Draft genome sequence of Escherichia coli MS499, isolated from the infected uterus of a postpartum cow with metritis

Robert J Goldstone, Richard Talbot, Hans-Joachim Schuberth, Olivier Sandra, I. Martin Sheldon, David G.E. Smith

Draft Genome Sequence of *Escherichia coli* MS499, Isolated from the Infected Uterus of a Postpartum Cow with Metritis

Robert J. Goldstone,[a] Richard Talbot,[b] Hans-Joachim Schuberth,[c] Olivier Sandra,[d,e] I. Martin Sheldon,[f] David G. E. Smith[a,g]

Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom[a]; The Roslin Institute, University of Edinburgh, Easter Bush, Midlothian, Scotland, United Kingdom[b]; Immunology Unit, University of Veterinary Medicine, Hannover, Germany[c]; INRA, UMR1198 Biologie du Développement et Reproduction, Jouy-en-Josas, France[d]; ENVA, Maisons Alfort, France[e]; Institute of Life Science, School of Medicine, Swansea University, Singleton Park, Swansea, United Kingdom[f]; Moredun Research Institute, Pentlands Science Park, Bush Loan, Edinburgh, Midlothian, United Kingdom[g]

Specific *Escherichia coli* strains associated with bovine postpartum uterine infection have recently been described. Many recognized virulence factors are absent in these strains; therefore, to define a prototypic strain, we report here the genome sequence of *E. coli* isolate MS499 from a cow with the postpartum disease metritis.

Address correspondence to David G. E. Smith, david.g.smith@glasgow.ac.uk.

*E*scherichia coli is a diverse species, and several of its pathotypes are well-characterized, including extraintestinal pathogenic *E. coli* (ExPEC). ExPEC infections of the udder and uterus of cattle have been of increasing significance, and *E. coli* is a cause of postpartum uterine disease (metritis) in cattle (1–3) and was recently proposed as a novel pathotype termed intrauterine *E. coli* or endometrial pathogenic *E. coli* (EnPEC) (2, 4). To date, genomic characterization of *E. coli* has focused on phylogenetic grouping, genotyping, and pathotyping based on the detection of defined virulence-associated factors. Among metritis isolates, phylogroup B1 strains have predominated (>65%), and these lack many virulence-associated factors typical of other *E. coli* pathotypes. Thus, we undertook genome sequencing of an EnPEC strain to define a prototype for this pathotype.

E. coli MS499 isolated from the uterus of an animal with metritis was previously confirmed as phylogroup B1 and possessed only *fyuA* of 17 fitness- and virulence-associated genes (4). The isolate was cultured on LB agar plates, and genomic DNA was prepared using the DNeasy blood and tissue kit (Qiagen) according to the manufacturer's instructions. Sequencing was carried out using 454 GS-FLX and Illumina GA IIx 120-bp paired-end sequencing. The reads were assembled using Velvet (5) and the Celera assembler with the best overlap graph (CABOG) (6), and gaps were closed using unmapped 454 and Illumina reads. The assembled 447 contigs totaled 4,883,598 bp, with a G+C content of 50.7%. The assembled contigs were subjected to preliminary annotation using the RAST resource (release 70) (7) and subsequently using the NCBI annotation tool. The genome was predicted to encode 4,708 proteins. Several contigs assembled into a large plasmid showing high similarity to plasmids that confer functions in host colonization, persistence, and pathogenicity in *Salmonella enterica* serovar Kentucky and avian pathogenic *E. coli* (APEC) (8).

Frequently, *E. coli* strains of phylogroup B1 are considered nonpathogenic or poorly pathogenic (9) though strain MS499 contains many pathogenicity-associated determinants, including multiple fimbriae, iron acquisition systems, and autotransporter proteins (type 5 secretion system [T5SS]), as well as flagella. Putative toxins were absent, with the exception of an ortholog of the *hylF* hemolysin. Like many EnPEC isolates, *E. coli* MS499 is resistant to β-lactams (4), and a gene encoding a TEM-1 family β-lactamase was present.

E. coli MS499 from bovine metritis has many of the characteristics typifying ExPEC, i.e., elements that may contribute to pathogenicity, fitness, and resistance, and this initial genome sequence supports the previous proposal of EnPEC as a novel *E. coli* pathotype. This first reported genome sequence for EnPEC provides a resource for further defining this novel *E. coli* pathotype and forms the basis for systematic studies. Comparative genomic analyses are under way and will be reported separately in order to further define this *E. coli* pathotype.

ACKNOWLEDGMENTS

REFERENCES

1. **Bonnett BN, Martin SW, Gannon VP, Miller RB, Etherington WG.** 1991. Endometrial biopsy in Holstein-Friesian dairy cows. III. Bacteriological analysis and correlations with histological findings. Can. J. Vet. Res. **55:** 168–173.
2. **Bicalho RC, Machado VS, Bicalho ML, Gilbert RO, Teixeira AG, Caixeta LS, Pereira RV.** 2010. Molecular and epidemiological characterization of bovine intrauterine *Escherichia coli*. J. Dairy Sci. **93:**5818–5830. http://dx.doi.org/10.3168/jds.2010-3550.
3. **Sheldon IM, Cronin J, Goetze L, Donofrio G, Schuberth HJ.** 2009. Defining postpartum uterine disease and the mechanisms of infection and immunity in the female reproductive tract in cattle. Biol. Reprod. **81:** 1025–1032. http://dx.doi.org/10.1095/biolreprod.109.077370.

4. **Sheldon IM, Rycroft AN, Dogan B, Craven M, Bromfield JJ, Chandler A, Roberts MH, Price SB, Gilbert RO, Simpson KW.** 2010. Specific strains of *Escherichia coli* are pathogenic for the endometrium of cattle and cause pelvic inflammatory disease in cattle and mice. PLoS One **5:**e9192. http://dx.doi.org/10.1371/journal.pone.0009192.

5. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. **18:**821–829. http://dx.doi.org/10.1101/gr.074492.107.

6. **Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G.** 2008. Aggressive assembly of pyrosequencing reads with mates. Bioinformatics **24:**2818–2824. http://dx.doi.org/10.1093/bioinformatics/btn548.

7. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics **9:**75. http://dx.doi.org/10.1186/1471-2164-9-75.

8. **Johnson TJ, Nolan LK.** 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. Microbiol. Mol. Biol. Rev. **73:**750–774. http://dx.doi.org/10.1128/MMBR.00015-09.

9. **Köhler CD, Dobrindt U.** 2011. What defines extraintestinal pathogenic *Escherichia coli*? Int. J. Med. Microbiol. **301:**642–647. http://dx.doi.org/10.1016/j.ijmm.2011.09.006.