Complete genome sequence of the epidemic and highly virulent CTX-M-15-producing H30-Rx subclone of Escherichia coli ST131

Paal S. Andersen  
Statens Serum Institut, Copenhagen, Denmark

Marc Stegger  
Statens Serum Institut, Copenhagen, Denmark

Maliha Aziz  
Translational Genomics Research Institute, Flagstaff, AZ

Tania L. Contente-Cuomo  
Translational Genomics Research Institute, Flagstaff, AZ

Henry S. Gibbons  
Northern Arizona University

See next page for additional authors

Follow this and additional works at: https://hsrc.himmelfarb.gwu.edu/sphhs_enviro_facpubs

Part of the Environmental Public Health Commons, and the Occupational Health and Industrial Hygiene Commons

Recommended Citation

This Journal Article is brought to you for free and open access by the Environmental and Occupational Health at Health Sciences Research Commons. It has been accepted for inclusion in Environmental and Occupational Health Faculty Publications by an authorized administrator of Health Sciences Research Commons. For more information, please contact hsrc@gwu.edu.
We report the complete genome sequence, including five complete plasmid sequences, of *Escherichia coli* ST131 isolate JJ1886. The isolate was obtained in 2007 in the United States from a patient with fatal urosepsis and belongs to the virulent, CTX-M-15-producing *H30-Rx* sublineage.

*Escherichia coli* sequence type 131 (ST131) has emerged as one of the most prevalent extraintestinal pathogenic *E. coli* lineages in circulation today (1). The ST131 *H30* lineage dominates among *Escherichia coli* strains and is associated with recurrent urinary tract infections, pyelonephritis, and sepsis (1, 2, 3, 4, 5). Phylogenomic analysis indicated that the dominant ST131 ESBL gene, *bla*<sub>CTX-M-15</sub>, is associated with an *H50* sublineage, designated *H50-Rx* (6). Here, we present the complete genome sequence of JJ1886, an ST131 *H50*-Rx isolate recovered from a patient with fatal urosepsis (7).

The genome sequence was assembled using 320 Mb of paired-end 100-bp HiSeq reads (Illumina, Hayward, CA), 225 Mb of 500-bp reads from the Roche Genome Sequencer FLX (Roche Diagnostics, Switzerland), and 187,000 reads (296 Mb) from the PacBio RS platform (Pacific Biosciences, Menlo Park, CA). The PacBio sequence reads were error corrected using HiSeq data (8). Sequences were assembled using multiple *de novo* assemblers, including MIRA (9), ABySS (10), and CLCbio (CLCbio, Denmark). Initial assembly of the Roche and HiSeq data yielded 52 chromosomal contigs after extension using PBjelly and PAGIT (11, 12). PacBio reads further reduced the number of chromosomal contigs to 18. Gaps were closed *in silico* using Genomics workbench 6.0.4 (CLCbio) by a combination of comparisons to *E. coli* reference genome sequences, with manual curation using HiSeq and PacBio data and verification using an optical map (OpGen, MD). Five complete plasmids were identified based on BLAST analysis and were verified using Illumina MiSeq sequencing (250-bp paired end) on isolated plasmids. The genome sequence was annotated using RAST (13).

The complete genome of JJ1886 comprises a 5,129,938-bp chromosome with a GC content of 50.8%, with 5,086 coding sequences, 88 tRNAs, and 22 rRNA features, plus five plasmids, pJJ1886-1 through pJJ1886-5, of sizes 1.6, 5.2, 5.6, 56, and 110 kb, respectively. According to ResFinder 1.4 (14), only pJJ1886-5 carries genes for resistance, including resistance to aminoglycosides and fluoroquinolones ([*aac(6’)-Ib, aac(6’)-Ib-cr, aac(6’)-31 (~85% sequence homology)*], beta-lactams ([*bla*<sub>OXA-23</sub>, *bla*<sub>TEM-1</sub>]), and chloramphenicol ([*catB3*]). However, the ESBL gene *bla*<sub>CTX-M-15</sub> was integrated into the JJ1886 chromosome via an incomplete Tn3 transposable element embedded within a lambda-like 58-kb prophage (identified by PHAST analysis [15]). This is in contrast with the only other complete, published ST131-*H30* genome sequence, that of NA114 (isolated from a patient with prostatitis [16]), which likely has *bla*<sub>CTX-M-15</sub> integrated into a previously described plasmid (17).

JJ1886 is the first complete genome sequence for a urosepsis ST131 isolate that includes all plasmids and has a chromosomal *bla*<sub>CTX-M-15</sub> integration site. The genome sequence will serve as a valuable resource for studies on the epidemiology and pathogenicity of the highly virulent ST131 lineage.

**Nucleotide sequence accession numbers.** The complete sequences of the chromosome of *E. coli* JJ1886 and its five plasmids, pJJ1886-1 through pJJ1886-5, have been deposited in GenBank (accession numbers CP006784, CP006785, CP006786, CP006787, CP006788, and CP006789, respectively).

**ACKNOWLEDGMENTS**

This material is based upon work supported by the Office of Research and Development, Medical Research Service, Department of Veterans Affairs, grant 101 CX000192 01 (to J. R. J.); NIH grant RC4-AI092828 (to E. V. S. and J. R. J.); the TGen Foundation (to L. B. P. and M. A.); the Statens Serum Institut (to P. A. and M. S.); the Augustinus Fonden (to P. A.); USAMRMC grant W81XWH-10-1-0753 (to L. B. P.); and the Transformational Medical Technologies Initiative of the Defense Threat Reduction Agency project no. CB2847 (to H. S. G.).

Patricia Stogsdill, Lonny Yarmus, Rob Owens, John Quinn, and Karen Lolans facilitated provision of *E. coli* JJ1886 and associated clinical data. Opinions presented here are those of the authors and do not represent the official policy or position of the U.S. Government or any of its agencies. This work has been cleared for public release.
REFERENCES


