

Draft Genome Sequence of Enterotoxigenic *Escherichia coli* Strain E24377A, Obtained from a Tribal Drinking Water Source in India

Ashok J. Tamhankar,^{a,b} Sandeep S. Nerkar,^{a,b} Prashant P. Khadake,^c Dadasaheb B. Akolkar,^d Sachin R. Apurwa,^d Uday Deshpande,^e Smita U. Khedkar,^f Cecilia Stålsby-Lundborg^b

Indian Initiative for Management of Antibiotic Resistance (IIMAR), Department of Environmental Medicine, R.D. Gardi Medical College, Ujjain, India^a; Department of Public Health Sciences (Global Health/IHCAR), Karolinska Institutet, Stockholm, Sweden^b; Life Technologies India Ltd., Thane, India^c; Datar Genetics Ltd., Nashik, India^d; Labindia (GOPD), Thane, India^e; Bac-Test Laboratory, Nashik, India^f

A.J.T. and S.S.N. contributed equally to this work.

Enterotoxigenic *Escherichia coli* (ETEC) is a major cause of diarrheal disease in humans and animals. Its dissemination can occur through water sources contaminated by it. Here, we report for the first time the draft genome sequence of ETEC strain E24377A, obtained from a tribal drinking water source in India.

Received 16 February 2015 Accepted 23 February 2015 Published 2 April 2015

Citation Tamhankar AJ, Nerkar SS, Khadake PP, Akolkar DB, Apurwa SR, Deshpande U, Khedkar SU, Stålsby-Lundborg C. 2015. Draft genome sequence of enterotoxigenic *Escherichia coli* strain E24377A, obtained from a tribal drinking water source in India. *Genome Announc* 3(2):e00225-15. doi:10.1128/genomeA.00225-15.

Copyright © 2015 Tamhankar 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 Ashok J. Tamhankar, ejetee@gmail.com.

Infection by enterotoxigenic *Escherichia coli* (ETEC) results in a large volume of watery diarrhea, generally known as traveler's diarrhea (1). Besides humans, the infection also causes diarrhea in animals (2). ETEC primarily colonizes the small intestine by way of the colonization factor antigen (CFA) pili; *E. coli* strain E24377A has been shown to contain CFA pili types CS1 and CS3. Two additional toxins are thought to be responsible for its virulence, a heat-stable enterotoxin and a heat-labile enterotoxin. Other virulence factors include those belonging to serotype O139:H28 and potential factors encoded on 6 plasmids. ETEC colonizes the lower gut of animals and survives when released into the natural environment, allowing widespread dissemination to new hosts. Since the spread can easily occur through drinking water, its presence in drinking water sources has great epidemiological and public health consequences, particularly in the case of tribal areas, where diarrheal disease management measures are largely inadequate. However, the whole-genome sequence of ETEC E24377A from drinking water sources from tribal areas has not been reported.

The complete genome sequences of three ETEC strains infecting humans, E24377A (3), H10407 (4), and B2C (5), as well as one ETEC strain, W25K, infecting piglets (6), have been published. Recently, draft whole-genome sequences of 10 serogroup O6 ETEC strains from historical and cruise ship outbreaks (7) were announced. Here, we report the draft genome sequence of strain E24377A, which is an O139:H28 serotype strain of ETEC that was obtained from an environmental sample intercepted from a drinking water well in a hilly tribal village in India. In this village, we reported an earlier occurrence of diarrheal cases and the contamination of tribal drinking water wells by fecal coliform and *E. coli* (8).

Genomic DNA was extracted from the *E. coli* O139:H28 strain E24377A using the GenElute bacterial genomic DNA kit (Sigma Aldrich), according to the manufacturer's protocol. The genomic

DNA was used to generate a library using the Ion Xpress fragment library kit and sequenced using a semiconductor-based sequencing platform (Ion Proton; Life Technologies). The sequencing generated approximately 3 million singleton reads, constituting 101-fold coverage of the genome. The reads were overlapped where possible and trimmed for quality using the PRINSEQ version 0.20.4 software. A *de novo* assembly of the overlapped and quality-trimmed reads was generated using the CLC Workbench. The final assembly consists of 453 contigs. The genome size of ETEC O139:H28 strain E24377A, intercepted from an aquatic environment, is estimated to be 4.9 Mb. The genome sequence was annotated using the RAST genome annotation server (9). An analysis of the genome showed that the environmental ETEC O139:H28 strain E24377A has 597 subsystems, 5,305 coding sequences, and 67 RNAs. Type II and type VI protein secretion systems were identified. Classical ETEC virulence genes, *eltA* and *eltB*, the genes encoding the two subunits of heat-labile enterotoxin, were also found.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession no. [JXRF00000000](https://www.ncbi.nlm.nih.gov/nuccore/JXRF00000000). The version described in this paper is the first version, [JXRF01000000](https://www.ncbi.nlm.nih.gov/nuccore/JXRF01000000).

ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council.

S.S.N. received a Ph.D. scholarship from the European Commission (Erasmus Mundus External Cooperation Window lot 13).

REFERENCES

1. Fleckenstein JM, Hardwidge PR, Munson GP, Rasko DA, Sommerfelt H, Steinsland H. 2010. Molecular mechanisms of enterotoxigenic *Escherichia coli* infection. *Microbes Infect* 12:89–98. <http://dx.doi.org/10.1016/j.micinf.2009.10.002>.
2. Frydendahl K. 2002. Prevalence of serogroups and virulence genes in *Escherichia coli* associated with postweaning diarrhoea and edema disease in

- pigs and a comparison of diagnostic approaches. *Vet Microbiol* 85: 169–182. [http://dx.doi.org/10.1016/S0378-1135\(01\)00504-1](http://dx.doi.org/10.1016/S0378-1135(01)00504-1).
3. Sahl JW, Rasko DA. 2012. Analysis of global transcriptional profiles of enterotoxigenic *Escherichia coli* isolate E24377A. *Infect Immun* 80: 1232–1242. <http://dx.doi.org/10.1128/IAI.06138-11>.
 4. Crossman LC, Chaudhuri RR, Beatson SA, Wells TJ, Desvaux M, Cunningham AF, Petty NK, Mahon V, Brinkley C, Hobman JL, Savarino SJ, Turner SM, Pallen MJ, Penn CW, Parkhill J, Turner AK, Johnson TJ, Thomson NR, Smith SG, Henderson IR. 2010. A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. *J Bacteriol* 192:5822–5831. <http://dx.doi.org/10.1128/JB.00710-10>.
 5. Madhavan TP, Steen JA, Hugenholtz P, Sakellaris H. 2014. Genome sequence of enterotoxigenic *Escherichia coli* strain B2C. *Genome Announc* 2(2):e00247-14. <http://dx.doi.org/10.1128/genomeA.00247-14>.
 6. Ren W, Liu G, Yin J, Chen S, Li T, Kong X, Peng Y, Yin Y, Hardwidge PR. 2014. Draft genome sequence of enterotoxigenic *Escherichia coli* strain W25K. *Genome Announc* 2(3):e00593-14. <http://dx.doi.org/10.1128/genomeA.00593-14>.
 7. Pattabiraman V, Bopp CA. 2014. Draft whole-genome sequences of 10 serogroup O6 enterotoxigenic *Escherichia coli* strains. *Genome Announc* 2(6):e01274-14. <http://dx.doi.org/10.1128/genomeA.01274-14>.
 8. Nerkar SS, Tamhankar AJ, Khedkar SU, Stålsby Lundborg CS. 2014. Quality of water and antibiotic resistance of *Escherichia coli* from water sources of hilly tribal villages with and without integrated watershed management—a one year prospective study. *Int J Environ Res Public Health* 11:6156–6170. <http://dx.doi.org/10.3390/ijerph110606156>.
 9. 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>.