

First Complete Genome Sequence of *Escherichia albertii* Strain KF1, a New Potential Human Enteric Pathogen

Krzysztof Fiedoruk,^a Tamara Daniluk,^a Izabela Swiecicka,^b Emilia Murawska,^b Malgorzata Sciepek,^a Katarzyna Leszczynska^a

Department of Microbiology, Medical University of Białystok, Białystok, Poland^a; Department of Microbiology, Institute of Biology, University of Białystok, Białystok, Poland^b

Escherichia albertii has been recently recognized as an emerging human and bird enteric pathogen. Here, we report the first complete chromosome nucleotide sequence of a clinical isolate of *E. albertii* strain KF1, which may provide information about the pathogenic potential of this new species and the mechanisms of evolution of enteropathogenic *Escherichia* spp.

Received 2 January 2014 Accepted 6 January 2014 Published 30 January 2014

Citation Fiedoruk K, Daniluk T, Swiecicka I, Murawska E, Sciepek M, Leszczynska K. 2014. First complete genome sequence of *Escherichia albertii* strain KF1, a new potential human enteric pathogen. *Genome Announc.* 2(1):e00004-14. doi:10.1128/genomeA.00004-14.

Copyright © 2014 Fiedoruk et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Krzysztof Fiedoruk, krzysztof.fiedoruk@umb.edu.pl.

Escherichia albertii is a new emerging enteric pathogen that has been associated with infections in humans and birds (1). This bacterium was first isolated from five Bangladeshi children with diarrhea, but it was initially identified as *Hafnia alvei* (2). However, further analysis has showed that those isolates were more closely related to the genera *Escherichia* and *Shigella*, and subsequently the isolates were reclassified in 2003 as *E. albertii* (3, 4). *E. albertii* is sometimes misidentified because it shares some virulence factors (e.g., intimin) with enteropathogenic *Escherichia coli* (EPEC) and enterohemorrhagic *E. coli* (EHEC) strains (5). Nevertheless, its pathogenic potential is poorly understood, and little is known about its prevalence in food and clinical samples. To date, only three draft genome sequences of *E. albertii* are available in GenBank, including the annotated scaffold representing the chromosome of *E. albertii* TW07627, originally isolated from children in Bangladesh.

Here, we report the complete chromosome sequence of *E. albertii* strain KF1, isolated together with *Campylobacter jejuni* and *C. difficile* from a stool sample from a 1-year-old child with acute diarrhea and identified according to Hyma et al. (6). The stool sample was collected with consent according to Polish law and approved by the Bioethics Commission of the Medical University of Białystok.

The genome of *E. albertii* strain KF1 was sequenced using three platforms, (i) an Ion Torrent 316 yielding 3,472,976 reads with an average length of 183 bp, (ii) an Illumina paired-end technology (2 × 250 bp) generating 756 Mb of data and 709,435 paired reads, and (iii) a Roche GS FLX 8-kb-long paired-end technology generating 103,198 reads with 35,043,425 bp of a raw sequence. *De novo* hybrid assembly was performed using the MaSuRCA genome assembler version 2.1.0 (7), resulting in 5 scaffolds of >1,000 bp. Gaps were initially filled *in silico* using the GapFiller software version 1.3 (8), and the remaining gaps were verified using PCR/sequencing (9).

The *E. albertii* KF1 genome consists of a 4,701,875-bp chromosome with a G+C content of 49.7% and four plasmids (3.8 kb, 5.5 kb, 77 kb, and 85 kb). The annotation was performed using the RAST server (10) and manually inspected for start codons and

potential pseudogenes. The chromosome contains 4,422 open reading frames (ORFs), 85 tRNAs, and 7 rRNA operons. In addition, 121 genes are currently designated pseudogenes. The preliminary analysis showed the presence of genes of the locus of enterocyte effacement (LEE), which contains the intimin gene (*eae*) and encodes other virulence factors, the cytolethal distending toxins CdtA/CdtB and invasion protein IbeA. The detailed report of a full analysis of plasmids will be presented in a future publication.

The work will provide a platform to identify new genes that may contribute to this pathogen's virulence or markers useful in reliable identification of *E. albertii* strains. Moreover, it may help to expand our knowledge about genomic variability and plasticity and the mechanisms of evolution of enteropathogenic *Escherichia* strains.

Nucleotide sequence accession number. The annotated chromosome sequence of *E. albertii* KF1 has been deposited in GenBank under accession number [CP007025](https://ncbi.nlm.nih.gov/nuccore/CP007025).

ACKNOWLEDGMENTS

This work was supported by project number 133-22889L from the Medical University of Białystok. We also thank the Ministry of Science and Higher Education in Poland for funds allocated to "Specific Scientific Equipment-2013."

REFERENCES

1. Nimri LF. 2013. *Escherichia albertii*, a newly emerging enteric pathogen with poorly defined properties. *Diagn. Microbiol. Infect. Dis.* 77:91–95. <http://dx.doi.org/10.1016/j.diagmicrobio.2013.06.028>.
2. Albert MJ, Faruque SM, Ansaruzzaman M, Islam MM, Haider K, Alam K, Kabir I, Robins-Browne R. 1992. Sharing of virulence-associated properties at the phenotypic and genetic levels between enteropathogenic *Escherichia coli* and *Hafnia alvei*. *J. Med. Microbiol.* 37:310–314. <http://dx.doi.org/10.1099/00222615-37-5-310>.
3. Janda JM, Abbott SL, Albert MJ. 1999. Prototypal diarrheagenic strains of *Hafnia alvei* are actually members of the genus *Escherichia*. *J. Clin. Microbiol.* 37:2399–23401.
4. Huys G, Cnockaert M, Janda JM, Swings J. 2003. *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. *Int. J. Syst. Evol. Microbiol.* 53(Pt 3):807–810. <http://dx.doi.org/10.1099/ijs.0.02475-0>.
5. Ooka T, Seto K, Kawano K, Kobayashi H, Etoh Y, Ichihara S, Kaneko

- A, Isobe J, Yamaguchi K, Horikawa K, Gomes TA, Linden A, Bardiau M, Mainil JG, Beutin L, Ogura Y, Hayashi T. 2012. Clinical significance of *Escherichia albertii*. *Emerg. Infect. Dis.* 18:488–492. <http://dx.doi.org/10.3201/eid1803.111401>.
6. Hyma KE, Lacher DW, Nelson AM, Bumbaugh AC, Janda JM, Strockbine NA, Young VB, Whittam TS. 2005. Evolutionary genetics of a new pathogenic *Escherichia* species: *Escherichia albertii* and related *Shigella boydii* strains. *J. Bacteriol.* 187:619–628. <http://dx.doi.org/10.1128/JB.187.2.619-628.2005>.
7. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29:2669–2677. <http://dx.doi.org/10.1093/bioinformatics/btt476>.
8. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol.* 13:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.
9. Murawska E, Fiedoruk K, Bideshi DK, Swiecicka I. 2013. Complete genome sequence of *Bacillus thuringiensis* subsp. *thuringiensis* strain IS5056, an isolate highly toxic to *Trichoplusia ni*. *Genome Announc.* 1(2): e00108-13. <http://dx.doi.org/10.1128/genomeA.00108-13>.
10. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.