

Draft Genome Sequences of the Enteroinvasive *Escherichia coli* Strains M4163 and 4608-58

Susan R. Leonard, David W. Lacher, Keith A. Lampel

Center for Food Safety and Applied Nutrition, Division of Molecular Biology, U.S. Food and Drug Administration, Laurel, Maryland, USA

We report here the draft genome sequences of enteroinvasive *Escherichia coli* (EIEC) O124:H30 strain M4163 isolated from imported French cheese and EIEC O143:H26 strain 4608-58. The assembled data determined that both strains contain multiple copies of the *ipaH* gene, as well as the pINV A form of the invasion plasmid.

Received 25 November 2014 Accepted 29 December 2014 Published 5 February 2015

Citation Leonard SR, Lacher DW, Lampel KA. 2015. Draft genome sequences of the enteroinvasive *Escherichia coli* strains M4163 and 4608-58. *Genome Announc* 3(1):e01395-14. doi:10.1128/genomeA.01395-14.

Copyright © 2015 Leonard 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 Keith A. Lampel, keith.lampel@fda.hhs.gov.

As an evolutionary intermediate between pathogenic *Escherichia coli* and *Shigella*, enteroinvasive *E. coli* (EIEC) invades intestinal epithelial cells but has not undergone the extensive gene decay observed in the *Shigella* genome (1–3). EIEC causes dysentery in infected individuals and has been associated with food-borne outbreaks (4–6). Here, we report the draft genome sequences of two EIEC isolates, M4163 and 4608-58. EIEC M4163 was isolated from imported Camembert cheese associated with a large outbreak in the United States in 1971 (4, 5), and EIEC 4608-58 was obtained from the Walter Reed National Military Medical Center collection (7, 8).

Genomic DNA was extracted from overnight cultures using the DNeasy blood and tissue kit (Qiagen). The sequencing libraries were prepared with the TruSeq DNA sample prep kit (Illumina) and sequenced on the Illumina MiSeq platform, generating 4,488,740 and 1,925,710 paired-end 250-bp reads for isolates M4163 and 4608-58, respectively. The draft genome sequences were assembled *de novo* with CLC Genomics Workbench version 7.0.4 (CLC bio), producing 294 contigs, resulting in a total genome size of 5,093,926 bp and an N_{50} of 60,982 bp for isolate M4163, and 266 contigs resulting in a total genome size of 5,043,882 bp and an N_{50} of 64,082 bp for isolate 4608-58. The draft genome sequences were annotated using RAST (9) and predicted to contain 5,203 and 4,977 coding sequences for EIEC M4163 and 4608-58, respectively.

Phylogenetic analysis utilizing backbone single-nucleotide polymorphisms (SNPs) from whole-genome sequences revealed that EIEC M4163 clusters with *E. coli* phylogroup A, while EIEC 4608-58 clusters with phylogroup E (10). Molecular serotyping using the *wzx*, *wzy*, and *fliC* loci indicates that M4163 and 4608-58 belong to serotypes O124:H30 and O143:H26, respectively. The genomes of both strains are approximately 500,000 bp larger than typical *Shigella* genomes. However, like *Shigella* genomes, both strains carry multiple copies of the *ipaH* gene. Sequence analysis of the invasion plasmid-associated genes *ipgD*, *mxiC*, and *mxiA* revealed that both EIEC strains possess the pINV A form of the plasmid (11). The 32-kb invasion region on these plasmids shows the greatest sequence similarity to the *Shigella boydii* type 18 strain

CDC 3083-94 (accession no. CP001062) for EIEC M4163 and to the *Shigella dysenteriae* type 2 strain 1012 (accession no. AAMJ00000000) for EIEC 4608-58. While *Shigella* isolates are lactose negative, EIEC isolates display variable ability to utilize lactose. EIEC 4608-58 harbors an intact *lac* operon, but the *lac* operon in the EIEC M4163 genome is truncated by an IS1 element. Consistent with this observation, plating on MacConkey agar demonstrated that EIEC 4608-58 and EIEC M4163 display positive and negative *lac* phenotypes, respectively.

The ability to further characterize these important pathogens will be enhanced with the addition of the sequences generated in this study, particularly since there are very few EIEC genome sequences that are publicly available.

Nucleotide sequence accession numbers. The draft genome sequences of *E. coli* M4163 and 4608-58 were deposited at DDBJ/EMBL/GenBank under the accession numbers JTCN00000000 and JTCO00000000, respectively. The versions described in this paper are JTCN01000000 and JTCO01000000.

ACKNOWLEDGMENTS

This research was funded in part by the Oak Ridge Institute for Science Education.

We thank Mark Mammel for his assistance with the backbone SNP analysis.

REFERENCES

- Peng J, Yang J, Jin Q. 2009. The molecular evolutionary history of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Infect Genet Evol* 9:147–152. <http://dx.doi.org/10.1016/j.meegid.2008.10.003>.
- Van den Beld MJC, Reubsat FAG. 2012. Differentiation between *Shigella*, enteroinvasive *Escherichia coli* (EIEC) and noninvasive *Escherichia coli*. *Eur J Clin Microbiol Infect Dis* 31:899–904. <http://dx.doi.org/10.1007/s10096-011-1395-7>.
- Parsot C. 2005. *Shigella* spp. and enteroinvasive *Escherichia coli* pathogenicity factors. *FEMS Microbiol Lett* 252:11–18. <http://dx.doi.org/10.1016/j.femsle.2005.08.046>.
- Tulloch EF, Jr, Ryan KJ, Formal SB, Franklin FA. 1973. Invasive enteropathic *Escherichia coli* dysentery. An outbreak in 28 adults. *Ann Intern Med* 79:13–17. <http://dx.doi.org/10.7326/0003-4819-79-1-13>.
- Marier R, Wells JG, Swanson RC, Callahan W, Mehlman IJ. 1973. An outbreak of enteropathogenic *Escherichia coli* foodborne disease traced to

- imported French cheese. *Lancet* 302:1376–1378. [http://dx.doi.org/10.1016/S0140-6736\(73\)93335-7](http://dx.doi.org/10.1016/S0140-6736(73)93335-7).
6. Gordillo ME, Reeve GR, Pappas J, Mathewson JJ, DuPont HL, Murray BE. 1992. Molecular characterization of strains of enteroinvasive *Escherichia coli* O143, including isolates from a large outbreak in Houston, Texas. *J Clin Microbiol* 30:889–893.
 7. Sansonetti PJ, d’Hauteville H, Formal SB, Toucas M. 1982. Plasmid-mediated invasiveness of *Shigella*-like *Escherichia coli*. *Ann Microbiol* 132: 351–355.
 8. DuPont HL, Formal SB, Hornick RB, Snyder MJ, Libonati JP, Sheahan DG, LaBrec EH, Kalas JP. 1971. Pathogenesis of *Escherichia coli* diarrhea. *N Engl J Med* 285:1–9. <http://dx.doi.org/10.1056/NEJM197107012850101>.
 9. 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>.
 10. Jaureguy F, Landraud L, Passet V, Diancourt L, Frapy E, Guigon G, Carbonnelle E, Lortholary O, Clermont O, Denamur E, Picard B, Nassis X, Brisse S. 2008. Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics* 9:560. <http://dx.doi.org/10.1186/1471-2164-9-560>.
 11. Lan R, Lumb B, Ryan D, Reeves PR. 2001. Molecular evolution of large virulence plasmid in *Shigella* clones and enteroinvasive *Escherichia coli*. *Infect Immun* 69:6303–6309. <http://dx.doi.org/10.1128/IAI.69.10.6303-6309.2001>.