



Draft Genome Sequences of Enterohemorrhagic and Enteropathogenic *Escherichia coli* Strains Isolated from Alpacas in Peru

Lenin Maturrano,^a Marjorie Aleman,^a Dennis Carhuaricra,^a Jorge Maximiliano,^a Juan Siuce,^a Luis Luna,^a Raul Rosadio^a

^aLaboratory of Molecular Biology and Genetics, Faculty of Veterinary Medicine, Universidad Nacional Mayor de San Marcos, Lima, Peru

ABSTRACT The draft genome sequences of two strains of *Escherichia coli*, isolated from alpacas in Peru, are reported here. ECA1 has been determined to be a strain of enterohemorrhagic *E. coli* and ECB1 a strain of enteropathogenic *E. coli*. These pathogens are responsible for hemolytic-uremic syndrome in humans and diarrhea in different mammals, respectively.

In Peru, alpacas are the principal economic resource of the Andean communities and they suffer from diarrhea associated with pathogenic *Escherichia coli* strains, especially enterohemorrhagic *E. coli* (EHEC) and enteropathogenic *E. coli* (EPEC) strains (1, 2). The gastrointestinal tract of animals, mainly cattle, is considered the most important reservoir of EHEC (3); this pathogen is responsible for serious human diseases, such as hemolytic-uremic syndrome (HUS) (3), which expresses Shiga toxins. EPEC strains are implicated as a cause of diarrhea in humans and in several species of animals (4), and they induce the attaching and effacing (A/E) lesions in the intestinal mucosa, which are mediated by the *eae* gene encoding an outer membrane protein called intimin (5).

In this report, we present the draft genome sequence of EHEC strain ECA1 and EPEC strain ECB1 isolated from alpacas in Peru on a farm in which shedding had been observed (6). Using PCR, strain ECA1 was confirmed to carry one of the virulence genes of EHEC, *stx*, encoding Shiga toxin (responsible for HUS), and strain ECB1 was confirmed to carry one of the key genes of EPEC, *eae*. Additionally, the *in silico* serotypes were obtained using SerotypeFinder (7) and determined to be O26:H11 and O76:H2 for ECA1 and ECB1, respectively.

Genomic DNA was isolated from an overnight culture in Luria-Bertani medium using the GeneJET genomic DNA purification kit (Thermo Fisher Scientific). Whole-genome sequencing was performed at Macrogen (Seoul, South Korea). An Illumina HiSeq platform was used to sequence the genomes using a paired-end read length of 2 × 101 bp, and genome libraries were constructed using a TruSeq DNA PCR-free library preparation kit (Illumina, Inc.). The quality of the raw data was checked with FastQC (8), and adapters were trimmed with Trimmomatic (9). The genomes were assembled with Velvet version 1.2.10 (10) and annotated with the NCBI Prokaryotic Genome Annotation Pipeline (11).

The EHEC strain ECA1 and EPEC strain ECB1 draft genomes were determined to be 5,341,965 and 4,928,478 bp in length, distributed as 247 and 179 contigs, with G+C contents of 50.4% and 50.6%, respectively. The genome of ECA1 contains a total of 5,782 genes, including 5,464 coding genes, 6 rRNAs, and 90 tRNAs. The genome of ECB1 contains a total of 5,307 genes, including 4,911 coding genes, 5 rRNAs, and 79 tRNAs.

Accession number(s). These whole-genome shotgun projects have been deposited in DDBJ/ENA/GenBank under the accession numbers [PEQH00000000](https://www.ncbi.nlm.nih.gov/nuccore/PEQH00000000) (ECA1) and [PEQI00000000](https://www.ncbi.nlm.nih.gov/nuccore/PEQI00000000) (ECB1). The versions described here are the first versions.

Received 5 November 2017 Accepted 14 November 2017 Published 4 January 2018

Citation Maturrano L, Aleman M, Carhuaricra D, Maximiliano J, Siuce J, Luna L, Rosadio R. 2018. Draft genome sequences of enterohemorrhagic and enteropathogenic *Escherichia coli* strains isolated from alpacas in Peru. *Genome Announc* 6:e01391-17. <https://doi.org/10.1128/genomeA.01391-17>.

Copyright © 2018 Maturrano et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Lenin Maturrano, amaturrano@unmsm.edu.pe.

ACKNOWLEDGMENTS

This study was supported by “Generación Científica-Becas Nacionales-Fortalecimiento de Programas de Doctorado en Universidades Peruanas” grant no. 215-2014-FONDECYT and by the Consejo Nacional de Ciencia y Tecnología (CONCYTEC).

REFERENCES

1. Cid D, Martín-Espada C, Maturrano L, García A, Luna L, Rosadio R. 2011. Diarrheagenic *Escherichia coli* strains isolated from young Peruvian alpacas (*Vicugna pacos*) with diarrhea. In Perez-Cabal MA, Gutiérrez JP, Cervantes I, Alcalde MJ (ed), Fibre production in South American camelids and other fibre animals. Wageningen Academic Publishers, Wageningen, the Netherlands. https://doi.org/10.3920/978-90-8686-727-1_29.
2. Mori L, Perales R, Rodríguez J, Shiva C, Koga Y, Choquehuanca G, Palacios C. 2014. Molecular identification of Shiga-toxin producing and enteropathogenic *Escherichia coli* (STEC and EPEC) in diarrheic and healthy young alpacas. *Adv Microbiol* 4:360–364. <https://doi.org/10.4236/aim.2014.47043>.
3. Karmali MA, Gannon V, Sargeant JM. 2010. Verocytotoxin-producing *Escherichia coli* (VTEC). *Vet Microbiol* 140:360–370. <https://doi.org/10.1016/j.vetmic.2009.04.011>.
4. Ochoa TJ, Contreras CA. 2011. Enteropathogenic *Escherichia coli* infection in children. *Curr Opin Infect Dis* 24:478–483. <https://doi.org/10.1097/QCO.0b013e32834a8b8b>.
5. Hartland EL, Leong JM. 2013. Enteropathogenic and enterohemorrhagic *E. coli*: ecology, pathogenesis, and evolution. *Front Cell Infect Microbiol* 3:15. <https://doi.org/10.3389/fcimb.2013.00015>.
6. Rimac R. 2015. Detección de cepas patógenas de *Escherichia coli* en alpacas neonaetas y niños pastores. DVM thesis. Facultad de Medicina Veterinaria, Universidad Nacional Mayor de San Marcos, Lima, Peru.
7. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. 2015. Rapid and easy *in silico* serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. *J Clin Microbiol* 53:2410–2426. <https://doi.org/10.1128/JCM.00008-15>.
8. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
9. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
10. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
11. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of standard operating procedures (SOPs) for (meta)genomic annotation. *Omic* 12:137–141. <https://doi.org/10.1089/omi.2008.0017>.