





# Draft Genome Sequences of Plant-Pathogenic *Klebsiella variicola* Strains Isolated from Plantain in Haiti

James C. Fulton,<sup>a</sup>  Jeannie M. Klein,<sup>a</sup> Sladana Bec,<sup>a</sup> Joubert Fayette,<sup>a,b</sup>  Karen A. Garrett,<sup>a,b</sup> Jeffrey B. Jones,<sup>a</sup>  Sujan Timilsina,<sup>a</sup>  Carrie L. Harmon<sup>a</sup>

<sup>a</sup>Department of Plant Pathology, University of Florida, Gainesville, Florida, USA

<sup>b</sup>Institute for Sustainable Food Systems, University of Florida, Gainesville, Florida, USA

James C. Fulton and Jeannie M. Klein contributed equally to this work. James C. Fulton collected strains and performed initial strain characterization, submitted sequences to NCBI, and led manuscript writing. Jeannie M. Klein extracted DNA, submitted it for sequencing, assembled genomes, submitted genomes for annotation, and wrote portions of the manuscript. Author order was determined based on project seniority.

**ABSTRACT** The genus *Klebsiella* includes pathogenic and nonpathogenic species. We report the 5.57-Mb genome sequences of two *Klebsiella variicola* strains, G18-1365 and G18-1376, isolated from symptomatic plantain plants in Haiti. These strains are genetically closely related (average nucleotide identity [ANI] > 99%) to the previously described type strain of *K. variicola*, DSM 15968.

A member of the *Enterobacteriaceae* family, the Gram-negative genus *Klebsiella* includes species found globally in a wide range of habitats, some of which are important animal and plant pathogens (1–3). Using strains isolated from plant and nosocomial patient samples, *K. variicola* was first classified as a distinct species in 2004 based on DNA-DNA hybridization and multilocus sequence analysis (MLSA) of six housekeeping genes (4). We previously reported isolation of *K. variicola* strains from symptomatic plantain plants exhibiting necrotic soft rot in several Haitian farms (5). The bacteria were preliminarily identified as *K. variicola* using MLSA of the *rRNA*, *rpoB*, *phoE*, and *infB* genes and phenotype microarray profiles (Biolog, Hayward, CA). To confirm that the strains were most closely related to *K. variicola*, a subset of strains was subjected to whole-genome sequencing.

Single colonies previously preserved in 20% glycerol and stored at  $-80^{\circ}\text{C}$  were grown for 24 h at  $28^{\circ}\text{C}$  on nutrient agar (BBL; Becton, Dickinson and Co., Franklin Lakes, NJ), and genomic DNA was extracted using a Wizard genomic DNA purification kit (Promega, Chicago, IL). Library preparation and genome sequencing were performed at the Microbial Genome Sequencing Center (Pittsburgh, PA). Tagmentation, as described by Baym et al. (6), was used to prepare the genomic libraries. A NextSeq 550 system (Illumina, San Diego, CA) was used to produce 151-bp paired-end reads. Trim Galore! v. 0.6.3 with default parameters was used to trim and pair the raw reads, which were then assembled into contigs with SPAdes v. 3.10.1 with k-mers 21, 33, 55, 77, 99, and 127. Contigs smaller than 500 bp and with less than 2.0 k-mer coverage were removed. Bowtie 2 v. 2.3.3 was used with default parameters to align the validated reads against the filtered contigs and output a sequence alignment map (SAM)-formatted alignment. SAMtools was used to convert the SAM files to binary alignment map (BAM) files; then, Pilon v. 1.22 was used with default parameters to polish the draft assemblies and output improved FASTA files.

Assembled sequences were annotated using the Department of Energy Joint Genome Institute's Integrated Microbial Genomes and Microbiomes annotation pipeline v. 5.0.3 (7, 8) and the National Center for Biotechnology Information's internal Prokaryotic Genome Annotation Pipeline (9). The final assembly for G18-1365 was 5,572,087 bp

**Citation** Fulton JC, Klein JM, Bec S, Fayette J, Garrett KA, Jones JB, Timilsina S, Harmon CL. 2020. Draft genome sequences of plant-pathogenic *Klebsiella variicola* strains isolated from plantain in Haiti. *Microbiol Resour Announc* 9:e00336-20. <https://doi.org/10.1128/MRA.00336-20>.

**Editor** David A. Baltrus, University of Arizona

**Copyright** © 2020 Fulton 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 Sujan Timilsina, [sujan.timilsina@ufl.edu](mailto:sujan.timilsina@ufl.edu), or Carrie L. Harmon, [clharmon@ufl.edu](mailto:clharmon@ufl.edu).

**Received** 6 April 2020

**Accepted** 19 June 2020

**Published** 16 July 2020

**TABLE 1** ANI<sup>a</sup> comparisons of sequenced *Klebsiella variicola* genomes with representative publicly available *Klebsiella variicola* assemblies using FastANI calculations

Strain	ANI (%) for strain:		
	At-22	DSM 15968	342
G18-1365	99.07	99.06	98.94
G18-1376	99.07	99.06	98.93

<sup>a</sup> ANI, average nucleotide identity.

long with a GC content of 57.35%, 43 contigs, and 87.03× coverage. The annotation for strain G18-1365 (IMG Genome ID 28356457) predicted 5,235 protein-coding genes and 173 RNA genes. The final assembly for G18-1376 was 5,572,027 bp long with a 57.35% GC content, 41 contigs, and 86.92× coverage. The annotation for strain G18-1376 (IMG Genome ID 2835651175) predicted 5,229 protein-coding genes and 178 RNA genes. Average nucleotide identity comparisons between newly sequenced strains G18-1365 and G18-1376 and *K. variicola* strains At-22 (10), DSM 15968 (11), and 342 (12) were computed using methods described by Jain et al. (13), and the results are shown in Table 1.

**Data availability.** These whole-genome sequences were deposited in GenBank with the raw data accession numbers [PRJNA589638](https://ncbi.nlm.nih.gov/submit/PRJNA589638) and [PRJNA589639](https://ncbi.nlm.nih.gov/submit/PRJNA589639) and assembly accession numbers [GCA\\_009729475.1](https://ncbi.nlm.nih.gov/submit/GCA_009729475.1) and [GCA\\_009729495.1](https://ncbi.nlm.nih.gov/submit/GCA_009729495.1) for G18-1365 and G18-1376, respectively.

## ACKNOWLEDGMENTS

This work was supported by funds from the University of Florida-IFAS Plant Diagnostic Center International Diagnostic Service, the National Science Foundation Graduate Research Fellowship Grant (number DGE-1842473), the University of Florida Graduate Student Fellowship, and the United States Agency for International Development (USAID) under cooperative agreement number AID-OAA-A-15-00039. The contents are the responsibility of the authors and do not necessarily reflect the views of USAID or the U.S. government.

## REFERENCES

- Berry GJ, Loeffelholz MJ, Williams-Bouyer N. 2015. An investigation into laboratory misidentification of a bloodstream *Klebsiella variicola* infection. *J Clin Microbiol* 53:2793–2794. <https://doi.org/10.1128/JCM.00841-15>.
- Seki M, Gotoh K, Nakamura S, Akeda Y, Yoshii T, Miyaguchi S, Inohara H, Horii T, Oishi K, Iida T, Tomono K. 2013. Fatal sepsis caused by an unusual *Klebsiella* species that was misidentified by an automated identification system. *J Med Microbiol* 62:801–803. <https://doi.org/10.1099/jmm.0.051334-0>.
- Fan H, Zeng L, Yang P, Guo Z, Bai T. 2016. First report of banana soft rot caused by *Klebsiella variicola* in China. *Plant Dis* 100:517. <https://doi.org/10.1094/PDIS-05-15-0586-PDN>.
- Rosenblueth M, Martínez L, Silva J, Martínez-Romero E. 2004. *Klebsiella variicola*, a novel species with clinical and plant-associated isolates. *Syst Appl Microbiol* 27:27–35. <https://doi.org/10.1078/0723-2020-00261>.
- Fulton JC, Bec S, Fayette J, Ploetz RC, Garrett KA, Harmon CL. 2020. First report of plantain soft rot caused by *Klebsiella variicola* in Haiti. *Plant Dis* 104:1851. <https://doi.org/10.1094/PDIS-10-19-2105-PDN>.
- Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* 10:e0128036. <https://doi.org/10.1371/journal.pone.0128036>.
- Chen I-MA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese N, White JR, Seshadri R, Smirnova T, Kirton E, Jungbluth SP, Woyke T, Eloe-Fadrosh EA, Ivanova NN, Kyrpides NC. 2019. IMG/M v. 5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* 47:D666–D677. <https://doi.org/10.1093/nar/gky901>.
- Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen I-MA, Pati A, Nielsen T, Markowitz VM, Kyrpides NC. 2015. The standard operating procedure of the DOE-JGI microbial genome annotation pipeline (MGAP v. 4). *Stand Genomic Sci* 10:86. <https://doi.org/10.1186/s40793-015-0077-y>.
- 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>.
- Pinto-Tomás AA, Anderson MA, Suen G, Stevenson DM, Chu FS, Cleland WW, Weimer PJ, Currie CR. 2009. Symbiotic nitrogen fixation in the fungus gardens of leaf-cutter ants. *Science* 326:1120–1123. <https://doi.org/10.1126/science.1173036>.
- Chen M, Li Y, Li S, Tang L, Zheng J, An Q. 2016. Genomic identification of nitrogen-fixing *Klebsiella variicola*, *K. pneumoniae* and *K. quasipneumoniae*. *J Basic Microbiol* 56:78–84. <https://doi.org/10.1002/jobm.201500415>.
- Fouts DE, Tyler HL, DeBoy RT, Daugherty S, Ren Q, Badger JH, Durkin AS, Huot H, Shrivastava S, Kothari S, Dodson RJ, Mohamoud Y, Khouri H, Roesch LFW, Krogfelt KA, Struve C, Triplett EW, Methé BA. 2008. Complete genome sequence of the N<sub>2</sub>-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. *PLoS Genet* 4:e1000141. <https://doi.org/10.1371/journal.pgen.1000141>.
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:1–8. <https://doi.org/10.1038/s41467-018-07641-9>.