



# Draft Genome Sequence of *Enterococcus faecalis* AS003, a Strain Possessing All Three Type II-a CRISPR Loci

Kevin B. Mulkerrins,<sup>a</sup> Casandra Lyons,<sup>a</sup>  Michael P. Shiaris<sup>a</sup>

<sup>a</sup>Shiaris Laboratory, Department of Biology, University of Massachusetts Boston, Boston, Massachusetts, USA

Kevin B. Mulkerrins and Casandra Lyons contributed equally to this work. Author order was determined in order of increasing seniority.

**ABSTRACT** *Enterococcus faecalis* is a clinically significant member of the human microbiome. Three CRISPR-Cas loci are located in conserved locations. Previous studies provide evidence that *E. faecalis* strains with functional CRISPR-Cas genes are negatively correlated with antibiotic resistance. Here, we report the genome sequence of an unusual strain possessing all three CRISPR-Cas loci.

*Enterococcus faecalis* is a widespread bacterial species found in environments ranging from soil to animal gut microbiomes (1, 2). An opportunistic pathogen, it is a common etiological agent of nosocomial infection (1, 3). It frequently undergoes horizontal gene transfer (HGT) associated with increased pathogenicity and acquired antibiotic resistance genes (ARGs) (4). Previous investigations of *E. faecalis* CRISPR loci have demonstrated a negative correlation between functional CRISPR loci and ARGs, presumably because functional CRISPR degrades foreign DNA as an immune response, thus impeding HGT (2, 5).

*E. faecalis* has three type II-A CRISPR-Cas loci in respectively conserved locations. From most to least common are CRISPR2, CRISPR1-Cas, and CRISPR3-Cas. CRISPR2 lacks *cas* genes and is assumed nonfunctional (6). The CRISPR type is determined by *cas* genes; however, CRISPR2 arrays show sequence similarity to CRISPR1 arrays (2, 6, 7). While enterococcal CRISPR loci can coincide, we have not observed other strains to possess all three; thus, AS003 was selected for sequencing.

Strain AS003 was cultured from an activated sludge sample at the Deer Island wastewater treatment plant in Boston, MA (8). Diluted activated sludge was filtered through 0.22- $\mu$ m membrane filters and then incubated at 35°C on *m-Enterococcus* agar (Difco). Colonies were streaked for isolation onto Enterococcosel agar (BBL) and grown at 35°C. Extraction cultures were grown in Trypticase soy broth without shaking. DNA extractions were performed using the MoBio UltraClean microbial DNA isolation kit. The variable region of the 16S rRNA gene was amplified using universal bacterial DNA primers (forward, 5'-CCTACGGGAGG CAGCAG-3'; reverse, 5'-ATTACCGCGGCTGCTGG-3') (9). Isolates were identified by 16S rRNA sequence matches in the Ribosomal Database Project (10). *Enterococcus faecalis* AS003 sequencing libraries were prepared from genomic DNA using the Nextera XT DNA library preparation kit and then indexed using the Nextera XT V2 index kit (Illumina, San Diego, CA). Libraries were purified and 350- to 500-bp fragments selected using the QIAquick gel extraction kit (Qiagen, Valencia, CA). Sequencing was performed on an Illumina MiSeq platform using a 500-cycle MiSeq V2 reagent kit.

A total of 1,670,410 250-bp reads were obtained for AS003. Genome assembly was performed using the A5-miseq pipeline, which trims and denoises the raw reads and then corrects the misassembled contigs to produce the final scaffolds (11). The assembly  $N_{50}$  score is 26,565 bp. The scaffolds were annotated by the National Microbial Pathogen Database Resource's Rapid Annotation using Subsystem Technology (RAST) service (12).

**Citation** Mulkerrins KB, Lyons C, Shiaris MP. 2021.

Draft genome sequence of *Enterococcus faecalis* AS003, a strain possessing all three type II-A CRISPR loci. Microbiol Resour Announc 10:e01449-20. <https://doi.org/10.1128/MRA.01449-20>.

**Editor** Julia A. Maresca, University of Delaware

**Copyright** © 2021 Mulkerrins 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 Michael P. Shiaris, michael.shiaris@umb.edu.

**Received** 18 December 2020

**Accepted** 26 February 2021

**Published** 18 March 2021

Assembly produced a 2,976,733-bp draft genome sequence consisting of 195 contigs. CRISPR1-Cas occurs on scaffold 115. CRISPR2 occurs on scaffold 25 and is confirmed by the cooccurrence of its flanking genes, cytochrome *d* ubiquinol oxidase subunit I (*cydA*), and transcriptional regulator, the AraC family. CRISPR3-Cas occurs on two scaffolds. The *cas* operon occurs on scaffold 124, and the array, confirmed by flanking gene *pstS*, occurs on scaffold 155. No mobile ARGs were detected in the genome sequence.

The presented draft assembly provides a rare example of an *E. faecalis* strain possessing all three canonical CRISPR loci. We had not observed the coincidence of CRISPR1 and CRISPR3 in *E. faecalis* prior to the isolation of this strain. The uncommon occurrence of CRISPR3 is expected to make this combination unlikely; however, a survey into the genome sequences available in public databases may reveal additional strains with both CRISPR1 and CRISPR3 or all three loci.

**Data availability.** This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number [JACJUJ000000000](https://www.ncbi.nlm.nih.gov/nuccore/JACJUJ000000000). The version described in this paper is the first version, [JACJUJ010000000](https://www.ncbi.nlm.nih.gov/nuccore/JACJUJ010000000). The BioProject and BioSample accession numbers are [PRJNA648426](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA648426) and [SAMN15641373](https://www.ncbi.nlm.nih.gov/biosample/SAMN15641373), respectively (13).

## ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## REFERENCES

1. Fisher K, Phillips C. 2009. The ecology, epidemiology and virulence of *Enterococcus*. *Microbiology* (Reading) 155:1749–1757. <https://doi.org/10.1099/mic.0.026385-0>.
2. Palmer KL, Gilmore MS. 2010. Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1:e00227-10. <https://doi.org/10.1128/mBio.00227-10>.
3. Rashid R, Cazenave-Gassiot A, Gao IH, Nair ZJ, Kumar JK, Gao L, Kline KA, Wenk MR. 2017. Comprehensive analysis of phospholipids and glycolipids in the opportunistic pathogen *Enterococcus faecalis*. *PLoS One* 12:e0175886. <https://doi.org/10.1371/journal.pone.0175886>.
4. Sanderson H, Ortega-Polo R, Zaheer R, Goji N, Amoako KK, Brown RS, Majury A, Liss SN, McAllister TA. 2020. Comparative genomics of multidrug-resistant *Enterococcus* spp. isolated from wastewater treatment plants. *BMC Microbiol* 20:20. <https://doi.org/10.1186/s12866-019-1683-4>.
5. Shehreen S, Chyou T-Y, Fineran PC, Brown CM. 2019. Genome-wide correlation analysis suggests different roles of CRISPR-Cas systems in the acquisition of antibiotic resistance genes in diverse species. *Philos Trans R Soc Lond B Biol Sci* 374:20180384. <https://doi.org/10.1098/rstb.2018.0384>.
6. Hullahalli K, Rodrigues M, Schmidt BD, Li X, Bhardwaj P, Palmer KL. 2015. Comparative analysis of the orphan CRISPR2 locus in 242 *Enterococcus faecalis* strains. *PLoS One* 10:e0138890. <https://doi.org/10.1371/journal.pone.0138890>.
7. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, Charpentier E, Cheng D, Haft DH, Horvath P, Moineau S, Mojica FJM, Scott D, Shah SA, Siksnys V, Terns MP, Venclovas Č, White MF, Yakunin AF, Yan W, Zhang F, Garrett RA, Backofen R, van der Oost J, Barrangou R, Koonin EV. 2020. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol* 18:67–83. <https://doi.org/10.1038/s41579-019-0299-x>.
8. Lyons C, Raustad N, Bustos MA, Shiaris M. 2015. Incidence of type II CRISPR1-Cas systems in *Enterococcus* is species-dependent. *PLoS One* 10:e0143544. <https://doi.org/10.1371/journal.pone.0143544>.
9. Muyzer G, de Waal EC, Uitterlinden AG. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* 59:695–700. <https://doi.org/10.1128/AEM.59.3.695-700.1993>.
10. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267. <https://doi.org/10.1128/AEM.00062-07>.
11. Coil D, Jospin G, Darling AE. 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* 31:587–589. <https://doi.org/10.1093/bioinformatics/btu661>.
12. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>.
13. Mulkerrins KB, Lyons CJ, Shiaris MP. 2020. Data from “Draft genome of *Enterococcus faecalis* strain AS003.” GenBank <https://www.ncbi.nlm.nih.gov/nuccore/JACJUJ000000000> (accession no. NZ\_JACJUJ000000000).