AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Draft Genome Sequences of *Escherichia albertii*, *Escherichia fergusonii*, and Strains Belonging to Six Cryptic Lineages of *Escherichia* spp.

Jayanthi Gangiredla,[a] Mark K. Mammel,[a] Tammy J. Barnaba,[a] Carmen Tartera,[a] Solomon T. Gebru,[a] Isha R. Patel,[a] Susan R. Leonard,[a] Michael L. Kotewicz,[a] Keith A. Lampel,[a] Christopher A. Elkins,[a] David W. Lacher[a]

[a]Division of Molecular Biology, Office of Applied Research and Safety Assessment, Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, Laurel, Maryland, USA

ABSTRACT We report here the genome sequences of 55 strains belonging to the genus *Escherichia* from multiple animal and environmental sources. These strains include representatives of *Escherichia albertii*, *Escherichia fergusonii*, and six additional genetically distinct lineages of *Escherichia* spp., one of which is newly discovered and is being reported for the first time here.

Escherichia coli is undoubtedly the most studied bacterial species in microbiology. However, other members of the genus do exist, with the most widely accepted species being *Escherichia albertii* (1) and *Escherichia fergusonii* (2). *Escherichia adecarboxylata*, *E. blattae*, *E. hermannii*, and *E. vulneris* were also once considered to be part of the genus but have since been reclassified (3–6). In 2009, Walk et al. reported the discovery of five genetically distinct "cryptic" lineages of *Escherichia* spp. that were phenotypically indistinguishable from *E. coli* (7, 8). As part of our continuing work to expand the observed genomic diversity of *E. coli*, the genomes of 55 strains belonging to these other lineages and species of *Escherichia* were sequenced. Many of the strains were previously misidentified as *E. coli*, demonstrating the need for further study and characterization of these lineages.

Pure cultures of each strain were grown aerobically overnight in Luria-Bertani broth at 37°C. Total genomic DNA was extracted from 1 ml of overnight culture using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany). DNA extractions were performed with the Qiagen QIAcube instrument using the manufacturer's Gram-negative bacteria protocol. Sequencing libraries were prepared with 1 ng of DNA by using the Nextera XT DNA sample prep kit (Illumina, San Diego, CA, USA) and sequenced on either an Illumina MiSeq or NextSeq platform. The resulting paired-end reads were quality controlled using FastQC ($Q > 30$) and *de novo* assembled using SPAdes 3.8.2 (9) or CLC Genomics Workbench 8.2.1 (CLC bio, Aarhus, Denmark).

Depth of coverage for the draft genomes ranged from 21× to 186×, with the genome sizes ranging from 4,006,242 to 5,544,922 bp. The number of contigs ranged from 29 to 333, while the N_{50} values ranged from 30,834 to 440,408 bp. Nine strains of *E. albertii* and two strains of *E. fergusonii* were sequenced as part of the collection. Preliminary phylogenetic analysis utilizing polymorphisms present within conserved core genes identified one strain, EC7003, as belonging to a new cryptic lineage, which we have designated lineage 6. The remaining 43 strains belong to cryptic lineages 1 through 5, with 10, 1, 3, 4, and 25 strains identified within each lineage, respectively. The strains were screened for the presence of the known or putative virulence factors *aggR*, *eae*, *ipaH*, LT, ST, stx_1, and stx_2. All 55 strains were negative for *aggR*, *ipaH*, and stx_1. Both *E. fergusonii* strains and all cryptic lineage 2, 3, 4, and 5 strains were negative for all seven factors investigated. All nine *E. albertii* strains and the cryptic lineage 6

Received 1 March 2018 **Accepted** 29 March 2018 **Published** 3 May 2018

Citation Gangiredla J, Mammel MK, Barnaba TJ, Tartera C, Gebru ST, Patel IR, Leonard SR, Kotewicz ML, Lampel KA, Elkins CA, Lacher DW. 2018. Draft genome sequences of *Escherichia albertii*, *Escherichia fergusonii*, and strains belonging to six cryptic lineages of *Escherichia* spp. Genome Announc 6:e00271-18. https://doi.org/10.1128/genomeA.00271-18.

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to David W. Lacher, david.lacher@fda.hhs.gov.

strain EC7003 possess the *eae* adhesin. The cryptic lineage 1 strains are more varied in their profiles, with five strains negative for all seven factors, three strains positive for LT and ST, one strain positive for LT only, and one strain positive for ST and stx_2.

Accession number(s). The draft genome assemblies were deposited in DDBJ/ENA/GenBank through the FDA's GenomeTrakr pipeline under BioProject number PRJNA230969 with accession numbers PTRD00000000 to PTTF00000000. The versions described in this announcement are the first versions, with the exception of PTSX00000000, which is the second version.

ACKNOWLEDGMENTS

The views expressed in this article are those of the authors and do not necessarily reflect the official policy of the Department of Health and Human Services, the U.S. Food and Drug Administration (FDA), or the U.S. Government. Reference to any commercial materials, equipment, or process does not in any way constitute approval, endorsement, or recommendation by the FDA.

Part of this study was supported by the ORISE Fellowship Program.

REFERENCES

1. 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 53:807–810. https://doi.org/10.1099/ijs.0.02475-0.
2. Farmer JJ, 3rd, Fanning GR, Davis BR, O'Hara CM, Riddle C, Hickman-Brenner FW, Asbury MA, Lowery VA, 3rd, Brenner DJ. 1985. *Escherichia fergusonii* and *Enterobacter taylorae*, two new species of *Enterobacteriaceae* isolated from clinical specimens. J Clin Microbiol 21:77–81.
3. Tamura K, Sakazaki R, Kosako Y, Yoshizaki E. 1986. *Leclercia adecarboxylata* gen. nov., comb. nov., formerly known as *Escherichia adecarboxylata*. Curr Microbiol 13:179–184. https://doi.org/10.1007/BF01568943.
4. Priest FG, Barker M. 2010. Gram-negative bacteria associated with brewery yeasts: reclassification of *Obesumbacterium proteus* biogroup 2 as *Shimwellia pseudoproteus* gen. nov., sp. nov., and transfer of *Escherichia blattae* to *Shimwellia blattae* comb. nov. Int J Syst Evol Microbiol 60:828–833. https://doi.org/10.1099/ijs.0.013458-0.
5. Hata H, Natori T, Mizuno T, Kanazawa I, Eldesouky I, Hayashi M, Miyata M, Fukunaga H, Ohji S, Hosoyama A, Aono E, Yamazoe A, Tsuchikane K, Fujita N, Ezaki T. 2016. Phylogenetics of family *Enterobacteriaceae* and proposal to reclassify *Escherichia hermannii* and *Salmonella subterranea* as *Atlantibacter hermannii* and *Atlantibacter subterranea* gen. nov., comb. nov. Microbiol Immunol 60:303–311. https://doi.org/10.1111/1348-0421.12374.
6. Alnajar S, Gupta RS. 2017. Phylogenomics and comparative genomic studies delineate six main clades within the family *Enterobacteriaceae* and support the reclassification of several polyphyletic members of the family. Infect Genet Evol 54:108–127. https://doi.org/10.1016/j.meegid.2017.06.024.
7. Walk ST, Alm EW, Gordon DM, Ram JL, Toranzos GA, Tiedje JM, Whittam TS. 2009. Cryptic lineages of the genus *Escherichia*. Appl Environ Microbiol 75:6534–6544. https://doi.org/10.1128/AEM.01262-09.
8. Walk ST. 2015. The "cryptic" *Escherichia*. EcoSal Plus 6. https://doi.org/10.1128/ecosalplus.ESP-0002-2015.
9. 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.