



Complete Genome Sequence of *Escherichia coli* 81009, a Representative of the Sequence Type 131 C1-M27 Clade with a Multidrug-Resistant Phenotype

Michele Mutti,^a Ágnes Sonnevend,^b Tibor Pál,^b Sini Junttila,^c Heinz Ekker,^d Bence Galik,^c Attila Gyenesei,^c Gábor Nagy,^a Eszter Nagy,^a Valéria Szijártó^a

^aArsanis Biosciences GmbH, Vienna, Austria

^bDepartment of Medical Microbiology and Immunology, United Arab Emirates University, Al Ain, United Arab Emirates

^cBioComp, Bioinformatics & Scientific Computing, Vienna Biocenter Core Facilities GmbH, Vienna, Austria

^dNext Generation Sequencing, Vienna Biocenter Core Facilities GmbH, Vienna, Austria

ABSTRACT The sequence type 131 (ST131)-H30 clone is responsible for a significant proportion of multidrug-resistant extraintestinal *Escherichia coli* infections. Recently, the C1-M27 clade of ST131-H30, associated with *bla*_{CTX-M-27}, has emerged. The complete genome sequence of *E. coli* isolate 81009 belonging to this clone, previously used during the development of ST131-specific monoclonal antibodies, is reported here.

Escherichia coli strain 81009 was isolated from the urine of a 76-year-old male patient treated at Tawam Hospital, Al Ain, United Arab Emirates (1), and has recently been deposited at the Polish Collection of Microorganisms (PCM-2857). This strain belongs to sequence type 131 (ST131) and expresses the O25b O-antigen (2). *E. coli* 81009 is resistant to monobactams, cephalosporins, streptomycin, and fluoroquinolones and carries the *bla*_{CTX-M-27} allele. The recently emerging ST131-H30 isolates expressing CTX-M-27 are considered to form a new clade (C1-M27) within this clone (3, 4).

Genomic DNA of *E. coli* 81009 was extracted with the Genomic-tip 100/G kit and Genome buffer set (Qiagen) from bacteria grown in LB supplemented with 100 µg/ml ampicillin to mid-log phase (optical density at 600 nm [OD₆₀₀], ~0.5). The genomic material was sequenced on a PacBio RS II platform (Pacific Biosciences, USA) and with the Illumina HiSeq 2500 platform with a paired-end (PE) 125-bp read length (Illumina, The Netherlands).

Hierarchical Genome Assembly Process 4 (HGAP4) (5) was used for *de novo* assembly of the PacBio reads, and Illumina short reads were aligned to the *de novo* assembled genome with the Burrows-Wheeler Alignment tool (BWA) (6), followed by indel realignment and base quality recalibration with Genome Analysis Toolkit (GATK) (7). Protein-coding genes were called with Prodigal within the Prokka annotation pipeline (8). Trusted protein sequences for the annotation step were obtained from *E. coli* ST131 EC958 (9). Additional coding and noncoding RNA genes, as well as signal peptides, were predicted by RNAmmer, ARAGORN, SignalP, and Infernal within the Prokka annotation pipeline. The PlasmidFinder version 1.3 server (10) and ResFinder version 2.1 server (11) were used to identify plasmid *repA* genes and antibiotic resistance determinants.

E. coli 81009 has a circular chromosome of 5,012,084 bp and one plasmid (pEC-81009) with a length of 135,720 bp. The chromosome has a GC content of 50.8% and contains 4,987 predicted genes encoding 4,668 proteins and 319 RNAs (including 87 tRNAs, 22 rRNAs, and 1 transfer-messenger RNAs [tmRNAs]). A putative function was assigned to 91.7% of the protein-coding genes. The extraintestinal pathogen designation of *E. coli*

Received 15 January 2018 Accepted 17 January 2018 Published 22 February 2018

Citation Mutti M, Sonnevend Á, Pál T, Junttila S, Ekker H, Galik B, Gyenesei A, Nagy G, Nagy E, Szijártó V. 2018. Complete genome sequence of *Escherichia coli* 81009, a representative of the sequence type 131 C1-M27 clade with a multidrug-resistant phenotype. Genome Announc 6:e00056-18. <https://doi.org/10.1128/genomeA.00056-18>.

Copyright © 2018 Mutti 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 Valéria Szijártó, valeria.szijarto@arsanis.com.

81009 was supported by an array of virulence determinants identified *in silico* (12), including those involved in adhesion (*upaB*, *crl*, *csgA*, *fimH*, and *iha*), biofilm formation (*bscA*, *agn43*, and F9 fimbriae), survival in body fluids (*iss* and *gad*), and toxin production (*senB* and *sat*). *E. coli* 81009 was also shown to carry the *H30* allele of the *fimH* gene. Additionally, an 11.8-kb region (M27PP1) specific to the C1-M27 clade of the *H30* subclone (4) was identified, confirming that strain 81009 belongs to the C1-M27 clade.

Plasmid pEC-81009 was found to belong to the IncFII-FIA-FIB group by *in silico* replicon typing (10). Among the 160 predicted genes carried on this plasmid, toxin/antitoxin systems (Phd/Doc, CcdBA, and Kid/Kis) and genes encoding resistance to aminoglycosides (*aadA5*, *strA*, and *strB*), sulfonamide (*sul1* and *sul2*), tetracycline [*tet(A)*], macrolides (*mphA*), trimethoprim (*dfpA17*), and the extended-spectrum β -lactamase (ESBL) gene *bla*_{CTX-M-27} were identified.

The availability of *E. coli* 81009 in a public repository and its well-characterized phenotype and genotype make it a suitable model organism for studying the ST131 C1-M27 clade.

Accession number(s). The complete sequences of the chromosome and plasmid of *E. coli* 81009 have been deposited in GenBank under accession numbers [CP021179](#) and [CP021180](#), respectively.

ACKNOWLEDGMENTS

We thank Katharina Ramoni for assistance in the preparation of the genomic DNA (gDNA) samples and Agnieszka Korzeniowska-Kowal from PCM for her help during the deposition of *E. coli* 81009.

The PacBio sequencing was performed at the ETHZ/UZH Functional Genomics Center, Zurich, Switzerland.

M.M., G.N., E.N., and V.S. are employees of Arsanis Biosciences GmbH and hold shares in Arsanis, Inc., the parent company of Arsanis Biosciences GmbH. S.J., H.E., B.G., and A.G. are employees of the Vienna Biocenter Core Facilities GmbH, a member of the Vienna Biocenter, Vienna, Austria. The Illumina sequencing and the assembly and annotation of the PacBio data were done at the Next Generation Sequencing and BioComp facilities at the Vienna Biocenter Core Facilities GmbH, a member of the Vienna Biocenter, Vienna, Austria.

REFERENCES

- Szjártó V, Pal T, Nagy G, Nagy E, Ghazawi A, Al-Haj M, El Kurdi S, Sonnevend A. 2012. The rapidly emerging ESBL-producing *Escherichia coli* O25-ST131 clone carries LPS core synthesis genes of the K-12 type. *FEMS Microbiol Lett* 332:131–136. <https://doi.org/10.1111/j.1574-6968.2012.02585.x>.
- Szjártó V, Lukaszewicz J, Gozdiewicz TK, Magyarics Z, Nagy E, Nagy G. 2014. Diagnostic potential of monoclonal antibodies specific to the unique O-antigen of multidrug-resistant epidemic *Escherichia coli* clone ST131-O25b:H4. *Clin Vaccine Immunol* 21:930–939. <https://doi.org/10.1128/CVI.00685-13>.
- Birgy A, Bidet P, Levy C, Sobral E, Cohen R, Bonacorsi S. 2017. CTX-M-27-producing *Escherichia coli* of sequence type 131 and clade C1-M27, France. *Emerg Infect Dis* 23:885. <https://doi.org/10.3201/eid2305.161865>.
- Matsumura Y, Pitout JDD, Gomi R, Matsuda T, Noguchi T, Yamamoto M, Peirano G, De Vinney R, Bradford PA, Motyl MR, Tanaka M, Nagao M, Takakura S, Ichihama S. 2016. Global *Escherichia coli* sequence type 131 clade with *bla*_{CTX-M-27} gene. *Emerg Infect Dis* 22:1900–1907. <https://doi.org/10.3201/eid2211.160519>.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Forde BM, Ben Zakour NL, Stanton-Cook M, Phan M-D, Totsika M, Peters KM, Chan KG, Schembri MA, Upton M, Beatson SA. 2014. The complete genome sequence of *Escherichia coli* EC958: a high quality reference sequence for the globally disseminated multidrug resistant *E. coli* O25b:H4-ST131 clone. *PLoS One* 9:e104400. <https://doi.org/10.1371/journal.pone.0104400>.
- Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>.
- Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640–2644. <https://doi.org/10.1093/jac/dks261>.
- Johnson JR, Murray AC, Gajewski A, Sullivan M, Snippes P, Kuskowski MA, Smith KE. 2003. Isolation and molecular characterization of nalidixic acid-resistant extraintestinal pathogenic *Escherichia coli* from retail chicken products. *Antimicrob Agents Chemother* 47:2161–2168. <https://doi.org/10.1128/AAC.47.7.2161-2168.2003>.