Complete Genome Sequence of the Neonatal-Meningitis-Associated *Escherichia coli* Strain CE10

Shuting Lu,¹ Xiaobing Zhang,¹ Yafang Zhu,¹ Kwang Sik Kim,² Jian Yang,^{1*} and Qi Jin^{1*}

State Key Laboratory for Molecular Virology and Genetic Engineering, Institute of Pathogen Biology, Chinese Academy Medical Sciences & Peking Union Medical College, Beijing 100176, China,¹ and Department of Pediatrics, Division of Pediatric Infectious Diseases, Johns Hopkins University School of Medicine, Baltimore, Maryland 21287²

Received 29 September 2011/Accepted 10 October 2011

Neonatal bacterial meningitis continues to be an important cause of mortality and morbidity worldwide. *Escherichia coli* possessing the K1 capsular polysaccharide is the most common Gram-negative pathogen causing neonatal meningitis. Here we present the complete genome sequence of neonatal meningitis-associated *E. coli* strain CE10, a unique K1 strain with a functional type III secretion system. Functional analysis of the genome should enhance our knowledge of the pathogenesis of neonatal *E. coli* K1 meningitis.

Despite advances in antibiotics, neonatal bacterial meningitis remains a significant threat to the health of newborns worldwide (3, 4). Neonatal meningitis-associated *Escherichia coli* (NMEC) is one of the leading causative agents of neonatal meningitis (4). *E. coli* strains possessing the K1 capsular polysaccharide are predominant among NMEC isolates, and the K1 strains isolated from cerebrospinal fluid (CSF) can be categorized into two distinct groups in terms of genomic background, which are likely to utilize different pathogenic mechanisms to cause meningitis (7). *E. coli* O7:K1 strain CE10 (previously named strain EC10), which was isolated from the CSF of a neonate with meningitis, is a representative of group 2 isolates from CSF (7). Strain CE10 was found to carry a complete gene set encoding the type III secretion system (T3SS), which was recently demonstrated to be involved in the invasion and intracellular survival of the bacterium in human brain microvascular endothelial cells (8). Most of the K1 isolates from CSF belong to phylogenetic group B2 and, to a less extent, to group D (7). Strain CE10 belongs to group D, and the unique characteristics of the strain make it an interesting focus for pathogenesis and evolution studies of *E. coli* K1 strains.

The whole genome was sequenced using a Roche/454 FLX genome sequencer system, and the raw data set contains 503,963 reads, totaling 204 Mbp (~38-fold coverage). The initial assembly, performed using Newbler software, yielded 144 large (>500-bp) contigs. Gaps were then closed on the Consed workbench, using genomic PCR and primer walking with conventional Sanger sequencing (2). Glimmer3 and tRNAscan-SE were used to predict protein-coding sequences (CDS) and tRNA genes, respectively (1, 5). Functional annotation of CDS was performed by BLAST searches of the GenBank nonredundant protein database followed by manual curation.

The whole genome of CE10 is composed of a 5,313,531-bp chromosome and four plasmids. The chromosome (50.6% GC content) is predicted to possess 5,009 CDS, 76 pseudogenes, 91 tRNAs, and 7 rRNA operons. It shares a colinear backbone of ~4.1 Mb with the

chromosome of strain S88, another *E. coli* K1 strain (O45:K1; B2 group). However, the results of genome-based phylogenetic analysis indicated that CE10 is closely related to uropathogenic *E. coli* (UPEC) strain IAI39 (O7:K1) rather than to other K1 strains. Indeed, a total of about 5.0 Mb (93.4%) of the CE10 chromosome is highly conserved with respect to the IAI39 chromosome, with identity > 99%, but the collinearity between them is severely interrupted by genomic translocations and inversions, which are probably recent events due to niche-specific adaptations.

Genome sequencing of CE10 confirmed the presence of *epr-epa-eiv* operons encoding the T3SS apparatus. Furthermore, we also identified genes scattered throughout the CE10 genome that encode putative T3SS effectors, including homologues of EspL, EspR, EspX, and EspY, which are experimentally confirmed T3SS effectors in enterohemorrhagic *E. coli* (6). Functional studies of the T3SS effectors may shed light on the pathogenic mechanism of the T3SS in strain CE10 that is employed for traversal of the blood-brain barrier and penetration into the central nervous system.

Nucleotide sequence accession numbers. Complete sequences of *E. coli* CE10 have been deposited in GenBank under accession no. CP003034 to CP003038.

This work was supported by the National Basic Research Program from the Ministry of Science and Technology of China (grant 2009CB522603) and by Beijing Nova Program grant 2009A67 (to J.Y.).

REFERENCES

1. Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**:4636–4641.
2. Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* **8**:195–202.
3. Kim, K. S. 2003. Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury. *Nat. Rev. Neurosci.* **4**:376–385.
4. Kim, K. S. 2008. Mechanisms of microbial traversal of the blood-brain barrier. *Nat. Rev. Microbiol.* **6**:625–634.
5. Lowe, T. M., and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**:955–964.
6. Tobe, T., et al. 2006. An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination. *Proc. Natl. Acad. Sci. U. S. A.* **103**:14941–14946.
7. Yao, Y., Y. Xie, and K. S. Kim. 2006. Genomic comparison of *Escherichia coli* K1 strains isolated from the cerebrospinal fluid of patients with meningitis. *Infect. Immun.* **74**:2196–2206.
8. Yao, Y., et al. 2009. The type III secretion system is involved in the invasion and intracellular survival of *Escherichia coli* K1 in human brain microvascular endothelial cells. *FEMS Microbiol. Lett.* **300**:18–24.

* Corresponding author. Mailing address: Institute of Pathogen Biology, CAMS & PUMC, 6 Rongjing Eastern Street, BDA, Beijing 100176, China. Phone: (8610) 6787 7735. Fax: (8610) 6787 7736. E-mail for Qi Jin: zdsys@vip.sina.com. E-mail for Jian Yang: yangj@ipbcbms.ac.cn.