

Complete Genome Sequence of the Avian Pathogenic *Escherichia coli* Strain APEC O78

Paul Mangiamele,^a Bryon Nicholson,^a Yvonne Wannemuehler,^a Torsten Seemann,^b Catherine M. Logue,^a Ganwu Li,^a Kelly A. Tivendale,^{a,c} Lisa K. Nolan^a

Department of Veterinary Microbiology and Preventive Medicine, College of Veterinary Medicine, VMRI 2, Iowa State University, Ames, Iowa, USA^a; Victorian Bioinformatics Consortium, Monash University, Clayton, Victoria, Australia^b; Veterinary Science, the University of Melbourne, Parkville, Victoria, Australia^c

Colibacillosis, caused by avian pathogenic *Escherichia coli* (APEC), is a significant disease, causing extensive animal and financial losses globally. Because of the significance of this disease, more knowledge is needed regarding APEC's mechanisms of virulence. Here, we present the fully closed genome sequence of a typical avian pathogenic *E. coli* strain belonging to the serogroup O78.

Received 15 January 2013 Accepted 18 February 2013 Published 21 March 2013

Citation Mangiamele P, Nicholson B, Wannemuehler Y, Seemann T, Logue CM, Li G, Tivendale KA, Nolan LK. 2013. Complete genome sequence of the avian pathogenic *Escherichia coli* strain APEC O78. *Genome Announc.* 1(2):e00026-13. doi:10.1128/genomeA.00026-13.

Copyright © 2013 Mangiamele 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 Lisa K. Nolan, lnolan@iastate.edu.

Colibacillosis, caused by avian pathogenic *Escherichia coli* (APEC), is one of the most significant infectious diseases affecting poultry (1–7). Poultry colibacillosis takes many forms, with systemic forms occurring most often (2). Collectively, these diseases result in annual multimillion-dollar losses due to mortality, decreased production, and condemnations (1, 3, 5). Indeed, colibacillosis poses a profound threat to one of humankind's cheapest sources of high-quality animal protein. Despite the importance of this disease, the mechanisms of APEC virulence remain largely unknown. Studies into APEC pathogenesis would be enhanced by public access to high-quality genomic sequences. To date three APEC sequences are publically available. The sequence of APEC O1, an O1:K1:H7 strain isolated from the lung of a turkey, is fully closed (9). Obtained from gelatinous edema lesions from a laying hen, a draft sequence of a Brazilian APEC strain, SCI-07, a member of the O nontypeable:H31 serotype, is in 68 contigs (10), and a sequence of an O78 strain (χ 7122) was recently released in 12 contigs (11). Here, we describe a fully closed and annotated sequence of an O78 strain with the idea that fully closed sequences representative of the most commonly isolated APEC serogroups, such as O1 and O78 strains, are needed to adequately support future colibacillosis research (1).

APEC O78 is an O78 strain isolated from the lung of a turkey clinically diagnosed with colibacillosis. Genomic sequencing was performed using complementary sequencing technologies, combining results obtained with a Roche/454 FLX genome sequencer (GS) instrument and an Illumina Hi-Seq 2000. The following datasets were used in the final assembly: (i) GS-FLX, with 590,773 shotgun reads totaling 237 Mb (~49-fold coverage); (ii) GS-FLX 8-kb mate-pair library with 474,583 shotgun reads totaling 168 Mb (~35-fold coverage) of which 330,857 were paired; and (iii) Illumina 100 bp paired-end library with 27,389,600 reads totaling 2,587 Mb (~539-fold coverage). Both 454-read sets were assembled *de novo* using Newbler 2.6 (Roche), and Illumina reads were assembled separately with Velvet 1.1 (8) and Illumina's

ELANDv2e assembler. The genome was closed using 454 assemblies as a "reference" sequence and the Illumina data to add depth, correct errors, and close gaps. Whole-genome optical mapping (OpGen, Gaithersburg, MD) was used to validate scaffolds and contig order. The assembly was confirmed using PCR and Sanger sequencing and validated by consistency of paired-end evidence from 454 and Illumina reads.

Annotation was automated using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP). The final version was checked against the previously completed Prokka 1.5.2 annotation.

The assembled genome consists of a single chromosome (4,798,435 bp; 50.68% GC content) and two plasmids, one 217.830 kb and the other 113.260 kb. The chromosome contains 4,696 protein-encoding genes, 88 tRNA-carrying genes, and 19 rRNA-carrying operons. The chromosome of APEC O78 is smaller than many other fully sequenced extraintestinal pathogenic *E. coli* (ExPEC) genomes, and its chromosomal structure appears different from those of other ExPEC genomes. Assessment of the implications of these differences is ongoing, but the addition of a fully closed genomic sequence of one of the commonly occurring serogroups among APEC significantly contributes to the toolset that can be used in studies of APEC pathogenesis and colibacillosis control.

Nucleotide sequence accession number. Complete sequences of APEC O78 have been deposited in GenBank under accession number [CP004009](https://www.ncbi.nlm.nih.gov/nuccore/CP004009).

ACKNOWLEDGMENT

This work was supported by grant USDA-NIFA award 0826675.

REFERENCES

1. Barnes HJ, Nolan LK, Vaillancourt J-P. 2008. Colibacillosis, p

- 691–716. In Saif YM (ed), Diseases of poultry, 12th ed. Iowa State University Press, Ames, IA.
2. Gross WB. 1984. Colibacillosis, p 270–278. In Hofstad MS, Barnes HJ, Calnek BW, Reid WM, Yoder HW (ed), Diseases of poultry, 8th ed. Iowa State Press, Ames, IA.
 3. Gross WG. 1994. Diseases due to *Escherichia coli* in poultry, p 237–259. In Gyles CL (ed), *Escherichia coli* in domestic animals and humans. CAB International, Wallingford, United Kingdom.
 4. Gyles CL. 1994. *Escherichia coli* in domestic animals and humans. CAB International, Wallingford, United Kingdom.
 5. Gyles GL. 1993. *Escherichia coli*, p 164–187. In Gyles CL, Thoen CO (ed), Pathogenesis of bacterial infections in animals, 2nd ed. Iowa State University Press, Ames, IA.
 6. Gyles CL, Fairbrother JM. 2004. *Escherichia coli*, p 193–223. In Gyles CL, Prescott JF, Songer JG, Thoen CO (ed), Pathogenesis of bacterial infections in animals, 3rd ed. Iowa State University Press, Ames, IA.
 7. Morris M. 1989. Poultry health issue. Poultry Times July 3:11.
 8. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829. doi:[10.1101/gr.074492.107](https://doi.org/10.1101/gr.074492.107).
 9. Johnson TJ, Kariyawasam S, Wannemuehler Y, Mangiamele P, Johnson SJ, Doetkott C, Skyberg JA, Lynne AM, Nolan LK, Nolan LK. 2007. The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. *J. Bacteriol.* 189:3228–3236.
 10. Rojas TC, Parizzi LP, Tiba MR, Chen L, Pereira GA, Vartul S, Yang J, Yu J, Silveira WD. 2012. Draft genome of a Brazilian avian-pathogenic *Escherichia coli* Strain and in silico characterization of virulence-related genes. *J. Bacteriol.* 194:3023.
 11. Dziva F, Hauser H, Connor TR, van Diemen PM, Prescott G, Langridge GC, Eckert S, Chaudhuri RR, Ewers C, Mellata M, Mukhopadhyay S, Curtiss R, III, Dougan G, Wieler LH, Thomson NR, Pickard DJ, Stevens MP. 2013. Sequencing and functional annotation of avian pathogenic *Escherichia coli* serogroup O78 strains reveal the evolution of *E. coli* lineages pathogenic for poultry via distinct mechanisms. *Infect. Immun.* 81:838–849.