genomeA
Journals.ASM.org

Complete Genome Sequences of *Escherichia coli* Strains 1303 and ECC-1470 Isolated from Bovine Mastitis

Andreas Leimbach,[a,b,c] Anja Poehlein,[b] Anika Witten,[d] Flemming Scheutz,[e] Ynte Schukken,[f,g] Rolf Daniel,[b] Ulrich Dobrindt[a,c]

Institute of Hygiene, University of Münster, Münster, Germany[a]; Department of Genomic and Applied Microbiology, Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August-University of Göttingen, Göttingen, Germany[b]; Institute for Molecular Infection Biology, Julius-Maximilians-University of Würzburg, Würzburg, Germany[c]; Institute for Human Genetics, University of Münster, Münster, Germany[d]; Statens Serum Institut, Copenhagen, Denmark[e]; Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York, USA[f]; GD Animal Health, Deventer, The Netherlands[g]

***Escherichia coli* is the leading causative agent of acute bovine mastitis. Here, we report the complete genome sequence of *E. coli* O70:H32 strain 1303, isolated from an acute case of bovine mastitis, and *E. coli* Ont:Hnt strain ECC-1470, isolated from a persistent infection.**

Received 9 February 2015 **Accepted** 13 February 2015 **Published** 26 March 2015

Citation Leimbach A, Poehlein A, Witten A, Scheutz F, Schukken Y, Daniel R, Dobrindt U. 2015. Complete genome sequences of *Escherichia coli* strains 1303 and ECC-1470 isolated from bovine mastitis. Genome Announc 3(2):e00182-15. doi:10.1128/genomeA.00182-15.

Address correspondence to Ulrich Dobrindt, dobrindt@uni-muenster.de.

The outcome and severity of *E. coli* intramammary infections were previously mainly associated with cow factors reacting to pathogen-associated molecular patterns rather than the genomic makeup of the infecting strain (1). Nevertheless, certain *E. coli* strains consistently cause an acute severe onset and others a mild chronic outcome (2, 3). Currently only the draft genome sequence of mastitis-associated *E. coli* O32:H37 strain P4 has been published (4).

E. coli 1303 was isolated from udder secretions of a cow with clinical mastitis (5) and *E. coli* ECC-1470 from a chronically infected cow (6). Both genomes were sequenced via whole-genome sequencing with the 454 FLX genome sequencer with GS20 chemistry (Roche Life Science, Mannheim, Germany) to a 27.8-fold or 13.4-fold coverage, respectively. Strain ECC-1470 was also sequenced with a 6-kb insert paired-end (PE) 454 sequencing library. Additionally, Nextera XT chemistry (Illumina, San Diego, CA, USA) for library preparation and a 101-bp PE sequencing run was used to sequence both strains on an Illumina HiScan SQ sequencer.

The 454 reads were *de novo* assembled with Newbler (v2.0.00.20 for 1303 and v2.3 for ECC-1470; Roche). The 454 and Illumina reads were *de novo* assembled using MIRA v3.4.0.1 (7). The hybrid assembly was combined with the initial Newbler assembly within the Gap4 software (v4.11.2) of the Staden package (8). Gaps were closed by primer walking via PCR and Sanger sequencing.

E. coli 1303 possesses a 4,948,797-bp and strain ECC-1470 a 4,803,751-bp chromosome. Each strain harbors an F-plasmid designated p1303_109 (108,501 bp) or pECC-1470_100 (100,061 bp), respectively. Additionally, strain 1303 contains a bacteriophage P1-like plasmid p1303_95 (94,959 bp) and a small cryptic plasmid p1303_5 (4,671 bp).

Annotation was done with Prokka v1.9 (9) and *E. coli* K-12 MG1655 (NC_000913.3) as reference. Annotations were manually curated by employing the Swiss-Prot, TrEMBL (10), IMG/ER (11), and Ecocyc databases (12). Open reading frame (ORF) finding was verified with YACOP v1 (13) and the reference strain MG1655's annotation using ACT v12.1.1 (14) for manual curation with Artemis v15.1.1 (15) and tbl2tab v0.1 (https://github.com/aleimba/bac-genomics-scripts/tree/master/tbl2tab). A total of 4,734 coding DNA sequences (CDS) were identified in *E. coli* 1303 with 22 rRNAs and 91 tRNAs (via tRNAscan-SE v1.3.1 [16]). The *E. coli* ECC-1470 genome includes 4,506 CDS with 22 rRNAs and 90 tRNAs.

By assigning multilocus sequence types (STs) using ecoli_mlst v0.3 (https://github.com/aleimba/bac-genomics-scripts/tree/master/ecoli_mlst) strains 1303 and ECC-1470 were allocated to phylogroups A (ST10) and B1 (ST847), respectively (17).

The most prominent virulence factors in both strains are the enterobactin siderophore, the group 4-capsule, and the *E. coli* type III secretion system 2. The genes *flu* (Ag43), *astA* (enteroaggregative *E. coli* heat-stable enterotoxin 1), *iss* (increased serum survival), an AMR-SSuT genomic island (antimicrobial resistance to streptomycin, sulfonamide, and tetracycline), and the second flagellar cluster, Flag-2, are only present in *E. coli* 1303. Putative virulence factors that are only present in strain ECC-1470 are two type VI secretion systems, the long polar fimbriae, Pix fimbriae, and the alternative flagellin Flk.

Nucleotide sequence accession numbers. The genome sequences have been deposited at DDBJ/ENA/GenBank under the accession numbers CP009166 to CP009169 (strain 1303) and CP010344 and CP010345 (strain ECC-1470).

ACKNOWLEDGMENTS

We thank I. Zude (Münster) for help with manual annotation and K. Gollnow (Göttingen) and B. Plaschke (Würzburg) for technical assistance.

This work was supported by the Deutsche Forschungsgemeinschaft (DO789/3-1 and DO789/4-1) and the Open Access Publishing Fund of the University of Münster.

REFERENCES

1. **Burvenich C, Van Merris V, Mehrzad J, Diez-Fraile A, Duchateau L.** 2003. Severity of *E. coli* mastitis is mainly determined by cow factors. Vet Res **34:**521–564. http://dx.doi.org/10.1051/vetres:2003023.
2. **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.
3. **Almeida RA, Dogan B, Klaessing S, Schukken YH, Oliver SP.** 2011. Intracellular fate of strains of *Escherichia coli* isolated from dairy cows with acute or chronic mastitis. Vet Res Commun **35:**89–101. http://dx.doi.org/10.1007/s11259-010-9455-5.
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. **Petzl W, Zerbe H, Günther J, Yang W, Seyfert HM, Nürnberg G, Schuberth HJ.** 2008. *Escherichia coli*, but not *Staphylococcus aureus* triggers an early increased expression of factors contributing to the innate immune defense in the udder of the cow. Vet Res **39:**18. http://dx.doi.org/10.1051/vetres:2007057.
6. **Dogan B, Klaessig S, Rishniw M, Almeida RA, Oliver SP, Simpson K, Schukken YH.** 2006. Adherent and invasive *Escherichia coli* are associated with persistent bovine mastitis. Vet Microbiol **116:**270–282. http://dx.doi.org/10.1016/j.vetmic.2006.04.023.
7. **Chevreux B.** 2005. MIRA: an automated genome and EST assembler. Ph.D. thesis. The Ruprecht-Karls-University, Heidelberg, Germany.
8. **Staden R, Beal KF, Bonfield JK.** 2000. The Staden package, 1998. Methods Mol Biol **132:**115–130.
9. **Seemann T.** 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30:**2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.
10. **UniProt Consortium.** 2014. Activities at the universal protein resource (UniProt). Nucleic Acids Res **42:**D191–D198. http://dx.doi.org/10.1093/nar/gkt1140.
11. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278. http://dx.doi.org/10.1093/bioinformatics/btp393.
12. **Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martínez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Schröder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus RP, Paulsen I, Karp PD.** 2013. EcoCyc: fusing model organism databases with systems biology. Nucleic Acids Res **41:**D605–D612. http://dx.doi.org/10.1093/nar/gks1027.
13. **Tech M, Merkl R.** 2003. YACOP: enhanced gene prediction obtained by a combination of existing methods. In Silico Biol **3:**441–451.
14. **Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J.** 2005. ACT: the Artemis comparison tool. Bioinformatics **21:**3422–3423. http://dx.doi.org/10.1093/bioinformatics/bti553.
15. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945. http://dx.doi.org/10.1093/bioinformatics/16.10.944.
16. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res **25:**955–964. http://dx.doi.org/10.1093/nar/25.5.0955.
17. **Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MC, Ochman H, Achtman M.** 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. Mol Microbiol **60:**1136–1151. http://dx.doi.org/10.1111/j.1365-2958.2006.05172.x.