




# Draft Genome Sequence of Enterohemorrhagic *Escherichia coli* O157:H7 Strain MC2 Isolated from Cattle in France

Pauline Auffret,<sup>a\*</sup> Audrey Segura,<sup>a</sup>  Yolande Bertin,<sup>a</sup> Christophe Klopp,<sup>b</sup> Olivier Bouchez,<sup>c</sup> Monique K  rour  dan,<sup>d</sup> Delphine Bibbal,<sup>d</sup> Hubert Brug  re,<sup>d</sup> Evelyne Forano<sup>a</sup>

Universit   Clermont Auvergne, INRA, MEDIS, Clermont-Ferrand, France<sup>a</sup>; Plateforme Bioinformatique Toulouse, Midi-Pyr  n  es UBIA, INRA, Auzeville Castanet-Tolosan, France<sup>b</sup>; GeT-PlaGe, Genotoul, INRA Auzeville, Castanet-Tolosan, France<sup>c</sup>; IRSD, Universit   de Toulouse, INSERM, INRA, ENVT, UPS, Toulouse, France<sup>d</sup>

**ABSTRACT** Enterohemorrhagic *Escherichia coli* (EHEC) with serotype O157:H7 is a major foodborne pathogen. Here, we report the draft genome sequence of EHEC O157:H7 strain MC2 isolated from cattle in France. The assembly contains 5,400,376 bp that encoded 5,914 predicted genes (5,805 protein-encoding genes and 109 RNA genes).

Enterohemorrhagic *Escherichia coli* (EHEC) with serotype O157:H7 is a major foodborne pathogen responsible for serious human diseases, such as hemolytic-uremic syndrome (HUS) (1). The gastrointestinal tract of cattle is the primary reservoir of EHEC O157:H7, which can be transmitted from cattle to humans by means of contaminated meat or unpasteurized milk (1). In this report, we present the draft genome sequence of EHEC O157:H7 strain MC2 isolated in France from cattle on a farm in which persistent shedding has been observed (2). Using real-time PCR (3, 4), MC2 was confirmed to belong to serotype O157:H7 and to carry key virulence genes of EHEC O157:H7: *eae*- $\gamma$ 1 encoding intimin (bacterial attachment to human epithelial cells) and *stx1* and *stx2* encoding Shiga toxins (responsible for HUS).

We sequenced *E. coli* MC2 because a better understanding of the nature of *E. coli* O157:H7 colonization of cattle is required to develop more effective preharvest food safety practices and to reduce the contamination of food of bovine origin.

*E. coli* MC2 was cultured on Luria-Bertani medium, and the genomic DNA was extracted using the DNeasy blood and tissue kit (Qiagen). Whole-genome sequencing was performed at the GeT-PlaGe core facility (INRA Toulouse, France [<http://get.genotoul.fr/>]). DNA-seq libraries were prepared according to Illumina's protocols (Illumina TruSeq Nano DNA LT library prep kit). DNA-seq experiments were performed on an Illumina MiSeq instrument using a paired-end read length of 2  $\times$  250 bp with the Illumina MiSeq reagent kit v2. The raw reads were stored in NG6 (5) and quality checked using FastQC (6). They were assembled with SPAdes v3.1.1 (7) using standard parameters and annotated with the NCBI Prokaryotic Genome Annotation Pipeline (8).

The resulting assembly consisted of 265 contigs. The MC2 genome comprised 5,400,376 bp with a G+C content of 50.34% and contained 5,805 protein-encoding genes (4,560 proteins with predicted functions and 4,144 assigned to cluster of orthologous groups (COG) of proteins using BLASTp) and 109 RNA genes. The contigs were aligned to the *E. coli* Sakai genome (NCBI Reference Sequence NC\_002695) and pO157 plasmid (NC\_002128) using Mauve (v2.3.1). The EHEC strain Sakai, isolated from a large-scale outbreak in Japan, is a standard reference for comparative genomic studies of O157:H7 *E. coli* (9). This analysis revealed that 94.19% of pO157 was covered by MC2 contigs (99.93% of homology). A total of 71 out of the 85 genes carried by

Received 4 September 2017 Accepted 6 September 2017 Published 5 October 2017

**Citation** Auffret P, Segura A, Bertin Y, Klopp C, Bouchez O, K  rour  dan M, Bibbal D, Brug  re H, Forano E. 2017. Draft genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 strain MC2 isolated from cattle in France. *Genome Announc* 5:e01097-17. <https://doi.org/10.1128/genomeA.01097-17>.

**Copyright**    2017 Auffret 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 Yolande Bertin, [yolande.bertin@inra.fr](mailto:yolande.bertin@inra.fr).

\* Present address: Pauline Auffret, Centre Iffremer du Pacifique, D  partement Ressources Biologiques et Environnement, Unit   Ressources Marines en Polyn  sie Fran  aise (RMPEF), Taravao, Tahiti, French Polynesia. P.A. and A.S. contributed equally to this work.

plasmid pO157 were present in the MC2 assembly (average homology, 99.74%), suggesting that MC2 carries at least one major plasmid homolog to pO157. Comparison of MC2 and Sakai genes using BLASTn (70% coverage and identity cutoff) demonstrated that 4,608 of the MC2 functional genes were present in *E. coli* O157:H7 Sakai, showing an average homology of 99.91%. Only 803 genes were unique to MC2. The high homology between the two genomes suggests that the bovine *E. coli* strain MC2 possesses the characteristic of an EHEC strain potentially responsible for clinical cases.

**Accession number(s).** This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number [NJDB00000000](https://www.ncbi.nlm.nih.gov/nuccore/NJDB00000000). The version described in this paper is version NJDB01000000.

## ACKNOWLEDGMENTS

This work was performed in collaboration with the GeT core facility, Toulouse, France, and was supported by France Génomique National infrastructure, funded as part of the “Investissement d’avenir” program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09) (the funders had no role in study design, data collection or interpretation or in submission of the work for publication). P.A. was supported by CASDAR (Compte d’Affectation Spéciale Développement Agricole et Rural; France Agrimer EDP 0914000451) and Interbev (Interprofession bétail et viande).

We are grateful to the Genotoul Bioinformatics Platform Toulouse Midi-Pyrénées (Genotoul Bioinfo) for providing computing and storage resources.

## REFERENCES

- 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>.
- Bibbal D, Loukiadis E, Kérouredan M, Ferré F, Dilasser F, Peytavin de Garam C, Cartier P, Oswald E, Gay E, Auvray F, Brugère H. 2015. Prevalence of carriage of Shiga toxin-producing *Escherichia coli* serotypes O157:H7, O26:H11, O103:H2, O111:H8, and O145:H28 among slaughtered adult cattle in France. *Appl Environ Microbiol* 81:1397–1405. <https://doi.org/10.1128/AEM.03315-14>.
- Auvray F, Lecureuil C, Dilasser F, Taché J, Derzelle S. 2009. Development of a real-time PCR assay with an internal amplification control for the screening of Shiga toxin-producing *Escherichia coli* in foods. *Lett Appl Microbiol* 48:554–559. <https://doi.org/10.1111/j.1472-765X.2009.02561.x>.
- Madic J, Peytavin de Garam C, Vingadassalon N, Oswald E, Fach P, Jamet E, Auvray F. 2010. Simplex and multiplex real-time PCR assays for the detection of flagellar (H-antigen) *fliC* alleles and intimin (*eae*) variants associated with enterohaemorrhagic *Escherichia coli* (EHEC) serotypes O26:H11, O103:H2, O111:H8, O145:H28 and O157:H7. *J Appl Microbiol* 109:1696–1705. <https://doi.org/10.1111/j.1365-2672.2010.04798.x>.
- Mariette J, Escudie F, Allias N, Salin G, Noiret C, Thomas S, Klopp C. 2012. NG6: integrated next generation sequencing storage and processing environment. *BMC Genomics* 13:462. <https://doi.org/10.1186/1471-2164-13-462>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- 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>.
- Hayashi T, Makino K, Ohnishi M, Kurokawa K, Ishii K, Yokoyama K, Han CG, Ohtsubo E, Nakayama K, Murata T, Tanaka M, Tobe T, Iida T, Takami H, Honda T, Sasakawa C, Ogasawara N, Yasunaga T, Kuhara S, Shiba T, Hattori M, Shinagawa H. 2001. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* 8:11–22. <https://doi.org/10.1093/dnares/8.1.11>.