Draft genome sequence of a Uropathogenic Escherichia coli (UPEC) isolate (ST38 O1:H15) from Pakistan, an emerging MDR sequence type with a high virulence profile

Draft genome sequence of a uropathogenic *Escherichia coli* isolate (ST38 O1:H15) from Pakistan, an emerging multidrug-resistant sequence type with a high virulence profile

Z. Rafaque[1,2], J. I. Dasti[1] and S. C. Andrews[2]

1) *Department of Microbiology, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan and* 2) *School of Biological Sciences, University of Reading, Reading, United Kingdom*

Abstract

Sequence type 38 is considered a uropathogenic *Escherichia coli*/enteroaggregative *E. coli* hybrid associated with multidrug resistance and urinary tract infections. The draft genome sequence of UEC59 from a woman in Pakistan revealed a 5 324 938 bp genome with 5386 coding sequences (CDS), 86 transfer RNA genes and multiple antibiotic resistance genes (bla_{TEM-1}, *CMY-2, sul1, sul2, dfrA17, tetA, mphA*) and mobile elements (*int1,* two transposons, 30 insertion sequence elements, one integrative conjugative element, four plasmids, five prophages), along with many virulence genes.

Keywords: Draft genome, MDR, Pakistan, ST38, UPEC, virulence
Original Submission: 21 August 2018; **Revised Submission:** 24 September 2018; **Accepted:** 12 October 2018
Article published online: 19 October 2018

Corresponding author: S. C. Andrews, School of Biological Sciences, University of Reading, Reading RG6 6AH, United Kingdom.
E-mail: s.c.andrews@reading.ac.uk

Urinary tract infections (UTI) are the second most common type of clinical infection globally. Uropathogenic *Escherichia coli* (UPEC) is the leading cause of UTI. It exhibits multidrug resistance (MDR) and harbours virulence factors enabling colonization of the urinary tract [1]. In developing nations, horizontal gene transfer has enabled the spread of MDR and acquisition of novel virulence-factor portfolios, exacerbating the problems caused by UPEC. Sequence type (ST) 38 is considered a hybrid of enteroaggregative *E. coli* (EAEC) and UPEC, and is frequently associated with UTIs [2]. Genome sequencing of ST38 isolates from developing nations offers the opportunity to identify emerging UPEC strains that are a potential global threat to human health.

UEC59 is a UPEC from a 73-year-old woman in Pakistan. Antimicrobial susceptibility profiling performed according to Clinical and Laboratory Standards Institute guidelines demonstrated resistance to cephalosporins, fluoroquinolones, trimethoprim and ampicillin, characterizing it as an extended-spectrum β-lactamase–producing MDR strain [3]. Phenotypic profiling of virulence characteristics identified it as a haemolytic, adherent, biofilm-producing and serum-resistant strain. Invasion assays in the uroepithelial cell line ATCC HTB5637 showed that UEC59 is capable of invasion.

Genome sequencing of UEC59 was performed by MicrobesNG, Birmingham, on MiSeq and HiSeq 2500 platforms with sequence coverage of 30×. *De novo* assembly of the reads was performed by SPAdes 3.11 [4]. The quality of genome assemblies was assessed by QUAST [5], genomic features were determined with PATRIC [6] and insertion sequence (IS) and integrative conjugative elements (ICE) were sought with IS finder and ICEberg [7,8]. A total of 146 contigs were obtained with a total genome size of 5 324 938 bp, a GC content of 50.55%, 5386 CDS and 86 transfer RNA genes. A yersiniabactin synthesis–associated ICE (similar to ICEEcoUMN026-1) of 65 732 bp was identified, as were a total of 30 IS elements (11 types).

UEC59 was identified as clonal type ST38 [9] and serotype O1:H15 [10]. Four plasmids were found [11]—*IncFII, IncFIA, IncFIB* and *Col156*—as were genes conferring resistance to extended-spectrum β-lactam drugs (bla_{TEM-1}, CMY-2), sulfonamides (*sul-1* and *sul-2*), trimethoprim (*dfrA17*), aminoglycosides

https://doi.org/10.1016/j.nmni.2018.10.004

(AadA5, APH(6)-Id, AAC(3)-IId), fluoroquinolone (mutation in *parC,* S80I), tetracycline (*tet*(*A*)), quaternary ammonium compounds (*qacEΔ1*) and macrolides (*mph*(*A*)). Moreover, an integron (*int1*) associated with *dfrA17* and transposons (Tn*2,* Tn*3*) associated with *tetA* were identified. Virulence genes present include those for adhesins and invasins (*fimH, papI, papGII, safC, kpsM*), toxins (*espC, ghoT*) and a hypothetical enterotoxin (*senB*). The *safC* gene found in UEC59 is homologous to the aggregative adherence fimbria II usher protein AafC of EAEC, other than that no EAEC-specific gene such as *aggR* is encoded by UEC59. However, studies have reported the presence of EAEC transport regulator gene *aggR* as a key feature in evolution of ST38 UPEC strains (from the United Kingdom, Netherland and Germany) [12]. However, ST38 UPEC from Saudi Arabia and Brazil have shown only 27% (4/15) and 16 UPEC strains out of 225 carrying EAEC-specific genes such as *aggR* and *aatA* respectively [2,13].

UEC59 also encodes iron-acquisition systems involved in the manufacture and utilization of the siderophores enterobactin, yersiniabactin and aerobactin; haeme (*hemSRT-chuY-hmuUV*) uptake; and ferrous iron (*feoABC* and *efeUOB*) utilization. There was a strong correlation between the various virulence factors identified phenotypically and those revealed from the genome sequence. The PHASTER tool [14] identified five intact (PHAGE_Entero_fiAA9, 36 kb; PHAGE_Entero_lambda, 48kb; PHAGE_Entero_lambda 28 kb; PHAGE_Vibrio_X29. 22 kb; PHAGE_Entero_BP, 17 kb) and three incomplete (PHAGE_Entero_N15, 8.5 kb; PHAGE_Salmon_118970, 213 kb; PHAGE_Salmon_118970, 27 kb) prophage regions. CRISPRCasfinder [15] confirmed two CRISPRs of more than ten direct repeats, but only one was associated with *cas* genes (CAS-TypeIE).

Nucleotide sequence accession numbers

This whole-genome draft shotgun sequence was deposited at GenBank under accession number QKMW00000000. The version described in this paper is version QKMW01000000.

Acknowledgements

We are thankful to Higher Education Commission, Pakistan, for supporting doctoral work of Z. Rafaque at Quaid-i-Azam University, Islamabad (213-53961-2BM2-093) and the Commonwealth Scholarship Commission for supporting her research at University of Reading (PKCN-2017-215).

Conflict of interest

None declared.

References

[1] Subashchandrabose S, Mobley HLT. Virulence and fitness determinants of uropathogenic *Escherichia coli*. Microbiol Spectrum 2015;3(4).
[2] Chattaway MA, Jenkins C, Ciesielczuk H, Day M, DoNascimento V, Day M, et al. Evidence of evolving extraintestinal enteroaggregative *Escherichia coli* ST38 clone. Emerg Infect Dis 2014;20:1935–7.
[3] Ali I, Rafaque Z, Ahmed S, Malik S, Dasti JI. Prevalence of multi-drug resistant uropathogenic *Escherichia coli* in Potohar region of Pakistan. Asian Pac J Trop Biomed 2016;6:60–6.
[4] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012;19:455–77.
[5] Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics 2013;29:1072–5.
[6] Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, et al. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res 2014;42(D1):D581–91.
[7] Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. Genome Biol 2011;12:R30.
[8] Bi D, Xu Z, Harrison EM, Tai C, Wei Y, He X, et al. ICEberg: a web-based resource for integrative and conjugative elements found in Bacteria. Nucleic Acids Res 2012;40(Database issue):D621–6.
[9] Jolley KA, Maiden MCJ. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinform 2010;11:595.
[10] Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and easy *in silico* serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. J Clin Microbiol 2015;53:2410–26.
[11] Carattoli A, Zankari E, Garcia-Fernandez A, Voldby Larsen M, Lund O, Villa L, et al. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 2014;58:3895–903.
[12] Alghoribi MF, Gibreel TM, Farnham G. Al Johani SM, Balkhy HH, Upton M. Antibiotic-resistant ST38, ST131 and ST405 strains are the leading uropathogenic *Escherichia coli* clones in Riyadh, Saudi Arabia. J Antimicrob Chemother 2015;70:2757–62.
[13] Abe CM, Salvador FA, Falsetti IN, Vieira MA, Blanco J, Blanco JE, et al. Uropathogenic *Escherichia coli* (UPEC) strains may carry virulence properties of diarrhoeagenic *E. coli*. FEMS Immunol Med Microbiol 2008;52:397–406.
[14] Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 2016;44(W1):W16–21.
[15] Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 2007;35(Suppl. 2):W52–7.