Complete Genome Sequence of *Enterococcus faecium* ATCC 700221

Peter T. McKenney,ᵃ Lilan Ling,ᵃ Guilin Wang,ᵇ Shrikant Mane,ᵇ Eric G. Pamerᵃ

Immunology Program, Infectious Diseases Service, Department of Medicine, Lucille Castori Center for Microbes, Inflammation, and Cancer, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York City, New York, USAᵃ; Yale Center for Genome Analysis, W. M. Keck Biotechnology Resource Laboratory, Yale University School of Medicine, Yale University, New Haven, Connecticut, USAᵇ

We report the complete genome sequence of a vancomycin-resistant isolate of *Enterococcus faecium* derived from human feces. The genome comprises one chromosome of 2.9 Mb and three plasmids. The strain harbors a plasmid-borne *vanA*-type vancomycin resistance locus and is a member of multilocus sequencing type (MLST) cluster ST-17.

Address correspondence to Peter T. McKenney, mckennep@mskcc.org, or Eric G. Pamer, pamere@mskcc.org.

*E*nterococcus faecium* ATCC 700221 is a *vanA*-positive fecal vancomycin-resistant enterococcus (VRE) isolate with high levels of resistance to several antibiotics, and it has been used in animal models of VRE intestinal colonization (1, 2). We produced a complete genome sequence of the strain in order to facilitate studies of genome dynamics and mutation accumulation during infection.

Colonies were grown overnight to exhaustion in brain heart infusion (BHI) liquid medium. DNA was extracted using phenol-chloroform extraction with bead beating and purified with AMPure XP beads. Libraries of 10 Kb were prepared by standard PacBio protocols (P/N 100-266-000-05) with BluePippin size selection. The samples were loaded with MagPure bead capture, sequencing was performed with P6-C4 chemistry, and data were captured for 4 h. For Illumina sequencing, purified DNA was sheared using a Covaris ultrasonicator and prepared for sequencing with a Kapa library preparation kit with Illumina TruSeq adapters to create 300 × 300-bp nonoverlapping paired-end reads (3).

PacBio reads from 2 single-molecule real-time (SMRT) cells were assembled in SMRT Analysis using HGAP_3 with default settings (4), yielding 4 contigs. They include a chromosome of 2.9 Mb and candidate plasmids of 189 kb, 64 kb, and 39 kb, respectively. The consensus was realigned to the PacBio reads using RS_resequencing_1 in SMRT Analysis, iteratively, for three rounds of self-correction by the Quiver algorithm, resulting in a consensus accuracy that was >99.999% for all contigs. PacBio data are known to be susceptible to single-base deletions in homopolymer runs during the calculation of circular consensus reads, and alignment of the Illumina reads to the PacBio consensus revealed 561 such deletions. We used Pilon to correct the PacBio consensus with the Illumina reads (5), which eliminated >90% of these deletions.

The genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (6). ATCC 700221 contains 2,725 predicted genes and belongs to multilocus sequencing type (MLST) cluster ST-17 (7, 8).

Plasmid-1 is 99% identical to pDO3 from *E. faecium* DO over 61% of its length (7) and contains a predicted hyl_{Efm} gene and *pilA–fms21-22* pilin gene cluster, which are thought to be virulence factors in mouse models (9–11). Plasmid-2 is 99% identical to the conjugative plasmid pMG1 over its entire length (12), and plasmid-3 is 99% identical to the Tn*1546*-containing VanA resistance plasmid pIS177 over 95% of its length (13). The VanA locus contains a predicted IS*1251* family insertion sequence between *vanRS* and *vanHAXYZ* (14).

PHAST predicts 6 complete prophages in ATCC 700221 (15). We note the presence of a large predicted prophage island between 1,624,385 bp and 1,729,248 bp. In this region, the average mapping quality values of PacBio reads aligned to the consensus is low, while the overall coverage is about 0.5 times greater (~450×) than that of the rest of the genome (~300×). This might indicate misassembly of a repetitive region. However, it is completely spanned by long reads with high mapping quality values, suggesting that the assembly may nonetheless be accurate in this region.

Nucleotide sequence accession numbers. The assembled sequences were submitted to GenBank under the accession numbers CP014449, CP014450, CP014451, and CP014452. The versions described in this announcement are the first versions.

FUNDING INFORMATION

REFERENCES

1. Patel R, Uhl JR, Kohner P, Hopkins MK, Cockerill FR. 1997. Multiplex PCR detection of *vanA*, *vanB*, *vanC*-1, and *vanC*-2/3 genes in enterococci. J Clin Microbiol **35**:703–707.
2. Ubeda C, Taur Y, Jenq RR, Equinda MJ, Son T, Samstein M, Viale A, Socci ND, van den Brink MRM, Kamboj M, Pamer EG. 2010. Vancomycin-resistant enterococcus domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. J Clin Invest **120**:4332–4341. http://dx.doi.org/10.1172/JCI43918.
3. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, No D, Liu H, Kinnebrew M, Viale A, Littmann E, van den Brink MRM, Jenq RR, Taur Y, Sander C, Cross JR, Toussaint NC, Xavier JB, Pamer EG. 2015. Precision microbiome reconstitution restores bile acid medi-

ated resistance to *Clostridium difficile*. Nature **517**:205–208. http://dx.doi.org/10.1038/nature13828.

4. **Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J.** 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods **10**:563–569. http://dx.doi.org/10.1038/nmeth.2474.

5. **Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM.** 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One **9**:e112963. http://dx.doi.org/10.1371/journal.pone.0112963.

6. **Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W.** 2013. Prokaryotic genome annotation pipeline. The NCBI handbook, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.

7. **Homan WL, Tribe D, Poznanski S, Li M, Hogg G, Spalburg E, Van Embden JDA, Willems RJL.** 2002. Multilocus sequence typing scheme for *Enterococcus faecium*. J Clin Microbiol **40**:1963–1971. http://dx.doi.org/10.1128/JCM.40.6.1963-1971.2002.

8. **Jolley KA, Maiden MC.** 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics **11**:595. http://dx.doi.org/10.1186/1471-2105-11-595.

9. **Arias CA, Panesso D, Singh KV, Rice LB, Murray BE.** 2009. Cotransfer of antibiotic resistance genes and a *hylEfm*-containing virulence plasmid in *Enterococcus faecium*. Antimicrob Agents Chemother **53**:4240–4246. http://dx.doi.org/10.1128/AAC.00242-09.

10. **Kim DS, Singh KV, Nallapareddy SR, Qin X, Panesso D, Arias CA, Murray BE.** 2010. The *fms21* (*pilA*)-*fms20* locus encoding one of four distinct pili of *Enterococcus faecium* is harboured on a large transferable plasmid associated with gut colonization and virulence. J Med Microbiol **59**:505–507. http://dx.doi.org/10.1099/jmm.0.016238-0.

11. **Rice LB, Laktičova V, Carias LL, Rudin S, Hutton R, Marshall SH.** 2009. Transferable capacity for gastrointestinal colonization in *Enterococcus faecium* in a mouse model. J Infect Dis **199**:342–349. http://dx.doi.org/10.1086/595986.

12. **Tanimoto K, Ike Y.** 2008. Complete nucleotide sequencing and analysis of the 65-kb highly conjugative *Enterococcus faecium* plasmid pMG1: identification of the transfer-related region and the minimum region required for replication. FEMS Microbiol Lett **288**:186–195. http://dx.doi.org/10.1111/j.1574-6968.2008.01342.x.

13. **Halvorsen EM, Williams JJ, Bhimani AJ, Billings EA, Hergenrother PJ.** 2011. Txe, an endoribonuclease of the enterococcal Axe-Txe toxin–antitoxin system, cleaves mRNA and inhibits protein synthesis. Microbiology **157**:387–397. http://dx.doi.org/10.1099/mic.0.045492-0.

14. **Handwerger S, Skoble J, Discotto LF, Pucci MJ.** 1995. Heterogeneity of the *vanA* gene cluster in clinical isolates of enterococci from the northeastern United States. Antimicrob Agents Chemother **39**:362–368. http://dx.doi.org/10.1128/AAC.39.2.362.

15. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS.** 2011. PHAST: a fast phage search tool. Nucleic Acids Res **39**:W347–W352. http://dx.doi.org/10.1093/nar/gkr485.