Whole-Genome Sequencing of Klebsiella pneumoniae BASUSDALSc45PDB48, a Unique Strain Capable of Growing in Pesticide-Containing Medium, Isolated from Soil in Bangladesh

Atikur Rahman, Sadiya Tahsin, Salek Ahmed Sajib, Mohammad Tanbir Habib, Khaled Mahmud Sujon, Khandaker Md, Khalid-Bin Ferdaus, K M F Hoque, Zennat Ferdousi

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Whole-Genome Sequencing of *Klebsiella pneumoniae* BASUSDALSc45PDB48, a Unique Strain Capable of Growing in Pesticide-Containing Medium, Isolated from Soil in Bangladesh

Md Atikur Rahman,[a] Sadiya Tahsin,[a] Salek Ahmed Sajib,[a]* Mohammad Tanbir Habib,[a] Khaled Mahmud Sujon,[a] Khandaker Md Khalid-Bin Ferdaus,[a] K. M. F. Hoque,[a] Zennat Ferdousi,[a] Md Abu Reza[a]

[a]Molecular Biology and Protein Science Laboratory, Department of Genetic Engineering and Biotechnology, University of Rajshahi, Rajshahi, Bangladesh

ABSTRACT We report the draft genome sequence of *Klebsiella pneumoniae* strain BASUSDALSc45PDB48, isolated from pesticide-contaminated soil; this strain showed the ability to grow in a medium with cypermethrin as the only carbon source. The genome assembly comprised 5,249,704 bp, with 128.17 Ns per 100 kbp, an N_{50} value of 5,035,968 bp, a GC content of 57.57%, and 5,349 annotated genes.

Klebsiella pneumoniae is a Gram-negative, encapsulated, nonmotile, rod-shaped bacterium. It can be found in soil, surface waters, human skin, mouth, intestine, etc. (1, 2). About 30% of all *K. pneumoniae* strains can fix nitrogen (3). They can also degrade pesticides and thus have high potential for practical application in bioremediation (4).

K. pneumoniae strain BASUSDALSc45PDB was isolated from pesticide-contaminated soil in Netrokona (24.87947 N, 90.70843 E), Bangladesh. The soil (1 g) was added to a 250-ml Erlenmeyer flask containing 100 ml minimal salt (MS) medium supplemented with cypermethrin (500 μg/ml), a commonly used pesticide in Bangladesh. The flask was incubated for 7 days on a rotary shaker (120 rpm) at 37°C. The enrichment culture was streaked onto agar-solidified minimal salt medium (HiMedia, India) supplemented with cypermethrin (500 μg/ml) and incubated at 37°C for 36 h. Finally, a single colony was isolated, and genomic DNA was extracted using the TIANamp bacteria DNA kit (Tiangen Biotech Co. Ltd., Beijing, China) according to the manufacturer's protocol. Whole-genome sequencing (WGS) and 16S rRNA sequencing were conducted by Invent Technologies Ltd. (Dhaka, Bangladesh). To identify the bacterial strain, the 16S rRNA region was amplified using the forward primer 8F (GAGTTTGATCCTGGCTCAG) and reverse primer 806R (GGACTACHVGGGTWTCTAAT). For taxonomic identification, the 16S sequence was queried against the nonredundant (nr)/nucleotide (nt) database using BLASTN (5) (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

For WGS, a genomic library was constructed for 151-bp paired-end sequencing using the Illumina DNA Prep library preparation kit (catalog number 20018705; San Diego, CA, USA) following the manufacturer's protocol. The prepared library was sequenced on the Illumina MiSeq platform. Genome sequencing yielded 1,410,224 reads overall, exceeding 52× coverage. The reads were quality checked using FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). No adapter sequences were found. Quality control of the raw sequencing data was performed using Trimmomatic v0.39 with default parameters except SLIDINGWINDOW:4:2 (http://www.usadellab.org/cms/?page=trimmomatic) (6). *De novo* assembly using SPAdes v3.14.1 (https://cab.spbu.ru/software/spades) (7) initially yielded 565 contigs with an N_{50} value of 147,594 bp. For the SPAdes assembly, the coverage threshold was calculated automatically using "--cov-cutoff auto," and the k-mer sizes were set to "-k 21,33,55,77," while the "--careful" option was selected to minimize the number of mismatches in the final contigs.

Editor Leighton Pritchard, SIPBS, University of Strathclyde

Address correspondence to Md Abu Reza, reza.gen@ru.ac.bd.

*Present address: Salek Ahmed Sajib, Université Paris-Saclay, CNRS, INRAE, Université Evry, Université de Paris, Institute of Plant Sciences Paris-Saclay (IPS2), Bat 630, Gif sur Yvette, France.

The authors declare no conflict of interest.

Received 9 July 2021
Accepted 22 November 2021
Published 9 December 2021

The resulting assembly was further mapped to the reference genome *K. pneumoniae* subsp. *pneumoniae* HS11286 (GenBank accession number NC_016845.1) using the scaffolding software RagTag v1.0.1 (https://github.com/malonge/RagTag), which produced 128.17 Ns per 100 kbp (6,600 Ns in total) and 420 scaffolds, with an N_{50} value of 5,035,968 bp (8). PlasmidFinder v2.0 (https://cge.cbs.dtu.dk/services/PlasmidFinder/) identified contig 288 as a single complete plasmid (IncN) (9). Default parameters were used for all software unless otherwise specified. The draft genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (10, 11). After scaffolding and gap estimation, the *Klebsiella pneumoniae* strain BASUSDALSc45PDB48 genome assembly size was 5,256,304 bp, with 5,249,704 bp of ungapped sequences (GC content, 57.57%), containing 5,195 protein coding genes, 70 tRNAs, 3 rRNAs, 16 noncoding RNAs, and a single plasmid.

Data availability. The *K. pneumoniae* strain BASUSDALSc45PDB48 genome sequence has been deposited at GenBank under accession numbers JAEKPA000000000.1 (genome) and JAEKPA010000277.1 (plasmid), BioProject accession number PRJNA685932, BioSample accession number SAMN17101190, and SRA accession number SRR13447672.

ACKNOWLEDGMENT

We thank the Bangladesh Academy of Sciences for support in the form of a project grant from the BAS-USDA Endowment Fund (project number BAS-USDA LSc 45, 2017 to 2020), as well as for constructive guidance.

REFERENCES

1. Bagley ST. 1985. Habitat association of Klebsiella species. Infect Control 6:52–58. https://doi.org/10.1017/s0195941700062603.
2. Ryan KJ, Ray CG, Ahmad N, Drew WL, Plorde JJ. 2010. Haemophilus and Bordetella, p 551–564. *In* Ryan KJ, Ray CG (ed), Sherris medical microbiology. McGraw-Hill, New York, NY.
3. Postgate J. 1998. Nitrogen fixation, 3rd ed. Cambridge University Press, Cambridge, UK.
4. Kwon GS, Kim JE, Kim TK, Sohn HY, Koh SC, Shin KS, Kim DG. 2002. Klebsiella pneumoniae KE-1 degrades endosulfan without formation of the toxic metabolite, endosulfan sulfate. FEMS Microbiol Lett 215:255–259. https://doi.org/10.1111/j.1574-6968.2002.tb11399.x.
5. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. 2008. NCBI BLAST: a better Web interface. Nucleic Acids Res 36:W5–W9. https://doi.org/10.1093/nar/gkn201.
6. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
7. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A, Sirotkin Y. 2013. Assembling genomes and mini-metagenomes from highly chimeric reads, p 158–170. *In* Deng M, Jiang R, Sun F, Zhang X (ed), Research in computational molecular biology. RECOMB 2013. Lecture notes in computer science, vol 7821. Springer, Berlin, Germany.
8. Alonge M, Soyk S, Ramakrishnan S, Wang X, Goodwin S, Sedlazeck FJ, Lippman ZB, Schatz MC. 2019. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. Genome Biol 20:224. https://doi.org/10.1186/s13059-019-1829-6.
9. 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.
10. 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.
11. 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.