

Complete Genome Sequences of Four *Escherichia coli* ST95 Isolates from Bloodstream Infections

Craig M. Stephens,^a Jeffrey M. Skerker,^{b,c} Manraj S. Sekhon,^d Adam P. Arkin,^{b,c} Lee W. Riley^d

Biology Department and Public Health Program, Santa Clara University, Santa Clara, California, USA^a; Department of Bioengineering, University of California, Berkeley, California, USA^b; Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA^c; Division of Infectious Diseases and Vaccinology, School of Public Health, University of California, Berkeley, California, USA^d

Finished genome sequences are presented for four *Escherichia coli* strains isolated from bloodstream infections at San Francisco General Hospital. These strains provide reference sequences for four major *fimH*-identified sublineages within the multilocus sequence type (MLST) ST95 group, and provide insights into pathogenicity and differential antimicrobial susceptibility within this group.

Received 11 September 2015 Accepted 28 September 2015 Published 5 November 2015

Citation Stephens CM, Skerker JM, Sekhon MS, Arkin AP, Riley LW. 2015. Complete genome sequences of four *Escherichia coli* ST95 isolates from bloodstream infections. *Genome Announc* 3(6):e01241-15. doi:10.1128/genomeA.01241-15.

Copyright © 2015 Stephens 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 Craig M. Stephens, cstephens@scu.edu.

The Gram-negative bacterium *Escherichia coli* can be associated with intestinal and extraintestinal infections. “ExPEC” (extraintestinal pathogenic *E. coli*) strains are a major cause of human urinary tract infections, as well as bloodstream infections arising as a consequence of pyelonephritis, a complication of urinary tract infections (1). Antimicrobial resistance among uropathogenic *E. coli* strains is an ongoing clinical challenge (2). Most ExPEC isolates derive from a small number of globally distributed pandemic lineages defined by multilocus sequence typing (MLST) (1). The ST95 MLST group is one of the most frequently observed, and is notable in showing a relatively low frequency of antimicrobial resistance. We report here complete genome sequences from ST95 strains representing four major sublineages associated with distinct alleles of the *fimH* gene (3, 4). These sequences provide references for comparative analysis of drug resistance and host niche adaptation within and beyond the ST95 clade.

The four *E. coli* strains examined here (SF-088, SF-166, SF-173, and SF-468) were isolated from patients with bloodstream infections at San Francisco General Hospital between 2007 and 2010 (4, 5). Sequencing used the Pacific Biosciences (PacBio) RSII platform (Pacific Biosciences, Menlo Park, CA, USA). DNA samples were initially purified using a Qiagen DNeasy tissue extraction kit (Qiagen, USA), followed by phenol-chloroform extraction, and sheared to an average fragment size of 20 kb (Diagenode, Liege, Belgium) before conversion into a sequencing library using the SMRTbell Template Prep Kit 1.0 (Pacific Biosciences). Each library was sequenced using P6C4 chemistry on a single-molecule real-time (SMRT) cell with a 240-min collection protocol. Reads were *de novo* assembled and polished using SMRT Analysis v2.3.0 and the HGAP2 or HGAP3 algorithms. For Illumina sequencing, libraries were prepared for 300 bp paired-end reads by incorporating index tags (Wafergen Biosystems). Libraries were sequenced on a MiSeq instrument using V3 chemistry. MiSeq data were used to correct errors in the HGAP assemblies using Pilon (6). Plasmids less than 10 kb in length were assembled solely from

MiSeq reads. Annotation of genomes used RAST (7) and the NCBI prokaryotic genome annotation pipeline.

The SF-468 (*fimH-1*) genome includes a 5.14-Mb chromosome and five plasmids (111, 94, 6.7, 4.1, and 1.6 kb). The SF-166 (*fimH-6*) genome is comprised of a 4.91-Mb chromosome and a 114-kb plasmid. The SF-088 (*fimH-9*) genome includes a 5.05-Mb chromosome and three plasmids (150, 5.2, and 3.9 kb). The SF-173 (*fimH-47*) genome is comprised of a 5.06-Mb chromosome and a 93-kb plasmid. The chromosomes are almost completely colinear, with size differences primarily due to prophage. The largest plasmids in each strain are conjugal IncF replicons, and with the exception of SF-166 carry multiple antimicrobial resistance genes. The 94-kb IncI1 plasmid in SF-468 also carries the *blaCTX-M-14* ESBL gene. Other than genes encoding the conjugal machinery in the IncF plasmids, there is relatively little similarity in the remaining plasmid content. More detailed analysis of virulence factors, plasmids, and antibiotic resistance among *E. coli* ST95 strains will be presented in subsequent publications.

Nucleotide sequence accession numbers. These sequences have been deposited at DDBJ/EMBL/GenBank under the following accession numbers: SF-468 chromosome and plasmids, CP012625 to CP012630; SF-173 chromosome and plasmid, CP012631 and CP012632; SF-166 chromosome and plasmid, CP012633 and CP012634; SF-088 chromosome and plasmids, CP012635 to CP012638.

ACKNOWLEDGMENTS

We thank the staff of the UC Davis Genome Center and the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley for their assistance in this project.

Support for this project was provided by the RB Roberts Fund for Drug-Resistant Infectious Disease Research. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, which is supported by NIH S10 Instrumentation grants S10RR029668 and S10RR027303.

REFERENCES

1. Riley LW. 2014. Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. *Clin Microbiol Infect* 20:380–390. <http://dx.doi.org/10.1111/1469-0691.12646>.
2. World Health Organization. 2014. Antimicrobial resistance: global report on surveillance. World Health Organization, Geneva, Switzerland.
3. Dias RCS, Moreira BM, Riley LW. 2010. Use of *fimH* single-nucleotide polymorphisms for strain typing of clinical isolates of *Escherichia coli* for epidemiologic investigation. *J Clin Microbiol* 48:483–488. <http://dx.doi.org/10.1128/JCM.01858-09>.
4. Adams-Sapper S, Diep BA, Perdreau-Remington F, Riley LW. 2013. Clonal composition and community clustering of drug-susceptible and -resistant *Escherichia coli* isolates from bloodstream infections. *Antimicrob Agents Chemother* 57:490–497. <http://dx.doi.org/10.1128/AAC.01025-12>.
5. Adams-Sapper S, Sergeevna-Selezneva J, Tartof S, Raphael E, Diep BA, Perdreau-Remington F, Riley LW. 2012. Globally dispersed mobile drug-resistance genes in Gram-negative bacterial isolates from patients with bloodstream infections in a US urban general hospital. *J Med Micro* 61: 968–974. <http://dx.doi.org/10.1099/jmm.0.041970-0>.
6. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <http://dx.doi.org/10.1371/journal.pone.0112963>.
7. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42: D206–D214. <http://dx.doi.org/10.1093/nar/gkt1226>.