Draft Genomic Sequences of Three *Escherichia coli* Sequence Type 131 Isolates (H45, H43ii, and H43iii) from Patients in Lagos, Nigeria

Yishan Yang,a Christopher H. Sommers,a Eyitayo O. Adenipekun,b,c Marina Ceruso,d Charlene R. Jackson,e Tiffanie A. Woodley,a John B. Barrett,a Lari M. Hiott,a Jonathan G. Frye,a Yanhong Liu*a

a U.S. Department of Agriculture, Agricultural Research Service, Eastern Regional Research Center, Wyndmoor, Pennsylvania, USA
b Department of Medical Microbiology and Parasitology, Olabisi Onabanjo University, College of Health Sciences, Sagamu, Ogun State, Nigeria
c Department of Medical Microbiology and Parasitology (Medical Laboratory Science), College of Medicine, University of Lagos, Lagos, Nigeria
d Department of Veterinary Medicine and Animal Production, University of Naples Federico II, Naples, Italy
e Bacterial Epidemiology and Antimicrobial Resistance Research Unit, USDA-ARS, Russell Research Center, Athens, Georgia, USA

**ABSTRACT** *Escherichia coli* sequence type 131 (ST131) has recently emerged as a leading multidrug-resistant pathogen that causes urinary tract and bloodstream infections in humans. Here, we report the draft genomic sequences of three *E. coli* ST131 isolates, H45, H43ii, and H43iii, from urine samples of patients in Lagos, Nigeria.

Extraintestinal pathogenic *Escherichia coli* (ExPEC) is a common cause of urinary tract infections (UTIs), bacteremia, and neonatal meningitis in humans (1). The widespread use of antimicrobials to treat human and animal infections and to enhance livestock growth results in dissemination of multidrug-resistant ExPEC strains, among which sequence type 131 (ST131) is the most frequent isolate (2, 3). The prevalence of *E. coli* ST131 is possibly attributable to its increased antimicrobial resistance, enhanced virulence, and greater propensity to transfer genetic materials compared to non-ST131 *E. coli* (4–6).

Three *E. coli* ST131 strains, H45, H43ii, and H43iii, were isolated from urine samples of patients in Lagos, Nigeria (5). Prior to whole-genome sequencing, the phylogenetic group and virulence factors of the three *E. coli* strains were determined in our laboratory using PCR methods (7, 8), which confirmed that all these strains belonged to phylogenetic group B2 and that they were ExPEC strains.

For whole-genome sequencing, genomic DNA was extracted using the DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA) from overnight cultures grown on Trypticase soy agar (TSA; Becton, Dickinson, and Company, Sparks, MD, USA) plates. The concentration of genomic DNA was determined using a Qubit 3.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) with Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kits (Thermo Fisher Scientific). Sequencing libraries were prepared using the Nextera DNA flex library prep kit (Illumina, San Diego, CA, USA) according to the manufacturer’s instructions. Prepared libraries were quantified, pooled, and denatured before paired-end sequencing (151 cycles with 150-bp read length) using the Illumina MiniSeq instrument. The quality of the sequence reads was assessed with FastQC version 1.0.0 (BaseSpace Labs, Illumina) (9), and the genome was *de novo* assembled using the SPAdes genome assembler version 3.9.0 (BaseSpace Labs) (10). Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 4.10 (11). Default parameters were used for all software unless otherwise specified. Genome cover-
age, genome size, number of paired-end reads, GC content, and other characteristics are shown in Table 1.

Serotype, multilocus sequence types (MLST), virulence genes, and antimicrobial resistance were determined using the E. coli Serotyping Pipeline version 1.0.2 (BaseSpace Labs) and the Bacterial Analysis Pipeline version 1.0.4 (BaseSpace Labs). Based on the sequencing data, all three isolates were serotype O25:H4 and belonged to E. coli ST131. The virulence factors of the three isolates included genes encoding serum resistance (iss), glutamate decarboxylase (gad), secreted autotransporter toxin (sat), IgA homologue adhesin (iha), and diffuse adherence fibrillar adhesion (fmaE). Antimicrobial resistance genes identified in all three isolates were a fluoroquinolone and aminoglycoside resistance gene [aac(6’)-Ib-cr], aminoglycoside resistance genes [aac(3)-IIa and adaA5], a sulfonamide resistance gene (sul1), a trimethoprim resistance gene (dfrA17), a tetracycline resistance gene (tet(A)), beta-lactam resistance genes (blaCTX-M-15, blaTEM-1B, and blaOXA-1), and a phenicol resistance gene (catB4). Additionally, three more antimicrobial resistance genes were found in isolates H45 and H43iii, which were aminoglycoside resistance genes (strA and strB) and a sulfonamide resistance gene (sul2).

Whole-genome sequencing is an effective tool for the identification and characterization of bacterial pathogens. The genomic data will be useful for understanding the dissemination and pathogenicity of E. coli ST131, as well as for facilitating the development of novel antimicrobial therapies.

Data availability. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession and BioProject numbers listed in Table 1. The versions described here are the first versions.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Agriculture (USDA) Agricultural Research Service National Program project 108 and Food Safety project 8072-42000-078-00D.

Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

REFERENCES

1. Poolman JT, Wacker M. 2016. Extraintestinal pathogenic Escherichia coli, a common human pathogen: challenges for vaccine development and progress in the field. J Infect Dis 213:6–13. https://doi.org/10.1093/infdis/jiv429.
2. Nicolas-Chanoine MH, Bertrand X, Madec JY. 2014. Escherichia coli ST131, an intriguing clonal group. Clin Microbiol Rev 27:543–574. https://doi.org/10.1128/CMR.00125-13.
3. Sarowska J, Futoma-Kloch B, Jama-Kmiecik A, Frej-Madrzak M, Ksiazczyk M, Bugla-Ploskonska G, Choroszy-Krol I. 2019. Virulence factors, prevalence and potential transmission of extraintestinal pathogenic Escherichia coli isolated from different sources: recent reports. Gut Pathog 11:10. https://doi.org/10.1186/s13099-019-0290-0.
4. Johnson JR, Johnston B, Ciabots C, Kuskowski MA, Castanheira M. 2010. Escherichia coli sequence type ST131 as the major cause of serious multidrug-resistant E. coli infections in the United States. Clin Infect Dis 51:286–294. https://doi.org/10.1086/653932.
5. Adenipekun EO, Jackson CR, Ramadan H, Iwalokun BA, Oyedeji KS, Frye JG, Barrett JB, Hiott LM, Woodley TA, Oluwadun A. 2016. Prevalence and multidrug resistance of Escherichia coli from community-acquired infections in Lagos, Nigeria. J Infect Dev Ctries 10:920–931. https://doi.org/10.3855/jidc.7997.
6. Chen SL, Ding Y, Apisarnthanarak S, Archuleta S, Omar SFS, De Pratim P, Koh TH, Chew KL, Aitai N, Suwantarat N, Velayuthan RD, Wong JGX, Lye DC. 2019. The higher prevalence of extended spectrum beta-lactamases among Escherichia coli ST131 in Southeast Asia is driven by expansion of a single, locally prevalent subclone. Sci Rep 9:13245. https://doi.org/10.1038/s41598-019-49467-5.
7. Johnson JR, Kuskowski MA, Gajewski A, Sahm DF, Karlowyka JA. 2004. Virulence characteristics and phylogenetic background of multidrug-resistant and antimicrobial-susceptible clinical isolates of Escherichia coli from across the United States, 2000–2001. J Infect Dis 190:1739–1744. https://doi.org/10.1086/425018.
8. Clermont O, Bonacorsi S, Bingen E. 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* 66:4555–4558. [https://doi.org/10.1128/aem.66.10.4555-4558.2000](https://doi.org/10.1128/aem.66.10.4555-4558.2000).

9. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. [http://www.bioinformatics.babraham.ac.uk/projects/fastqc.](http://www.bioinformatics.babraham.ac.uk/projects/fastqc.)

10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, 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](https://doi.org/10.1089/cmb.2012.0021).

11. Tatusova T, DiCuccio M, Badreddin 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](https://doi.org/10.1093/nar/gkw569).