

Genome Sequences of Two Bovine Mastitis-Causing *Escherichia coli* Strains

Florent Kempf,^{a,c} Valentin Loux,^d Pierre Germon^{b,c}

INRA, UMR 1282 Infectiologie et Santé Publique, CIRM-Bactéries Pathogènes, Nouzilly, France^a; INRA, UMR 1282 Infectiologie et Santé Publique, Immunité et Infections Mammaires, Nouzilly, France^b; Université François Rabelais de Tours, UMR 1282 Infectiologie et Santé Publique, Tours, France^c; INRA, UR 1077 MIG, Jouy-en-Josas, France^d

***Escherichia coli* is one of the main pathogenic agents causing inflammatory infections in the bovine udder. Here, we report the draft genome sequences of two strains isolated from different cases of clinical mastitis.**

Received 20 February 2015 Accepted 26 February 2015 Published 9 April 2015

Citation Kempf F, Loux V, Germon P. 2015. Genome sequences of two bovine mastitis-causing *Escherichia coli* strains. *Genome Announc* 3(2):e00259-15. doi:10.1128/genomeA.00259-15.

Copyright © 2015 Kempf 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/4.0/).

Address correspondence to Pierre Germon, pierre.germon@tours.inra.fr.

Bovine mastitis caused by *Escherichia coli* is a problem of major concern in the dairy industry. This well-known model bacterium has been described as either commensal or pathogenic, depending on the strain and environment considered. Nonetheless, except for some strains involved in recent human epidemic outbreaks, much remains unknown about the genetic background behind the phenotypic diversity of *E. coli* (1). In particular, mastitis strains are generally considered opportunistic pathogens. Several studies have focused on their possible genetic specificities that determine traits that are relevant for pathogenic potential (invasion capacity, virulence, and long-term persistence) (2), but so far, no virulence gene specific for a mastitis-causing *E. coli* isolate has been identified. Interestingly, clinical observations have suggested that mastitis strains might be clustered within a distinct pathotype (mammary pathogenic *E. coli* [MPEC] [3]).

E. coli model mastitis strain P4 was recently sequenced (4). In the present study, we considered two strains isolated from cases of clinical mastitis (D6-113.11 and D6-117.07), which may constitute a first corpus for comparative genomics studies. Further studies based on these strains may reveal bovine-associated genetic specificities possibly enabling a wider pathotype definition than that expected under the MPEC hypothesis.

High-quality genomic DNA was extracted using silica-based columns (NucleoSpin tissue; Macherey-Nagel, Hoerdt, France). For each strain, two sequencing runs were performed on an Illumina HiSeq 2000 genome analyzer (GATC, Konstanz, Germany) from a paired-end library and a 3-kb-insert mate-pair DNA library. Paired-end sequencing yielded between 62.7 and 19.6 million reads, totaling 3.196 and 1.031 Gb for D6-117.07 and D6-113.11, respectively. For each strain, the genome coverage was >220×, using the *E. coli* K-12 MG1655 genome size as a reference. Mate-pair library sequencing yielded between 11.3 and 18.9 million reads, totaling 0.057 and 0.963 Gb for D6-117.07 and D6-113.11, respectively (coverage range, 12 to 208×). The initial assembly and read scaffolding were performed using Velvet 1.2.07 (5) and were followed by automatic genome annotation on the AGMIAL platform (6).

The resulting draft genome information is as follows: D6-117.07 has 27 scaffolds and 4,795,662 bp, with a scaffold N_{50} of 2,494,862 bp, and D6-113.11 has 48 scaffolds totaling

5,097,170 bp, with a scaffold N_{50} of 715,015. Contigs/scaffolds <500 bp were considered potential artifacts and were removed from the study. Automatic annotation yielded 4,557 (D6-117.07) and 4,804 (D6-113.11) nonredundant coding sequences.

The strains are available upon request at the International Center for Microbial Resources—Bacterial Pathogens (CIRM-BP [http://www6.inra.fr/cirm_eng/Pathogenic-Bacteria]) under collection reference numbers and 2IM-260 (D6-117.07) and CIRM-BP-549 (D6-113.11).

Nucleotide sequence accession numbers. The draft genome sequences have been deposited at the European Nucleotide Archive (ENA) under study accession numbers [CCCP000000000](https://ena.ebi.ac.uk/ena/record/CCCP000000000) (D6-117.07) and [CCCO000000000](https://ena.ebi.ac.uk/ena/record/CCCO000000000) (D6-113.11).

ACKNOWLEDGMENT

This work was supported by the National Institute of Agronomic Research (INRA AIP “Bio-Resources”).

REFERENCES

1. Touchon M, Hoede C, Tenaillon O, Barbe V, Baeriswyl S, Bidet P, Bingen E, Bonacorsi S, Bouchier C, Bouvet O, Calteau A, Chiapello H, Clermont O, Cruveiller S, Danchin A, Diard M, Dossat C, Karoui ME, Frapy E, Garry L, Ghigo JM, Gilles AM, Johnson J, Le Bouguenec C, Lescat M, Mangenot S, Martinez-Jehanne V, Matic I, Nassif X, Oztas S, Petit MA, Pichon C, Rouy Z, Ruf CS, Schneider D, Tourret J, Vacherie B, Vallet D, Medigue C, Rocha EP, Denamur E. 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* 5:e1000344. <http://dx.doi.org/10.1371/journal.pgen.1000344>.
2. Blum SE, Leitner G. 2013. Genotyping and virulence factors assessment of bovine mastitis *Escherichia coli*. *Vet Microbiol* 163:305–312. <http://dx.doi.org/10.1016/j.vetmic.2012.12.037>.
3. Shpigel NY, Elazar S, Rosenshine I. 2008. Mammary pathogenic *Escherichia coli*. *Curr Opin Microbiol* 11:60–65. <http://dx.doi.org/10.1016/j.mib.2008.01.004>.
4. Blum S, Sela N, Heller ED, Sela S, Leitner G. 2012. Genome analysis of bovine-mastitis-associated *Escherichia coli* O32:H37 strain P4. *J Bacteriol* 194:3732. <http://dx.doi.org/10.1128/JB.00535-12>.
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. Bryson K, Loux V, Bossy R, Nicolas P, Chaillou S, van de Guchte M, Penaud S, Maguin E, Hoebeke M, Bessières P, Gibrat JF. 2006. AGMIAL: implementing an annotation strategy for prokaryote genomes as a distributed system. *Nucleic Acids Res* 34:3533–3545. <http://dx.doi.org/10.1093/nar/gkl471>.