi.org/10.1093/nar/gkm160>.
- Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, Wilkinson AC, Finn RD, Griffiths-Jones S, Eddy SR, Bateman A. 2009. Rfam: updates to the RNA families database. *Nucleic Acids Res* 37:D136–D140. <https://doi.org/10.1093/nar/gkn766>.
- Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRfinder: a Web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35:W52–W57. <https://doi.org/10.1093/nar/gkm360>.
- Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580. <https://doi.org/10.1093/nar/27.5.573>.
- Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as

- a reference resource for gene and protein annotation. *Nucleic Acids Res* 44:D457–D462. <https://doi.org/10.1093/nar/gkv1070>.
15. Galperin MY, Makarova KS, Wolf YI, Koonin EV. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43:D261–D269. <https://doi.org/10.1093/nar/gku1223>.
  16. The UniProt Consortium. 2015. UniProt: a hub for protein information. *Nucleic Acids Res* 43:D204–D212. <https://doi.org/10.1093/nar/gku989>.
  17. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene Ontology: tool for the unification of biology. *Nat Genet* 25:25–29. <https://doi.org/10.1038/75556>.
  18. Bairoch A, Apweiler R. 1997. The SWISS-PROT protein sequence data bank and its supplement TrEMBL. *Nucleic Acids Res* 25:31–36. <https://doi.org/10.1093/nar/25.1.31>.
  19. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 44:D286–D293. <https://doi.org/10.1093/nar/gkv1248>.