Complete Genome Sequence of the *Escherichia coli* PMV-1 Strain, a Model Extraintestinal Pathogenic *E. coli* Strain Used for Host-Pathogen Interaction Studies

Francesc Peris-Bondia,[a] Eric Muraille,[b] Laurence Van Melderen[a]

Laboratoire de Génétique et Physiologie Bactérienne, IBMM, Faculté des Sciences, Université Libre de Bruxelles (ULB), Brussels, Belgium[a]; Laboratoire de Parasitologie, Faculté de Médecine, Université Libre de Bruxelles (ULB), Brussels, Belgium[b]

***Escherichia coli* is a highly versatile species, causing diverse intestinal and extraintestinal infections. Here, we present the complete genome sequence of PMV-1, an O18:K1 extraintestinal pathogenic *E. coli* (ExPEC) strain that is used as a model for peritonitis in mice and was useful for deciphering the innate immune response triggered by ExPEC infections.**

Address correspondence to Laurence Van Melderen, lvmelder@ulb.ac.be.

*E*scherichia coli causes a broad spectrum of disease that relies on the acquisition of specific virulence factors, which confer an increased ability to adapt to new niches. In that context, the genome of the *E. coli* PMV-1 strain and its plasmids were sequenced. This strain is an appropriate model for host-pathogen interactions since it has been proven to be highly virulent after intraperitoneal administration in mouse models (1, 2). The whole genome was sequenced to a depth of 40× coverage using pyrosequencing by Macrogen Inc. *E. coli* strain UTI89 was used as a backbone to perform the assembly by MIRA (3) and the Staden package (4). Phylogenetic analysis was inferred using ClonalFrame (5) with three iterations between samples and representative genomes of *E. coli*. Protein-coding sequences (CDSs) and noncoding RNAs (ncRNAs) were identified using Prokka (6) and were compared with Glimmer (7) results. The functional annotation of the CDSs was performed using BLASTp (8) against the nr protein database (9) and VFDB (10). Genes for tRNAs, transfer-messenger RNA (tmRNA), rRNAs, and other small RNAs were identified using the Rfam database (11) and Infernal (12). In order to identify the genomic islands (GIs), a reciprocal best-hit blast with ≥70% identity was performed versus *E. coli* K-12 MG1655. GIs were defined by ≥10 consecutive CDSs that were not homologous to K-12 CDSs. Prophages were detected using PHAST (13).

The genome of PMV-1 comprises a circular chromosome of 4,984,940 bp and two plasmids of 98,864 bp and >127,123 bp (in two contigs of 83,288 and 43,835 bp). Phylogenetic analysis identified *E. coli* IHE3034, an extraintestinal pathogenic *E. coli* (ExPEC) O18:K1 strain (14), as the closest whole-genome-sequenced strain. The 98-kb plasmid is similar to the IncI pHUSEC41-1 plasmid (15) and comprises 106 open reading frames (ORFs). It lacks a 7,237-bp region flanked by transposase- and resolvase-encoding genes containing predicted β-lactamase, streptomycin phosphotransferase, and aminoglycoside resistance genes. This region is replaced by an IncQ incompatibility system, arsenic resistance, phosphoglucosamine mutase, and sulfonamide resistance genes. The other two scaffolds are part of an incomplete sequenced plasmid similar to pECOS88 (16). This IncFI-IncFII plasmid is associated with neonatal meningitis cases and harbors iron uptake systems (aerobactin, salmochelin, and *sitA*). Thirty-seven GIs of >5 kb carrying virulence factors, such as genes for S-fimbriae and iron acquisition systems and capsule-encoding genes, as well as a putative type VI secretion system (T6SS) and several predicted *hcp* genes (17), were identified. In addition, 5 GIs containing antibiotic resistance genes (encoding β-lactamase, efflux pumps, and multidrug transporters) were detected. Among the 9 prophages that were identified, 3 carry virulence factors, an iron-manganese transport system (*sit*), a putative *O*-acetyltranferase (*neuO*) (18), and type 1 fimbriae (*fim*). Despite the cytotoxic necrotizing factor 1 (CNF1) known to be involved in blood-brain barrier crossing (19), this strain carries all the virulence factors responsible for meningitis and sepsis (20).

Nucleotide sequence accession numbers. The annotated genome sequences of *E. coli* PMV-1 have been deposited in public databases under the accession no. HG428755 for the complete genome and under accession no. HG428756 and CBTO010000001 to CBTO010000002 for *E. coli* PMV-1 plasmids pHUSEC41-1-like and pECOS88-like, respectively.

REFERENCES

1. **Sewnath ME, Olszyna DP, Birjmohun R, ten Kate FJ, Gouma DJ, van der Poll T.** 2001. IL-10-deficient mice demonstrate multiple organ failure and increased mortality during *Escherichia coli* peritonitis despite an accelerated bacterial clearance. J. Immunol. **166:**6323–6331.
2. **van Westerloo DJ, Giebelen IA, Florquin S, Daalhuisen J, Bruno MJ, de Vos AF, Tracey KJ, van der Poll T.** 2005. The cholinergic anti-inflammatory pathway regulates the host response during septic peritonitis. J. Infect. Dis. **191:**2138–2148.
3. **Chevreux B.** 2005. MIRA: an automated genome and EST assembler. Ph.D. thesis. The Ruprecht-Karls-University, Heidelberg, Germany.

4. **Bonfield JK, Smith KF, Staden R.** 1995. A new DNA sequence assembly program. Nucleic Acids Res. **23**:4992–4999.

5. **Didelot X, Falush D.** 2007. Inference of bacterial microevolution using multilocus sequence data. Genetics **175**:1251–1266.

6. **Seemann T.** Prokka: prokaryotic genome annotation system. Victorian Bioinformatics Consortium, Clayton, Australia. http://vicbioinformatics.com/.

7. **Delcher AL, Harmon D, Kasif S, White O, Salzberg SL.** 1999. Improved microbial gene identification with Glimmer. Nucleic Acids Res. **27**:4636–4641.

8. **Ye J, Mcginnis S, Madden TL.** 2006. BLAST: improvements for better sequence analysis. Nucleic Acids Res. **34**:6–9.

9. **Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA, Wagner L.** 2003. Database resources of the National Center for Biotechnology. Nucleic Acids Res. **31**:28–33.

10. **Chen L, Xiong Z, Sun L, Yang J, Jin Q.** 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res. **40**:D641–D645. doi:10.1093/nar/gkr989.

11. **Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, Eddy SR, Gardner PP, Bateman A.** 2012. Rfam 11.1: 10 years of RNA families. Nucleic Acids Res. **41**:D226–D232. doi:10.1093/nar/gks1005.

12. **Nawrocki EP, Kolbe DL, Eddy SR.** 2009. Infernal 1.0: inference of RNA alignments. Bioinformatics **25**:1335–1337.

13. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS.** 2011. PHAST: a fast phage search tool. Nucleic Acids Res. **39**:W347–W352. doi:10.1093/nar/gkr485.

14. **Moriel DG, Bertoldi I, Spagnuolo A, Marchi S, Rosini R, Nesta B, Pastorello I, Corea VA, Torricelli G, Cartocci E, Savino S, Scarselli M, Dobrindt U, Hacker J, Tettelin H, Tallon LJ, Sullivan S, Wieler LH, Ewers C, Pickard D, Dougan G, Fontana MR, Rappuoli R, Pizza M, Serino L.** 2010. Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **107**:9072–9077.

15. **Künne C, Billion A, Mshana SE, Schmiedel J, Domann E, Hossain H, Hain T, Imirzalioglu C, Chakraborty T.** 2012. Complete sequences of plasmids from the hemolytic-uremic syndrome-associated *Escherichia coli* strain HUSEC41. J. Bacteriol. **194**:532–533.

16. **Peigne C, Bidet P, Mahjoub-Messai F, Plainvert C, Barbe V, Médigue C, Frapy E, Nassif X, Denamur E, Bingen E, Bonacorsi S.** 2009. The plasmid of *Escherichia coli* strain S88 (O45:K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E. coli* plasmids and is associated with high-level bacteremia in a neonatal rat meningitis model. Infect. Immun. **77**:2272–2284.

17. **Cascales E.** 2008. The type VI secretion toolkit. EMBO Rep. **9**:735–741.

18. **King MR, Vimr RP, Steenbergen SM, Spanjaard L, Plunkett G, Blattner FR, Vimr ER.** 2007. Escherichia coli K1-specific bacteriophage CUS-3 distribution and function in phase-variable capsular polysialic acid O acetylation. J. Bacteriol. **189**:6447–6456.

19. **Khan NA, Wang Y, Kim KJ, Chung JW, Wass CA, Kim KS.** 2002. Cytotoxic necrotizing factor-1 contributes to *Escherichia coli* K1 invasion of the central nervous system. J. Biol. Chem. **277**:15607–15612.

20. **Croxen MA, Finlay BB.** 2010. Molecular mechanisms of *Escherichia coli* pathogenicity. Nat. Rev. Microbiol. **8**:26–38.