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	Genome Announcement
Article Title (within 20 words without abbreviations)	Complete genome sequence of *Enterococcus faecium* strain GB_C_05 with potential characteristics applicable as a bacteriocin-producing probiotic feed additive
Running Title (within 10 words)	Complete genome sequence of *Enterococcus faecium* strain GB_C_05
Author	Juyoun Kang [1#], Hyunok Doo[1#], Jinok Kwak[1#], Eun Sol Kim[1, 2], Gi Beom Keum[1], Yejin Choi[1], Haram Kim[1], Yeongjae Chae[1], Sheena Kim[1], Hyeun Bum Kim[1*], and Ju-Hoon Lee[3*]
Affiliation	[1] Department of Animal Biotechnology, Dankook University, Cheonan 31116, Korea [2] Division of Infectious Diseases, Department of Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill 27599, USA [3] 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)	Juyoun Kang (https://orcid.org/0000-0002-3974-2832) Hyunok Doo (https://orcid.org/0000-0003-4329-4128) Jinok Kwak (https://orcid.org/0000-0003-1217-3569) Eun Sol Kim (https://orcid.org/0000-0001-8801-421X) Gi Beom Keum (https://orcid.org/0000-0001-6006-9577) Yejin Choi (https://orcid.org/0000-0002-7434-299X) Haram Kim (https://orcid.org/0009-0002-7504-5249) Yeongjae Chae (https://orcid.org/0009-0004-5573-1465)

ACCEPTED

	Sheena Kim (https://orcid.org/0000-0002-5410-1347) 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 2025
Acknowledgements	Not applicable.
Availability of data and material	The complete genome sequences of *Enterococcus faecium* strain GB_C_05 was deposited in GenBank under the accession numbers CP142862.1 and CP142861.1. The BioSample accession number is SAMN39489522, and BioProject accession number is PRJNA1066497
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Kang J, Doo H, Kim HB, Lee JH Data curation: Kim ES, Choi Y, Kim S Formal analysis: Kang J, Doo H, Kim ES, Choi Y, Chae Y Methodology: Kwak J, Keum GB Validation: Kwak J, Keum GB, Kim H, Kim S Writing - original draft: Kang J, Doo H, Kwak J Writing - review & editing: Kang J, Doo H, Kim S, Kim ES, Keum GB, Kwak J, Choi Y, Kim H, Chae Y, Kim HB, and Lee JH
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

ACCEPTED

4

5 **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 Biotechnology, 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

6

7 **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

8

9

ACCEPTED

Complete genome sequence of *Enterococcus faecium* strain GB_C_05 with potential characteristics applicable as a bacteriocin-producing probiotic feed additive.

Juyoun Kang [1#], Hyunok Doo[1#], Jinok Kwak[1#], Eun Sol Kim[1, 2], Gi Beom Keum[1], Yejin Choi[1], Haram Kim[1], Yeongjae Chae[1], Sheena Kim[1], Hyeun Bum Kim[1*], and Ju-Hoon Lee[3*]

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

[2] Division of Infectious Diseases, Department of Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill 27599, USA

[3] 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

ACCEPTED

Abstract (up to 150 words)

The whole genome of *Enterococcus faecium* GB_C_05, a strain isolated from Sikhye, a traditional Korean rice beverage, was successfully sequenced and analyzed using Oxford Nanopore Technologies. The complete genome sequence of GB_C_05 contains a circular chromosome with a total length of 2,575,440 base pairs (bp) and a guanine + cytosine (GC) content of 38.2%, along with one circular plasmid, which has a length of 230,283 bp and a GC content of 35.2%. Annotation of the GB_C_05 genome revealed 2,756 protein-coding sequences (CDSs), 70 tRNAs, and 18 rRNAs on the chromosome. Notably, CDSs related to bacteriocin synthesis (Enterocin) and carbohydrate metabolism (α-galactosidase, β-glucosidase and α-L-arabinofuranosidase), as well as genes potentially involved in probiotic-associated functions such as adhesion and colonization, were identified. This comprehensive study presents the complete genome sequence of *Enterococcus faecium* GB_C_05, providing insight into the diverse additives utilized in animal farming to enhance nutritional quality and livestock productivity.

Keywords (3 to 6):

Whole genome sequencing, *Enterococcus faecium*, Probiotics, Feed additive

ACCEPTED

The main text

The genus *Enterococcus* ranks as the third largest lactic acid bacteria (LAB) group, following *Lactobacillus* and *Streptococcus* [1]. Certain strains of *Enterococcus faecium*(*E. faecium*), when used as probiotics, have been shown to contribute to immunomodulation within the intestinal mucosa and to aid in the development of the digestive system. It is widely employed in the livestock industry as a substitute feed additive to enhance animal growth, particularly in pig and poultry farming [1-3]. *E. faecium* strain GB_C_05 was isolated from Sikhye, a traditional Korean rice beverage, obtained from a local market in Cheonan, South Korea. *E. faecium* GB_C_05 was cultured in Enterococcosel (MBcell, Seoul, South Korea) broth at 37°C for 24 hours. The genomic DNA of *Enterococcus faecium* GB_C_05 was extracted from the cell pellet obtained from a 24-hour culture using the G-spin™ Genomic DNA Extraction Kit (for Bacteria) (Invitrogen, Waltham, MA, USA). The concentration of the extracted DNA was determined using the Qubit™ dsDNA HS Assay Kit (Invitrogen). Libraries were constructed using the Ligation Sequencing Kit V14 (Oxford Nanopore Technologies, Oxford, UK) according to the manufacturer's instructions. The purified library was loaded into a MinION flow cell (R10.4.1) (Oxford Nanopore Technologies, Oxford, UK) and sequenced for 22 hours using a MinION sequencer (Oxford Nanopore Technologies, Oxford, UK). Oxford Nanopore sequencing produced 128,994 long reads, for a total of 375,852,265 base pairs. The extracted raw data was demultiplexed, and the adapters were trimmed using Porechop (version 0.2.4), followed by read quality adjustment using Chopper (version 0.7.0) [4]. Assembly was performed using Canu (version 1.8) and Flye (version 2.9.2) tools, and errors occurring in nanopore sequencing data were identified and corrected through Homopolish polisher (version 0.4.1) [4, 5]. Evaluation of the assembled genome was conducted using Quality Assessment Tool for Genome Assemblies (QUAST) (version 5.0.2) and Benchmarking Universal Single-Copy Orthologs (BUSCO) (version 5.4.6) [6, 7]. The web-based annotation tools RAST (version 2.0) and EggNOG-mapper (version 2.0) were used to analyze the data and identify key genes and metabolic pathways [8, 9]. Virulence and antibiotic resistance genes were identified using Virulence Factor Database (VFDB) and ResFinder (version 4.4.0) [10, 11]. Bacteriocin genes were explored using the Bagel 4 web software [10].

The chromosome of *E. faecium* strain GB_C_05 comprises 2,575,440 bp with a GC content of 38.2%, and contains 2,756 predicted protein-coding sequences, along with 18 rRNA genes and 70 tRNA genes. In addition, a circular plasmid, 230,283 bp in length and with a GC content of 35.2%, was identified separately from the chromosome. Additionally, the plasmid contained 391 CDSs, with no tRNA or rRNA genes identified. The most abundant COG categories, excluding ‘Unknown function [S]’, were ‘Carbohydrate transport and metabolism [G]’ (254 genes,

ACCEPTED

10.28%) and ‘Replication, recombination, and repair [L]’ (254 genes, 10.28%), comprising a total of 20.56%. This was followed by ‘Transcription [K]’ (250 genes, 10.11%). The genome map and COG functional classification of *E. faecium* GB_C_05 are shown in Fig. 1A and 1B. Genes encoding enzymes essential for carbohydrate transport and metabolism, such as α-galactosidase (EC 3.2.1.22), β-glucosidase (EC 3.2.1.21), and α-L-arabinofuranosidase (EC 3.2.1.55), were identified. This genetic composition suggests the potential for efficient carbohydrate utilization and energy extraction from various carbohydrate substrates. Genes related to carbohydrate metabolism may help improve feed digestibility and enhance livestock productivity, providing a crucial competitive advantage in the livestock industry [3]. The genome of *E. faecium* strain GB_C_05 contains bacteriocin gene clusters encoding bacteriocin-like inhibitors, including Enterocin A, Listeriocine 743A, Enterocin P, Enterocin SE-K4 and Enterolysin A. Enterocin, produced by *Enterococcus*, is a small antibacterial peptide known to exhibit broad-spectrum inhibitory activity against spoilage bacteria and foodborne pathogens [1]. Among them, the structural peptide of Enterocin P was predicted to contain an N-terminal signal sequence, suggesting the possibility of extracellular secretion and functional activity. Although signal peptides were not detected in the remaining candidates, their localization within organized operon-like gene clusters, which include structural, immunity, and transporter components, may still imply potential antimicrobial functions. These features are consistent with previously reported enterococcal bacteriocin operons and suggest that these gene clusters may encode functionally active bacteriocins, although further experimental validation is required to confirm their phenotypic expression (Table 2) [1]. The genes associated with probiotic features, such as bacteriocin production, acid and bile salt tolerance, epithelial cell adhesion, and stress response, are detailed in Table 1.

In the complete genome of *E. faecium* strain GB_C_05, the species-specific antibiotic resistance genes *aac(6')-Ii* and *msr(C)* were detected on the chromosome rather than on a plasmid, suggesting a low likelihood of their transmission to other microorganisms [11]. In the VFDB results, a total of 15 genes associated with virulence factors were identified in the chromosome. It contains genes *acm*, *sagA*, *sgrA*, and *pilB*, which are adherence-related genes and are involved in biofilm formation, and these genes are commonly found in Enterococcus. The presence of these genes may confer advantages to the strain by facilitating effective gut colonization, enhancing adhesion to the intestinal epithelium, and providing protection against harmful bacteria [10]. Notably, key virulence determinants such as gelatinase (*gelE*), cytolysin (*cyl*), and vancomycin resistance genes (*vanA*, *vanB*) were not detected. While experimental validation is necessary, the absence of these major virulence markers may suggest a potential safety profile for *E. faecium* GB_C_05 as a probiotic candidate.

ACCEPTED

In summary, although experimental validation is needed to confirm the phenotypic expression of genes encoding enzymes essential for carbohydrate transport and metabolism, such as α-galactosidase, their presence suggests the potential for supporting beneficial microbial activity and contributing to carbohydrate metabolism. Such functional traits may help enhance the nutritional value of livestock products and support the strain's possible application as a feed additive. Therefore, the whole genome analysis of *Enterococcus faecium* GB_C_05 is expected to unlock various application possibilities in the livestock industry and the field of feed additives.

ACCEPTED

References

1. Ben Braïek O, Smaoui S. Enterococci: Between Emerging Pathogens and Potential Probiotics. Biomed Res Int. 2019;2019.https://doi.org/10.1155/2019/5938210

2. Franz CMAP, Huch M, Abriouel H, Holzapfel W, Gálvez A. Enterococci as probiotics and their implications in food safety. Int J Food Microbiol. 2011;151(2):125-40.https://doi.org/10.1016/j.ijfoodmicro.2011.08.014

3. Keum GB, Pandey S, Kim ES, Doo H, Kwak J, Ryu S, et al. Understanding the Diversity and Roles of the Ruminal Microbiome. J Microbiol. 2024;62(3):217-30.10.1007/s12275-024-00121-4

4. Landman F, Jamin C, de Haan A, Witteveen S, Bos J, van der Heide HG, et al. Genomic surveillance of multidrug-resistant organisms based on long-read sequencing. medRxiv. 2024:2024.02. 18.24301916.https://doi.org/10.1101/2024.02.18.24301916

5. Lee JY, Kong M, Oh J, Lim J, Chung SH, Kim JM, et al. Comparative evaluation of Nanopore polishing tools for microbial genome assembly and polishing strategies for downstream analysis. Sci Rep. 2021;11(1).https://doi.org/10.1038/s41598-021-00178-w

6. 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

7. 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

8. 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.https://doi.org/10.1186/1471-2164-9-75

9. Cantalapiedra CP, Hernández-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

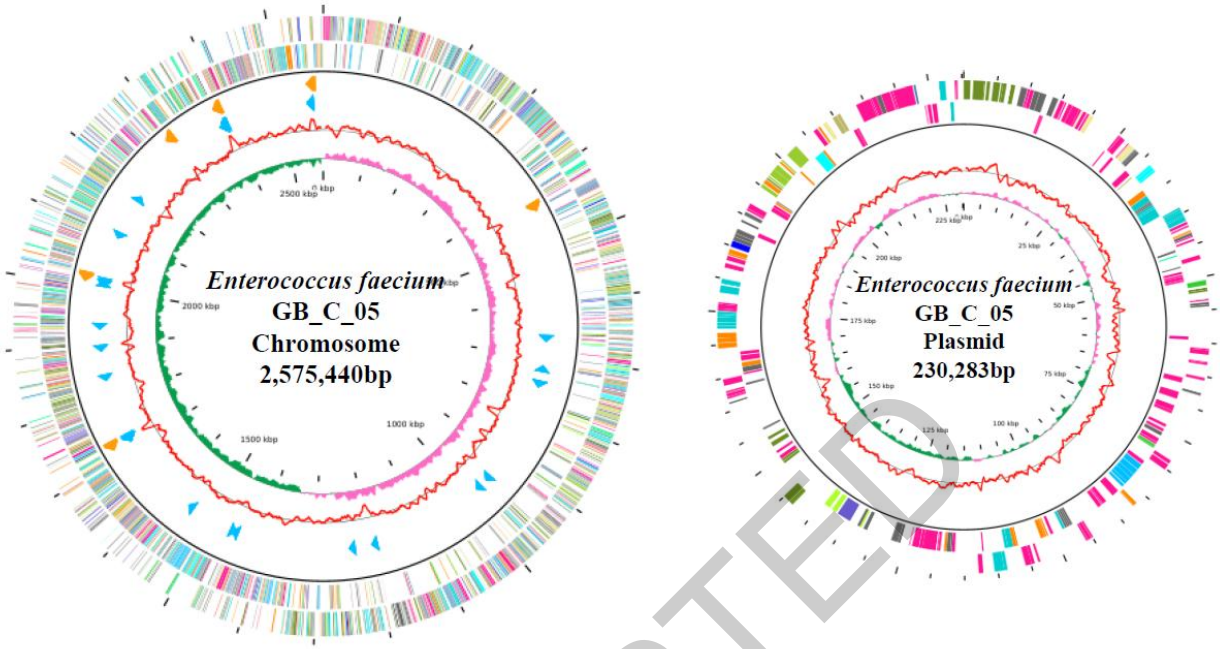
10. Kim D, Kim SA, Jo NG, Bae JH, Nguyen MT, Jo YM, et al. Phenotypic and genomic analyses of bacteriocin-producing probiotic Enterococcus faecium EFEL8600 isolated from Korean soy-meju. Front Microbiol. 2023;14.https://doi.org/10.3389/fmicb.2023.1237442

11. Urshev Z, Yungareva T. Initial safety evaluation of Enterococcus faecium LBB.E81. Biotechnol Biotec Eq. 2021;35(1):11-7.https://doi.org/10.1080/13102818.2020.1840438

ACCEPTED

142 **Table and Figure**

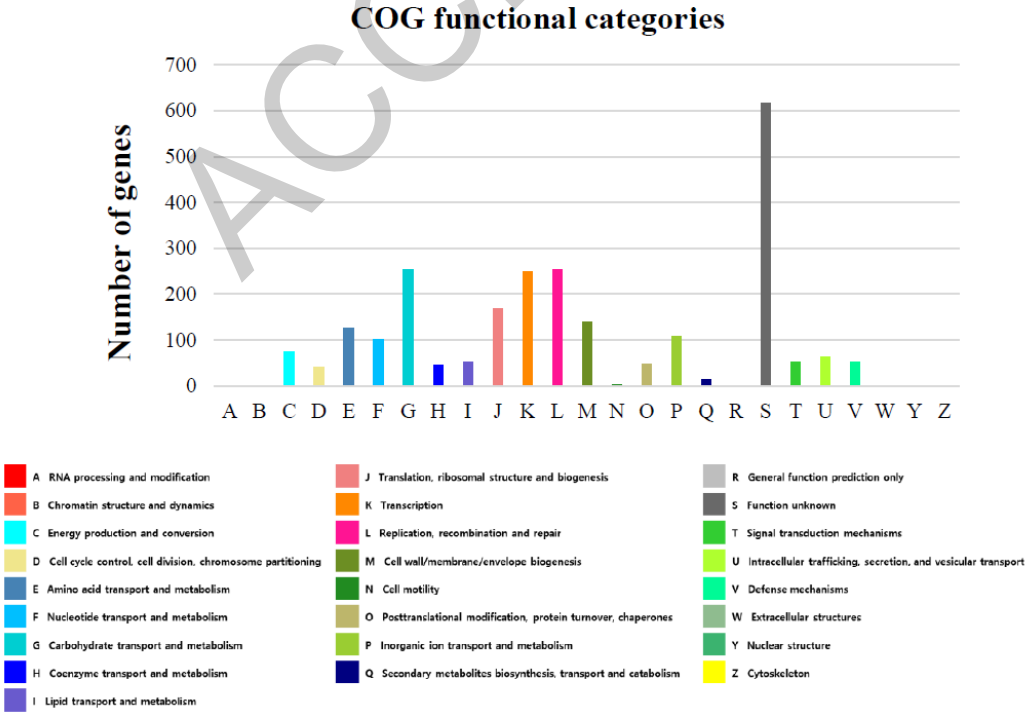
A



Type	BioProject	BioSample	Accession No.	Length (bp)	GC content (%)	CDSs	tRNA	rRNA
Chromosome	PRJNA1066497	SAMN39489522	CP142862.1	2,575,440	38.2	2756	70	18
Plasmid			CP142861.1	230,283	35.2	391	-	-

143

B



144

ACCEPTED

Figure 1. Genome map of *Enterococcus faecium* strain GB_C_05 and the functional categorization of predicted protein coding genes. In the visualization, the outer ring signifies the positions of all annotated gene coding regions (ORFs), while the inner ring highlighted in red represents the guanine + cytosine (GC) content. GC skew is indicated by pink and green color variations, and rRNA and tRNA operons are marked with orange and sky-blue arrows, respectively. (A) Circular genome map of Enterococcus faecium GB_C_05 with annotated ORFs color-coded according to Clusters of Orthologous Groups (COG). (B) COG functional categorization of predicted protein-coding sequences.

ACCEPTED

153 **Table 1. Predicted CDSs involved in probiotic potency in *Enterococcus faecium* GB_C_05**

Categories	Related protein	Start position	End position
Bacteriocin	Enterocin_A	2180744	2200936
	Listeriocine_743A	35153	55273
	Enterocin_P	50144	70294
	Enterocin_SE-K4	153539	173719
	Enterolysin_A	120245	140788
pH	Alkaline phosphatase synthesis transcriptional regulatory protein PhoP	891851	892555
	ATP synthase subunit A	704929	706710
	ATP synthase subunit B	706703	708078
	ATP synthase subunit C	703620	704606
	ATP synthase gamma chain	743042	743944
	ATP synthase epsilon chain	745387	745809
Bile	Choloylglycine hydrolase	882795	883769
Temperature	Copper chaperone	261689	261916
	Chaperone protein DnaJ	1040454	1041740
	Chaperone protein DnaK	1038474	1040303
	Chaperone protein ClpB	1149148	1151757
	60 kDa chaperone	2166820	2168445
Oxidation	Glutathione reductase	2543903	2545249
	Glutathione peroxidase	435175	435645
	Glutathione biosynthesis bifunctional protein	106865	109132
	NADH peroxidase	333723	335018
	NADH dehydrogenase	2474259	2474888
	Thioredoxin	244216	244536
	Thioredoxin reductase	800216	801142
	Quinone oxidoreductase	643735	644718

154

ACCEPTED

Table 2. Summary of identified bacteriocin-related gene clusters in *Enterococcus faecium* GB_C_05.

Identified Bacteriocin (AOI Class)	Location	Signal peptide	Match(%)	Molecular weight (kDa)
Enterocin A	Chromosome	NO	100	3.74
Listeriocin 743A	Plasmid	NO	100	4.84
Enterocin P	Plasmid	Yes	100	7.81
Enterocin SE-K4	Plasmid	NO	N/A*	N/A*
Enterolysin A	Plasmid	NO	35.93	44.11

"*" indicates values not assigned due to the absence of a confidently matched structural protein, although the cluster was predicted in the BAGEL4 AOI region.

ACCEPTED