



Draft Genome Sequences of Six *Enterococcus faecalis* Strains Isolated from Malaysian Clinical and Environmental Origins

Diane Sunira Daniel,^{a,b} Han Ming Gan,^{a,b} Sui Mae Lee,^{a,b}  Gary A. Dykes,^c Sadequr Rahman^{a,b}

School of Science, Monash University, Bandar Sunway, Malaysia^a; Tropical Medicine and Biology Platform, Monash University, Bandar Sunway, Malaysia^b; School of Public Health, Curtin University, Perth, Western Australia, Australia^c

ABSTRACT *Enterococcus faecalis* is known to cause a variety of nosocomial infections, including urinary tract infections. Antibiotic resistance and virulence properties in this species are of public concern. The draft genome sequences of six *E. faecalis* strains isolated from clinical and environmental sources in Malaysia are presented here.

Enterococcus faecalis is an opportunistic pathogen that is often recovered from urinary tract infections (UTIs). *E. faecalis* is known to cause infections mainly due to the expression of virulence factors associated with adherence of mucosal and abiotic surfaces (1). The number of complete or draft genome sequences available for *E. faecalis* as of April 2017 is 503, comprising the bulk of enterococcal genome sequences available. However, there is a poor representation of genomic sequences for enterococci from Malaysia with only approximately seven assemblies reported (2). In this study, six *E. faecalis* strains, designated S12, S13, S14, S15, S16, and S17, were sequenced. These six strains were previously isolated from water sources, farm animal feces, and UTI patients in Malaysia and selected on the basis of different pulsed-field gel electrophoresis pulsotypes reported in a previously published experiment (3), and different biofilm and attachment properties (our unpublished data).

Genomic DNA was extracted using a GF-1 bacterial DNA extraction kit (Vivantis, Malaysia), tagged with Nextera XT (Illumina, USA) according to the manufacturer's instructions and sequenced on the MiSeq desktop sequencer located at the Monash University Malaysia Genomics Facility (2 × 250-bp run configuration). The raw reads generated were trimmed (quality score limit of 0.05) and assembled *de novo* using CLC Genomics Workbench version 7.0 (CLC bio, Denmark). Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline and the Rapid Annotations using Subsystems Technology server (4, 5). The identification of acquired antibiotic resistance and virulence genes was performed with the web tools ResFinder and VirulenceFinder, respectively (6, 7). Contigs coding for each antibiotic resistance and virulence gene (represented by contig accession number) were filtered based on 90% identity to the reference sequence and are presented in Table 1.

Genomic statistics—mean coverage, N_{50} contig length, number of contigs, assembly size, number of coding sequences, and number of tRNAs and rRNAs for the six assembled genomes are provided in Table 1. An orthologous average nucleotide identity tool based on OrthoANI values revealed a >98% similarity of all six strains to the whole genome of *E. faecalis* ATCC 19433 (PRJNA157741) (8).

Strain S14 harbors the complete gene cassette for vancomycin resistance, corroborating a previous wet lab report that recorded a MIC value of 64 μg/mL (3), whereas

Received 1 May 2017 Accepted 2 May 2017 Published 15 June 2017

Citation Daniel DS, Gan HM, Lee SM, Dykes GA, Rahman S. 2017. Draft genome sequences of six *Enterococcus faecalis* strains isolated from Malaysian clinical and environmental origins. *Genome Announc* 5:e00553-17. <https://doi.org/10.1128/genomeA.00553-17>.

Copyright © 2017 Daniel 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 Sadequr Rahman, sadequr.rahman@monash.edu.

TABLE 1 Summary characteristics of whole-genome assemblies of six *E. faecalis* strains isolated from various sources in Malaysia^a

Parameter	Characteristic for strain:					
	S12	S13	S14	S15	S16	S17
Accession no.	NBDR00000000	NBDS00000000	NBDW00000000	NBDV00000000	NBDU00000000	NBDT00000000
Isolation source	River water	Chicken feces	Chicken feces	River water	UTI patient	UTI patient
Mean coverage (×)	53	70	63	27	80	89
G+C content (%)	37.3	37.4	37.3	37.5	37.5	37.6
<i>N</i> ₅₀ (bp)	38,596	57,066	152,272	36,471	83,501	33,430
No. of contigs	140	133	92	158	110	229
Genome size (bp)	3,002,129	2,934,970	3,065,309	2,877,096	2,856,164	2,861,645
No. of CDSs	2,921	2,857	3,271	2,738	2,698	2,685
No. of tRNAs	59	42	55	37	51	49
No. of rRNAs	4	3	3	4	5	7
Contig accession no. for ^b :						
<i>vanBHRXWY</i>	NA	NA	NBDW01000006	NA	NA	NA
<i>tetM</i>	NBDR01000040	NA	NA	NA	NA	NA
<i>gelE</i>	NBDR01000048	NBDS01000053	NBDW01000028	NBDV01000016	NBDU01000023	NBDT01000088
<i>efaA</i>	NBDR01000100	NBDS01000066	NBDW01000002	NBDV01000033	NBDU01000049	NBDT01000053
<i>ace</i>	NBDR01000038	NBDS01000079	NBDW01000009	NBDV01000014	NBDU01000054	NBDT01000152
<i>ebpC</i>	NBDR01000050	NBDS01000024	NBDW01000009	NBDV01000060	NA	NBDT01000066
<i>ebpR</i>	NBDR01000050	NBDS01000024	NBDW01000009	NA	NBDU01000043	NBDT01000066
<i>eep</i>	NBDR01000032	NBDS01000004	NBDW01000006	NBDV01000036	NBDU01000015	NBDT01000113

^aAbbreviations: UTI, urinary tract infection; CDSs, coding sequences; NA, not available.

^bThe accession numbers of the reference sequences for each gene are as follows: *vanBHRXWY*, [KC489787](https://doi.org/10.1093/nar/gkw569); *tetM*, [U09422](https://doi.org/10.1093/nar/gkw569); *gelE*, [DQ845100](https://doi.org/10.1093/nar/gkw569); *efaA*, [JF512477](https://doi.org/10.1093/nar/gkw569); *ace*, [HQ003827](https://doi.org/10.1093/nar/gkw569); *ebpC*, [KJ710255](https://doi.org/10.1093/nar/gkw569); *ebpR*, [EF646762](https://doi.org/10.1093/nar/gkw569); and *eep*, [AF152237](https://doi.org/10.1093/nar/gkw569). Contigs containing the corresponding gene(s) exhibit ≥90% identity to their respective reference sequence.

strain S12 was found to have a tetracycline resistance gene (*tetM*) with a MIC value of 32 μg/mL (3). In addition, all six strains harbor genes coding for virulence factors—gelatinase production (*gelE*), endocarditis antigen (*efaA*), collagen adhesion (*ace*), and biofilm-associated pili (*eep*)—that are associated with adherence and invasion of the host tissue among enterococcal strains (9). Genes *ebpC* and *ebpR*, which are also biofilm-associated virulence factors, were identified in all strains except S16 and S15, respectively.

Accession number(s). The whole genome shotgun project of *E. faecalis* strains S12, S13, S14, S15, S16, and S17 have been deposited in DDBJ/ENA/GenBank under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

The project was supported by the School of Science, Monash University Malaysia. We are grateful to Yin Peng Lee from the Monash University Malaysia Tropical Medicine and Biology Multidisciplinary Platform for her assistance in next-generation sequencing.

REFERENCES

- Maryam D, Ozlem GE, M NA, Fatma NY, Evrim GA, Nefise A. 2014. The interactions between *esp*, *fsr*, *gelE* genes and biofilm formation and pfge analysis of clinical *Enterococcus faecium* strains. *Afr J Microbiol Res* 8:129–137. <https://doi.org/10.5897/AJMR2013.6257>.
- Daniel DS, Lee SM, Gan HM, Dykes GA, Rahman S. 2017. The public health risks of multiple-drug resistant (MDR) *Enterococcus* spp. in Southeast Asia. *Appl Environ Microbiol* 81:6090–6097. <https://doi.org/10.1128/AEM.01741-15>.
- Daniel DS, Lee SM, Gan HM, Dykes GA, Rahman S. 2017. Genetic diversity of *Enterococcus faecalis* isolated from environmental, animal and clinical sources in Malaysia. *J Infect Public Health* [Epub ahead of print.] <https://doi.org/10.1016/j.jiph.2017.02.006>.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodvovskiy M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640–2644. <https://doi.org/10.1093/jac/dks261>.
- Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J Clin Microbiol* 52:1501–1510. <https://doi.org/10.1128/JCM.03617-13>.
- Lee I, Kim YO, Park SC, Chun J. 2015. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* [Epub ahead of print.] <https://doi.org/10.1099/ijsem.0.000760>.
- Anderson AC, Jonas D, Huber I, Karygianni L, Wölber J, Hellwig E, Arweiler N, Vach K, Wittmer A, Al-Ahmad A. 2015. *Enterococcus faecalis* from food, clinical specimens and oral sites: prevalence of virulence factors in association with biofilm formation. *Front Microbiol* 6:1534. <https://doi.org/10.3389/fmicb.2015.01534>.