JAST (Journal of Animal Science and Technology) TITLE PAGE

Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research Article
Article Title (within 20 words without abbreviations)	Complete genome sequence of *Enterococcus faecium* strain AK_C_05 with potential characteristics applicable in livestock industry
Running Title (within 10 words)	Complete genome sequence of *Enterococcus faecium* strain AK_C_05
Author	Hyunok Doo[1#], Jin Ho Cho[2#], Minho Song[3#], Eun Sol Kim[1], Sheena Kim[1], Gi Beom Keum[1], Jinok Kwak[1], Sriniwas Pandey[1], Sumin Ryu[1], Yejin Choi[1], Juyoun Kang[1], Hyeun Bum Kim[1*], and Ju-Hoon Lee[4*]
Affiliation	1 Department of Animal Resources Science, Dankook University, Cheonan 31116, Korea 2 Division of Food and Animal Science, Chungbuk National University, Cheongju 28644, Korea 3 Division of Animal and Dairy Science, Chungnam National University, Daejeon 34134, Korea 4 Department of Food Animal Biotechnology, Department of Agricultural Biotechnology, Center for Food and Bioconvergence, Seoul National University, Seoul 08826, Korea
ORCID (for more information, please visit https://orcid.org)	Hyunok Doo: https://orcid.org/0000-0003-4329-4128 Jin Ho Cho: https://orcid.org/0000-0001-7151-0778 Minho Song: https://orcid.org/0000-0002-4515-5212 Eun Sol Kim: https://orcid.org/0000-0001-8801-421X Sheena Kim: https://orcid.org/0000-0002-5410-1347 Gi Beom Keum: https://orcid.org/0000-0001-6006-9577 Jinok Kwak: https://orcid.org/0000-0003-1217-3569 Sriniwas Pandey: https://orcid.org/0000-0002-6947-3469 Sumin Ryu: https://orcid.org/0000-0002-1569-3394 Yejin Choi: https://orcid.org/0000-0002-7434-299X Juyoun Kang: https://orcid.org/0000-0002-3974-2832 Hyeun Bum Kim: https://orcid.org/0000-0003-1366-6090 Ju-Hoon Lee: https://orcid.org/0000-0003-0405-7621
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This research was supported by a grant (22193MFDS538) from Ministry of Food and Drug Safety in 2022
Acknowledgements	Not applicable.
Availability of data and material	The complete genome sequences of *Enterococcus faecium* strain AK_C_05 were deposited in GeneBank under the accession numbers CP128994.1 and CP128995.1. The BioSample accession number is SAMN35654454, and BioProject accession number is PRJNA980926
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Doo H, Kim HB, Lee JH Data curation: Keum GB, Choi Y, Kang J Formal analysis: Kim ES, Kim S, Keum GB, Ryu S Methodology: Cho JH, Song M Validation: Kim S, Kwak J, Pandey S Writing - original draft: Doo H, Cho JH, Song M Writing - review & editing: Doo H, Cho JH, Song M, Kim ES, Kim S, Keum GB, Kwak J, Pandey S, Ryu S, Choi Y, Kang J, Kim HB, Lee JH
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Hyeun Bum Kim
Email address – this is where your proofs will be sent	hbkim@dankook.ac.kr
Secondary Email address	
Address	Department of Animal Resources Science, Dankook University, Cheonan 31116, Korea
Cell phone number	+82-10-3724-3416
Office phone number	+82-41-550-3653
Fax number	+82-41-565-2940

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Ju-Hoon Lee
Email address – this is where your proofs will be sent	juhlee@snu.ac.kr
Secondary Email address	
Address	Department of Food Animal Biotechnology, Department of Agricultural Biotechnology, Center for Food and Bioconvergence, Seoul National University, Seoul 08826, Korea
Cell phone number	+82-10- 9678-5529
Office phone number	+82-2-880-4854
Fax number	+82-2-875-5095

Complete genome sequence of *Enterococcus faecium* strain AK_C_05 with potential characteristics applicable in livestock industry

Hyunok Doo[1#], Jin Ho Cho[2#], Minho Song[3#], Eun Sol Kim[1], Sheena Kim[1], Gi Beom Keum[1], Jinok Kwak[1], Sriniwas Pandey[1], Sumin Ryu[1], Yejin Choi[1], Juyoun Kang[1], Hyeun Bum Kim[1*] and Ju-Hoon Lee[4*]

[1] Department of Animal Resources Science, Dankook University, Cheonan 31116, Korea

[2] Division of Food and Animal Science, Chungbuk National University, Cheongju 28644, Korea

[3] Division of Animal and Dairy Science, Chungnam National University, Daejeon 34134, Korea

[4] Department of Food Animal Biotechnology, Department of Agricultural Biotechnology, Center for Food and Bioconvergence, Seoul National University, Seoul 08826, Korea

# Equal contributors

* Corresponding authors

Hyeun Bum Kim

Department of Animal Resources Science, Dankook University, Cheonan 31116, Korea

Tel: +82-41-550-3653

Email: hbkim@dankook.ac.kr

Ju-Hoon Lee

Department of agricultural biotechnology. Seoul National University, Seoul 08826, Korea

Tel: +82-2-880-4854

Email: juhlee@snu.ac.kr

(Unstructured) Abstract (up to 350 words)

The *Enterococcus faecium* (*E. faecium*) strain AK_C_05 was isolated from cheonggukjang, the Korean traditional food, collected from a local market in South Korea. In this report, we presented the complete genome sequence of *E. faecium* strain AK_C_05. The genome of *E. faecium* strain AK_C_05 genome consisted of one circular chromosome (2,691,319 bp) with a guanine + cytosine (GC) content of 38.3% and one circular plasmid (177,732 bp) with a guanine + cytosine (GC) content of 35.48%. The Annotation results revealed 2,827 protein-coding sequences (CDSs), 18 rRNAs, and 68 tRNA genes. It possesses genes, which encodes enzymes such as alpha-galactosidase (EC 3.2.1.22), beta-glucosidase (EC 3.2.1.21) and alpha-L-arabinofuranosidase (EC 3.2.1.55) enabling efficient utilization of carbohydrates. Based on Clusters of Orthologous Groups analysis, *E. faecium* strain AK_C_05 showed specialization in carbohydrate transport and metabolism indicating the ability to generate energy using a variety of carbohydrates.

Keywords (3 to 6):

Enterococcus faecium, Livestock, Carbohydrates

The main text

The Enterococci bacteria belong to lactic acid bacteria (LAB) group, which can be found in fermented foods[1]. Especially, *Enterococcus faecium* is also utilized as probiotics, which could enhance the microbial balance in animals[2]. Despite of safety concerns regarding its use as probiotics, recent research has explored the use of *Enterococcus faecium* as a feed additive for livestock to enhance growth performance[1, 3].

In the present study, the *E. faecium* strain AK_C_05 was isolated from homemade cheonggukjang, the Korean traditional food, collected from a local market in Cheonan (36.802917° N, 127.149796° E), Chungcheongnam-do, South Korea. Then, the whole genome sequencing was performed to understand the genomic characteristics of *E. faecium* strain AK_C_05 as a potential probiotic in the livestock industry. The *E. faecium* strain AK_C_05 was cultivated in Enterococcosel broth (MBcell, Seoul, South Korea) at 37℃ for 24 hours. Genomic DNA was extracted from the cultured *E. faecium* pellet using CTAB DNA extraction method. The complete genome of the *E. faecium* AK_C_05 was sequenced using the Oxford Nanopore Technologies MinION platform at eGnome (Seoul, South Korea). Briefly, library preparation was performed using Native barcoding Sequencing Kit (SQK_NBD114.24, V14) following the manufacturer's instructions (Oxford Nanopore Technologies, Oxford, UK). The prepared library was loaded into the MinION MK1b sequencing device (Oxford Nanopore) equipped with a MinION flow cell (MIN114, R10.4.1, Oxford Nanopore). The Oxford Nanopore sequencing produced 79,247 of long reads, resulting in a total of 572,297,864 base pairs. De novo assemble was performed using a Flye assembler v2.9.2, followed by polishing using the Homopolish polisher v0.4.1. The quality of genome assembly was assessed using Quality Assessment Tool for Genome Assemblies (QUAST) v5.2.0[4]. The quantitative assessment of the genome completeness was conducted using the Benchmarking Universal Single-Copy Orthologs (BUSCO) v5.4.6[5]. To annotate and predict the protein coding genes, rRNA, and tRNA genes of *E. faecium strain* AK_C_05, the Rapid Annotation using Subsystem Technology (RAST) v2.0 tool was utilized[6]. The functional categorization of all predicted protein coding genes was performed using the Clusters of Orthologous Groups (COGs)-based EggNOG-mapper v2.0[7]. Furthermore, the presence of virulence factors and antibiotic resistance in *E. faecium* strain AK_C_05 was predicted using the BLASTn method, with reference to the Virulence Factor Database (VFDB) and the Comprehensive Antibiotic Resistance Database (CARD)[8, 9].

The complete genome of the *E. faecium* strain AK_C_05 contain one circular chromosome (2,691,319 bp) with a guanine + cytosine (GC) content of 38.3% and one circular plasmid (177,732 bp) with a guanine + cytosine (GC) content of 35.48%. A total of 2,827 predicted protein-coding sequence, 18 rRNA genes, and 68 tRNA genes

were identified in *E. faecium* strain AK_C_05. The most abundant COGs category, excluding Function unknown [S], was Carbohydrate transport and metabolism [G], which accounted for 235 genes, representing 10.4% of the total genes identified. The genome feature and map of *E. faecium* strain AK_C_05 were presented in Table 1, Figure 1A and 1B.

Based on its specific focus on carbohydrate transport and metabolism, *E. faecium* strain AK_C_05 possesses genes and enzymes, such as alpha-galactosidase (EC 3.2.1.22), beta-glucosidase (EC 3.2.1.21) and alpha-L-arabinofuranosidase (EC 3.2.1.55), that enable efficient utilization of carbohydrates and the capacity to derive energy from diverse carbohydrate substrates. This characteristic makes *E. faecium* strain AK_C_05 a potential candidate for application in the livestock industry. The complete genome of *E. faecium* strain AK_C_05 has indicated the presence of the antibiotic resistance gene aac (6')-Ii in the chromosome and not in the plasmid, confirming that there is no potential for transmission of the resistance gene to other microorganisms. In the plasmid of *E. faecium* strain AK_C_05, the filA gene was detected, while no other virulence factors were identified. Interestingly, the filA gene's ability to facilitate adhesion to the cell wall is regarded as a beneficial trait for probiotics[10]. Overall, our results indicate that *E. faecium* AK_C_05 could be a promising functional probiotic for improving growth performance in the livestock industry.

References (Vancouver or NLM style)

1. Franz CM, Stiles ME, Schleifer KH, Holzapfel WH. Enterococci in foods--a conundrum for food safety. Int J Food Microbiol. 2003;88(2-3):105-22;https://doi.org/10.1016/s0168-1605(03)00174-0

2. Bhardwaj A, Gupta H, Kapila S, Kaur G, Vij S, Malik RK. Safety assessment and evaluation of probiotic potential of bacteriocinogenic Enterococcus faecium KH 24 strain under in vitro and in vivo conditions. Int J Food Microbiol. 2010;141(3):156-64;https://doi.org/10.1016/j.ijfoodmicro.2010.05.001

3. Sato Y, Kuroki Y, Oka K, Takahashi M, Rao S, Sukegawa S, et al. Effects of Dietary Supplementation With Enterococcus faecium and Clostridium butyricum, Either Alone or in Combination, on Growth and Fecal Microbiota Composition of Post-weaning Pigs at a Commercial Farm. Front Vet Sci. 2019;6:26;https://doi.org/10.3389/fvets.2019.00026

4. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072-5;https://doi.org/10.1093/bioinformatics/btt086

5. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210-2;https://doi.org/10.1093/bioinformatics/btv351

6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:75;https://doi.org/10.1186/1471-2164-9-75

7. Cantalapiedra CP, Hernandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. Mol Biol Evol. 2021;38(12):5825-9;https://doi.org/10.1093/molbev/msab293

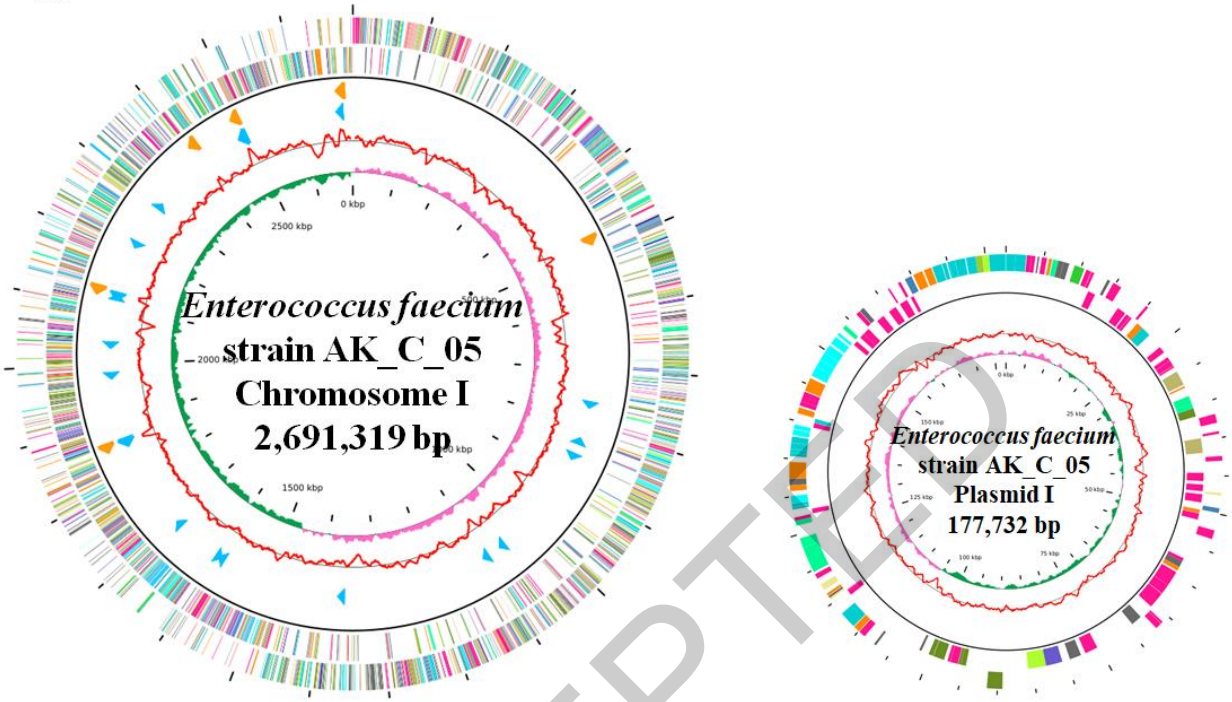
8. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 2005;33(Database issue):D325-8;https://doi.org/10.1093/nar/gki008

9. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother. 2013;57(7):3348-57;https://doi.org/10.1128/AAC.00419-13

10. Li B, Zhan M, Evivie SE, Jin D, Zhao L, Chowdhury S, et al. Evaluating the Safety of Potential Probiotic Enterococcus durans KLDS6.0930 Using Whole Genome Sequencing and Oral Toxicity Study. Front Microbiol. 2018;9:1943;https://doi.org/10.3389/fmicb.2018.01943

Tables and Figures

A



Type	Contig size (bp)	GC-content (%)	Protein-coding genes	tRNA	rRNA
Chromosome	2,691,319	38.3	2,629	68	18
Plasmid	177,732	35.48	198	-	-

B

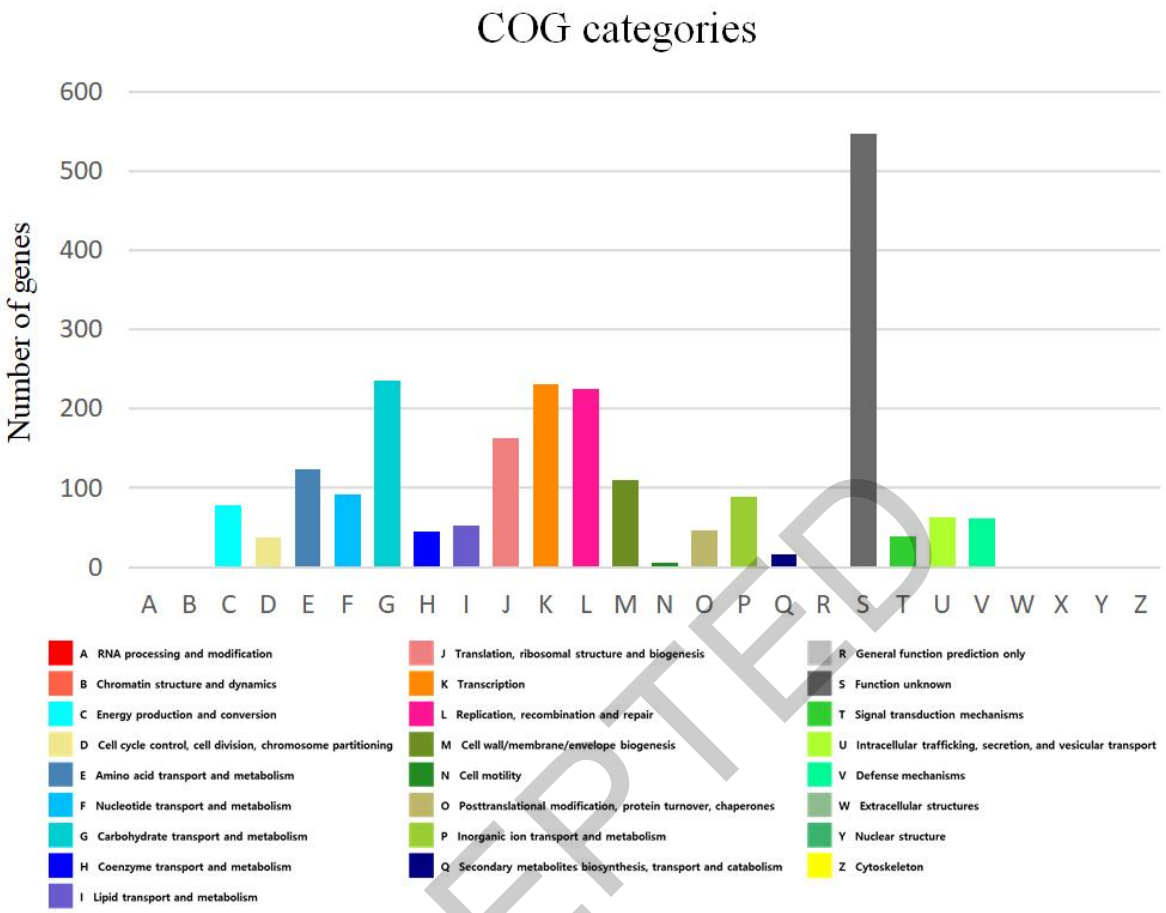


Figure 1. Genome map of *Enterococcus faecium* strain AK_C_05 and the functional categorization of predicted protein coding genes. The outer ring represents the positions of all annotated gene coding regions (ORFs), while the inner ring in red indicates the guanine + cytosine (GC) content. Peaks in pink and green indicate GC skew. The orange and sky-blue arrows represent rRNA and tRNA operons, respectively. The annotated ORFs are color-coded based on their Clusters of Orthologous Groups (COG) assignments in Figure 1A. The COG functional categories of the predicted protein coding genes are represented in Figure 1B.

136

137 **Table 1. Genome features of *Enterococcus faecium* strain AK_C_05**

Property	Term	
	Chromosome	**Plasmid**
Contig length (bp)	2,691,319 bp	177,732 bp
No. of contig	1 (chromosome)	1 (plasmid)
Guanine + cytosine (G + C)	38.3	35.48
Protein-coding genes	2,629	198
rRNA genes	18	-
tRNA genes	68	-
Genbank Accession No.	CP128995	CP128994

138