



# Genome Sequences of Enteroinvasive *Escherichia coli* Sequence Type 6, 99, and 311 Strains Acquired in Asia Pacific

 Rajat Dhakal,<sup>a</sup> Qinning Wang,<sup>a</sup> Peter Howard,<sup>a</sup> Vitali Sintchenko<sup>a,b</sup>

<sup>a</sup>Centre for Infectious Diseases and Microbiology-Public Health, Institute of Clinical Pathology and Medical Research-New South Wales Health Pathology, Westmead, New South Wales, Australia

<sup>b</sup>Marie Bashir Institute for Infectious Diseases and Biosecurity, Sydney Medical School, The University of Sydney, Sydney, New South Wales, Australia

**ABSTRACT** Laboratory diagnosis of enteroinvasive *Escherichia coli* (EIEC) remains difficult and limits the availability of EIEC genomes to the research community. We report the draft genome sequences of three EIEC strains which represent three distinct sequence types and serotypes circulating in Asia Pacific and causing enterocolitis in humans.

Enteroinvasive *Escherichia coli* (EIEC) is closely related to *Shigella* species and causes enterocolitis syndrome, similar to shigellosis (1). However, the diversity and genomics of EIEC strains remain poorly understood due to the small number of available isolates and genomes and the difficulty of laboratory differentiation of EIEC from *Shigella* spp. Culture-independent direct testing (CIDT) on stool from patients with diarrhea has improved the recognition of EIEC-associated disease. The ironic reason for this has been the selection of the *ipaH* gene, which is present in *Shigella* species and EIEC, as the target for CIDT. This report contributes three new EIEC genomes representative of distinct subtypes and serotypes of EIEC causing gastrointestinal infections in humans in Asia Pacific.

Fecal isolates ( $n = 3$ ) from New South Wales (NSW) (including one isolate from a patient with Pacific Islands travel history) were isolated in diagnostic laboratories throughout NSW and sent to the Centre for Infectious Diseases and Microbiology-Public Health (CIDM-PH), a public health microbiology reference laboratory in Sydney, Australia, for further characterization. Whenever required, isolates were grown in blood agar for 24 h at 37°C. We used our algorithm, which includes confirmation of the *ipaH* gene in isolates, identification of *E. coli* using the BD Phoenix automated identification and susceptibility testing system (Becton, Dickinson), and use of our in-house multiplex PCR assay (2), to identify and subtype EIEC. We then performed whole-genome sequencing (WGS) for in-depth characterization of the strains. Genomic DNA (gDNA) was extracted using the Geneaid Presto Mini gDNA bacteria kit. The extracted DNA was fragmented and tagged for multiplexing with Nextera XT DNA sample preparation kits (Illumina) and sequenced in house using the Illumina NextSeq 500 platform with  $2 \times 150$ -bp reads. Default parameters were used for the software used for analysis except where otherwise stated. Trimmomatic (v.0.36) (3) was used for cleaning fastq reads (LEADING: 3, TRAILING: 3, SLIDINGWINDOW: 4:20, MINLEN: 36). The Achtman 7-gene multilocus sequence typing (MLST) scheme in Enterbase and Nullarbor pipeline v.2.0.20181010 (4) containing ResFinder (v.3.1) and ABRicate (v.0.8) were used for subtyping and antibiotic resistome analysis. The Virulence Factors Database (VFDB), Snippy (v.4.3.5), and SKESA (v.2.3.0) were employed for virulome analysis, core genome analysis, and genome assembly, respectively.

The serotypes of these strains were identified as O121, O28, and O96, respectively. The genome sequencing statistics are presented in Table 1. The Achtman 7-gene MLST

**Citation** Dhakal R, Wang Q, Howard P, Sintchenko V. 2019. Genome sequences of enteroinvasive *Escherichia coli* sequence type 6, 99, and 311 strains acquired in Asia Pacific. *Microbiol Resour Announc* 8:e00944-19. <https://doi.org/10.1128/MRA.00944-19>.

**Editor** Julie C. Dunning Hotopp, University of Maryland School of Medicine

**Copyright** © 2019 Dhakal 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 Rajat Dhakal, [rajat.dhakal@health.nsw.gov.au](mailto:rajat.dhakal@health.nsw.gov.au).

**Received** 8 August 2019

**Accepted** 16 August 2019

**Published** 5 September 2019

**TABLE 1** Genome sequencing results

| Sample            | Coverage depth (fold) | Yr of isolation | Size (bp) | GC content (%) | No. of contigs | WGS-inferred serotype | Achtman 7-gene MLST | $N_{50}$ (bp) | No. of reads |
|-------------------|-----------------------|-----------------|-----------|----------------|----------------|-----------------------|---------------------|---------------|--------------|
| 19-0562-EIEC-0001 | 122                   | 2017            | 4,864,058 | 50.6           | 349            | O121                  | 6                   | 28,116        | 3,645,936    |
| 18-0562-010       | 102                   | 2018            | 4,797,511 | 50.5           | 385            | O28                   | 311                 | 39,521        | 2,980,100    |
| 19-0562-EIEC-0006 | 148                   | 2019            | 5,083,019 | 50.3           | 246            | O96                   | 99                  | 100,937       | 4,284,180    |

results showed that they belonged to sequence types 6, 311, and 99, respectively. Their subtypes from our multiplex PCR (2) were 3, 1, and 4 (common subtype). All strains contained the *ipgD* and *mxiA* genes associated with virulence. These two genes have been used to study various forms of pINV plasmids in EIEC strains (1). In addition, 19-0562-EIEC-0001 included a *gyrA* (D87N) mutation conferring *in vitro* resistance to fluoroquinolone antibiotics.

**Data availability.** The genome sequences have been deposited in GenBank under the accession no. [VMTQ000000000](https://ncbi.nlm.nih.gov/nuccore/VMTQ000000000), [VNKK000000000](https://ncbi.nlm.nih.gov/nuccore/VNKK000000000), and [VNKL000000000](https://ncbi.nlm.nih.gov/nuccore/VNKL000000000). The versions described in this paper are the first versions, VMTQ010000000, VNKK010000000, and VNKL010000000, respectively. The SRA accession numbers for these samples are [SRR9903352](https://www.ncbi.nlm.nih.gov/sra/SRR9903352), [SRR9903508](https://www.ncbi.nlm.nih.gov/sra/SRR9903508), and [SRR9903561](https://www.ncbi.nlm.nih.gov/sra/SRR9903561), respectively.

### ACKNOWLEDGMENTS

This research was funded by CIDM-PH.

We thank the scientists at the NSW Public Health Pathogen Genomics Unit at CIDM-PH for the sequencing of clinical strains. We are also grateful to the staff at the NSW Enteric Reference Laboratory for providing the isolates for analysis.

### REFERENCES

- Lan R, Alles MC, Donohoe K, Martinez MB, Reeves PR. 2004. Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infect Immun* 72:5080–5088. <https://doi.org/10.1128/IAI.72.9.5080-5088.2004>.
- Dhakal R, Wang Q, Lan R, Howard P, Sintchenko V. 2018. Novel multiplex PCR assay for identification and subtyping of enteroinvasive *Escherichia coli* and differentiation from *Shigella* based on target genes selected by comparative genomics. *J Med Microbiol* 67:1257–1264. <https://doi.org/10.1099/jmm.0.000784>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Seemann T, Goncalves da Silva A, Bulach DM, Schultz MB, Kwong JC, Howden BP. 2019. Nullarbor. <https://github.com/tseemann/nullarbor>.