



OPEN

## High-quality metagenome-assembled genomes from proximal colonic microbiomes of synbiotic-treated korean native black pigs reveal changes in functional capacity

Jaehoon Jung<sup>1,2</sup>, Andrew W. Bugenyi<sup>3,4</sup>, Ma-Ro Lee<sup>5</sup>, Yeon-Jae Choi<sup>6</sup>, Ki-Duk Song<sup>7</sup>, Hak-Kyo Lee<sup>5,6,7</sup>, Young-Ok Son<sup>8,10</sup>, Dong-Sun Lee<sup>9,10</sup>, Sang-Chul Lee<sup>11</sup>, Young-June Son<sup>11</sup> & Jaeyoung Heo<sup>5</sup>✉

Synbiotics are feed supplements with the potential to promote health and productivity in pigs partly, through modulation of the intestinal microbiome. Our study used shotgun sequencing and 16S rRNA gene sequencing techniques to characterize the effect of a synbiotic containing three *Lactobacillus* species and a fructo-oligosaccharide on the proximal colonic microbiome of 4- to 7-month-old Korean native black gilts. With shotgun sequencing we constructed unique metagenome-assembled genomes of gut microbiota in Native Black Pig for the first time, which we then used for downstream analysis. Results showed that synbiotic treatment did not alter microbial diversity and evenness within the proximal colons, but altered composition of some members of the *Lactobacillaceae*, *Enterococcaceae* and *Streptococcaceae* families. Functional analysis of the shotgun sequence data revealed 8 clusters of orthologous groups (COGs) that were differentially represented in the proximal colonic microbiomes of synbiotic-treated Jeju black pigs relative to controls. In conclusion, our results show that administering this synbiotic causes changes in the functional capacity of the proximal colonic microbiome of the Korean native black pig. This study improves our understanding of the potential impact of synbiotics on the colonic microbiome of Korean native black pigs.

### Abbreviations

cfu	Colony forming units
PCR	Polymerase Chain Reaction
rRNA	Ribosomal RNA (ribose nucleic acids)
DNA	Deoxyribose nucleic acids

<sup>1</sup>Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul 151-742, Republic of Korea. <sup>2</sup>eGnome, 26 Beobwon-ro, Songpa-gu, Seoul 05836, Republic of Korea. <sup>3</sup>Department of Agricultural Convergence Technology, Jeonbuk National University, Jeonju 54896, Republic of Korea. <sup>4</sup>National Agricultural Research Organization, Mbarara, Uganda. <sup>5</sup>Department of Animal Biotechnology, Jeonbuk National University, Jeonju 54896, Republic of Korea. <sup>6</sup>International Agricultural Development and Cooperation Center, Jeonbuk National University, Jeonju 54896, Korea. <sup>7</sup>The Animal Molecular Genetics and Breeding Center, Jeonbuk National University, Jeonju 54896, Republic of Korea. <sup>8</sup>Department of Animal Biotechnology, Faculty of Biotechnology, College of Applied Life Sciences and Interdisciplinary Graduate Program in Advanced Convergence Technology and Science, Jeju National University, Jeju 63243, Republic of Korea. <sup>9</sup>Faculty of Biotechnology, College of Applied Life Sciences and Interdisciplinary Graduate Program in Advanced Convergence Technology and Science, Jeju National University, Jeju 63243, Republic of Korea. <sup>10</sup>Jeju Microbiome Research Center, Jeju National University, Jeju, Jeju Special Self-Governing Province 63243, Republic of Korea. <sup>11</sup>Cronex Co., Cheongju 28174, Republic of Korea. ✉email: jyheo@jbnu.ac.kr

NCBI	National Center for Biotechnology Information
KEGG	Kyoto Encyclopedia of Genes and Genomes
MAG	Metagenome-assembled genomes
AMR	Antimicrobial resistance
COG	Cluster of Orthologous Groups

The Jeju Black Pig is a small, black domestic pig breed that is native to the Korean island of Jeju where it has been conserved for generations<sup>1</sup>. The breed is well adapted to foraging in the rough landscape of the island, where it endures windy conditions and an overall harsh climate<sup>2</sup>. Meanwhile the evolutionary pressure in modern commercial pig farming systems on the mainland has led to the dominance of exotic and commercially developed breeds. Modern commercial breeds have been bred for high performance traits such as prolificacy, mothering ability, number of piglets born, weight gain, growth rate, back fat thickness, lean meat, among others. Despite a low genetic potential for weight gain and overall performance in Jeju black pigs, their pork products have a unique taste which makes them highly sought after and relatively pricier on the market<sup>3,4</sup>. In recent years, there have been efforts to transform Jeju pork production from a traditional backyard system to a more productive, industrialized system while continuing to preserve the breed. Innovative management strategies are being employed to optimize the productivity of this native pig breed. Among these strategies is the use of probiotics, prebiotics and synbiotics to promote health and weight gain in the farmed pigs.

Synbiotics are a combination of probiotics and prebiotics that are used as feed additives or supplements in order to grant health benefits to the host<sup>5</sup>. Probiotics have been defined as “live microorganisms that, when administered in adequate amounts, confer a health benefit on the host”<sup>6</sup>. Prebiotics are dietary components, usually plant fiber that are undigestible by the host but selectively support the growth or activity of a limited group of beneficial microbial genera or species in the gut<sup>7</sup>. These synbiotic combinations modify intestinal microbial composition and influence the development and functioning of the gut with putative benefits to health and performance of pigs. Probiotics containing members of the genus *Lactobacillus* have been used to modulate the intestinal microbial flora and promote gut health, weight gain and overall performance of pigs<sup>8–10</sup>. This genus is a core member of the pig’s intestinal microbial ecosystem<sup>11</sup>. *Lactobacilli* are of interest because of their inhibitory role on some potentially pathogenic bacteria in the swine gut. They produce lactic acid which has a disruptive effect on the walls of gram-negative bacteria, such as the *E. coli* and salmonella which are key pathogens in domestic pigs<sup>12,13</sup>. This in addition to their production of antimicrobial peptides<sup>14</sup> and modulation of immune response in pigs<sup>15</sup> makes *Lactobacillus*-based probiotics potentially good alternatives to antibiotics in pig farming. Besides contributing to the host’s immune response to enteric pathogens, some strains of *Lactobacillus* promote nutrient digestibility by producing dietary enzymes such as phytases and amylases which then enables breakdown of indigestible dietary components<sup>16</sup>. Several studies have shown that combining prebiotics such as fructo-oligosaccharides with probiotics synergistically enhances the benefits of the probiotic<sup>17,18</sup>. These effects are most important in the colonic environment.

As a site with the highest number of microorganisms, the proximal colon plays very important roles not only in the excretion of waste but also in the absorption of water and electrolytes as well as absorption of vitamins produced by its complex and rich microbial flora<sup>19</sup>. A healthy balance of the microbial ecosystem within the colon is central to its role in the swine gut in that colonic mucosa is the main surface where the gut microbiota and the immune system interact<sup>20</sup>. Therefore, manipulation of this environment to promote the health of Native Black pigs is of interest. The potential of probiotics to promote the productivity of these pigs, can be understood further, by studying their effects on the proximal colonic microbial ecosystem. Metataxonomic techniques such as targeted sequencing of the 16S ribosomal RNA gene and metagenomic techniques like, shotgun sequencing, are frequently used to study microbial composition and infer biological functions in such communities<sup>21,22</sup>. The former approach is cheaper since, by targeting one gene (and typically, a fragment of it), the method imposes lower demands on sequencing coverage, compared to the latter, which probes all prevailing genes in an environment. However, techniques that target marker genes, are limited when studying functional profiles of microbial communities<sup>22</sup>. Even though computational techniques can be used to predict functionality from say, 16SrRNA gene-based taxonomies<sup>23</sup>, shotgun sequencing offers better resolution of functional gene composition within microbial environments<sup>22</sup>. Use of shotgun sequencing data is highly dependent on the completeness of the reference databases for the environment in question. To date, however, the microbiome of Jeju black pig has not been sufficiently studied to support a metagenomic assembly approach.

In this study, we used both shotgun and 16S rRNA sequencing techniques to explore, for the first time, the proximal colonic microbiota of Jeju Native Black pigs. We used a bioinformatic approach to construct a custom database for the Jeju Black pig proximal colonic samples to improve taxonomic classification of shotgun sequence data. We then evaluated the effect of a synbiotic formulation containing *Lactobacillus buchneri* NLRI-1201, *L. plantarum* NLRI-101, *L. casei* DK128 and the prebiotic, fructo-oligosaccharide on the microbial composition within the proximal colon.

## Methods

**Animals and sample collection.** All animal experiments were conducted in a facility of contract research organization at Cronex Co., Ltd. (Cheongju, Korea), with following the regulations of the Cronex Co., Ltd. Animal Care and Use Committee (CRONEX-IACUC 202002007). Jeju black pigs (M-Pig®) were purchased from Cronex Co., Ltd. and housed in Cronex Animal SPF (Specific Pathogen Free) facility for whole experimental period, where was environmentally controlled with temperature and relative humidity at 22 ± 2 °C and 50 ± 10%, respectively, under a 12-h light–dark cycle. Each cage (4 pigs per cage, 1.0 m<sup>2</sup> per pig) was equipped with plastic mesh floor, two round feeders, and two nipple drinkers. Samples were collected from 7-month-old female Jeju

Nutrient	Basal diet
<b>Ingredients, g/kg</b>	
Corn	392.5
Wheat	200.0
Soybean meal	137.3
Wheat bran	104.0
Soybean hull	54.1
Rice bran	40.0
Fiber feed	20.0
Limestone	15.6
Tallow	10.0
Mono-dicalcium phosphate	9.0
NaCl	5.0
Choline chloride	2.2
Acidifier	2.0
Tryptophan	2.0
Lysin	1.8
Toxin-binder	0.5
Vitamin Premix <sup>a</sup>	2.0
Mineral Premix <sup>b</sup>	2.0
Total	1000.00
<b>Analyzed crude nutrient and energy contents</b>	
Moisture, %	12.98
Crude protein, %	15.65
Crude fat, %	3.58
Crude ash, %	2.83
Crude fiber, %	5.64
Gross energy, Kcal/kg	3836.00

**Table 1.** Composition, nutrient and energy contents of basal diet for Jeju native pig. <sup>a</sup>The vitamin premix provided per kg diet: vitamin A, 12,000 IU; vitamin D3, 2,000 IU; vitamin E, 100 IU; vitamin K3, 4.5 mg; vitamin B1, 2 mg; vitamin B2, 7 mg; vitamin B3, 45 mg; pantothenic acid, 30 mg; vitamin B6, 4.5 mg; Biotin, 0.5 mg; Folic acid, 3.5 mg; vitamin B12, 0.03 mg; antioxidant, 6.6 mg. <sup>b</sup>The mineral premix provided per kg diet: Fe, 150 mg; Zn, 85 mg; Mn, 37 mg; Cu, 11 mg; Co, 2 mg; I, 0.3 mg; Se, 0.15 mg.

black pigs following a 3-month-long experimental study. At the start of the study, eight 4-month-old pigs were randomly recruited into 2 groups ( $n=4$ , in each group). Each group was subjected to a separate diet: either a basal diet (Control group) or synbiotics-supplemented diet (Treatment group). The experimental animals had similar mean body weight both at the start ( $32.8 \pm 4.5$  and  $32.6 \pm 2.5$  kg, respectively) and at the end of the experiment ( $60.4 \pm 1.3$  and  $56.5 \pm 6.8$  kg, respectively). All animals were monitored daily by caretakers and the investigating team to ensure their health and welfare. During health monitoring, the care takers observed the animals for signs of GIT conditions (reduced appetite, diarrhea, constipation, rectal prolapses), urogenital tract conditions (discoloration in urine, vaginal discharge) as well as respiratory conditions (coughing, sneezing, rhinitis). All animals remained healthy throughout the study.

The formulation of basal diet is presented in Table 1. The treatment group was allowed ad libitum access to a basal diet supplemented with a synbiotic containing *Lactobacillus buchneri* NLRI-1201 ( $1.2 \times 10^8$  colony forming units (cfu)/kg feed), *L. plantarum* NLRI-101 ( $1.6 \times 10^8$  cfu/kg feed), *L. casei* DK128 ( $1.4 \times 10^8$  cfu/kg feed), and fructo-oligosaccharide (5 g/kg feed). The multi-probiotic and fructo-oligosaccharide from *Saccharum barberi* were purchased from SunBio (Gunpo-Si, Gyeonggi-do, South Korea) and Samyang Corporation (Seongnam-Si, South Korea), respectively. At the end of the 3 months (12 weeks), 8 pigs were sacrificed (4 pigs from each group) and their proximal colonic contents collected for metagenomic analysis. All methods were performed in accordance with the relevant guidelines and regulations of the Cronex Co., Ltd. Animal Care and Use Committee. Genomic DNA (gDNA) from the 8 samples was isolated using DNA extraction kit (NucleoSpin DNA Stool Kit, Macherey–Nagel). After DNA extraction, the metagenome was sequenced and analyzed as the flowchart shown in Fig. S1.

**16S rRNA metagenome sequencing.** From the isolated gDNA, the V3–V4 hypervariable region of the 16S rRNA gene was amplified using the universal prokaryotic primers Bakt\_341F (5'-CCTACGGGNGGCWGC AG-3') and Bakt\_805R (5'-GACTACHVGGGTATCTAATCC-3')<sup>24</sup>. Amplicon libraries were prepared, pooled, and loaded onto an Illumina MiSeq flow-cell using a 600-cycle reagent cartridge (v3 reagent kit). Sequencing was done using a  $2 \times 301$  bp paired-end sequencing method.

**Shotgun metagenome sequencing.** In this technique, isolated gDNA was sheared using the S220 Focused-ultrasonicator (Covaris, Inc., Adaptive Focused Acoustics, Massachusetts, US). Library preparation was performed using MGIEasy DNA library prep kit (MGI, Shenzhen, China) according to the manufacturer's instructions. Briefly, After Size-selection of fragmented gDNA using AMPure XP magnetic beads, the fragmented gDNA was end-repaired and A-tailed at 37 °C for 30 min, and 65 °C for 15 min. Indexing adapters were ligated to the ends of the DNA fragments at 23 °C for 60 min. After cleanup of adapter-ligated DNA, PCR was performed to enrich those DNA fragments that have adapter molecules. Thermocycler conditions were as follows: 95 °C for 3 min, 7 cycles of 98 °C for 20 s, 60 °C for 15 s, and 72 °C for 30 s, with a final extension at 72 °C for 10 min. The double stranded library is quantified using QuantiFluor ONE dsDNA System (Promega, Madison, WI, USA). The library was then circularized at 37 °C for 30 min, and then digested at 37 °C for 30 min, followed by cleanup of circularization product. To make DNA nanoball (DNB), the library was incubated at 30 °C for 25 min using DNB enzyme. Finally, the libraries were quantified using QuantiFluor ssDNA System (Promega, Madison, WI, USA). Sequencing of the prepared DNB was conducted on the MGISEQ-2000 system (MGI, Shenzhen, China) with 150 bp paired-end reads. All the sequencing process was conducted by LAS, Inc. (Gimpo, South Korea).

**16S rRNA metagenome analysis.** Denoising and merging, the read pairs were processed using DADA2 package (q2-dada2 v2020.8.0)<sup>25</sup> (with options *-p-trunc-len-f 275 -p-trunc-len-r 225*). Denoised features were classified by Qiime2 q2-feature-classifier plugin<sup>26</sup> using a naïve Bayes classifier trained on the V3-V4 hypervariable region of 16S rRNA gene sequences in the Silva Database v138<sup>27</sup>. This denoising and taxonomic classification process was done using the Qiime2 (v2020.08) software<sup>28</sup>. Alpha and beta diversity metrics were calculated using normalized/rarefied sub-samples (14,397 features) of the denoised reads. Analysis of alpha and beta diversity was achieved using Qiime2 diversity core-metrics-phylogenetic plugin and results were visualized in R<sup>29</sup> using the package ggplot2<sup>30</sup>. Following Thorsen and colleagues<sup>31</sup> recommendations for small datasets, we normalized the features using Trimmed Means of M-values (TMM), and explored differentially abundant taxonomy using edgeR<sup>32</sup>.

**De novo assembly and metagenome-assembled genome (MAG) construction.** Above all, adapters and low-quality bases (Q < 20) were trimmed using trimmomatic v0.39<sup>33</sup>, and the subsequent reads were mapped to the *sus scrofa* (Sscrofa11.1) reference genome using a read alignment tool, bowtie2 v2.2.6<sup>34</sup>. Host genome sequences were removed by extracting pair-unmapped reads. To reconstruct metagenome-assembled genomes (MAGs) from our samples, the extracted reads were used as input for de novo assembly using the metaSPAdes v3.14.1 tool<sup>35</sup>.

In order to group the short sequence reads into draft genomes (metagenome binning), reads were first remapped to the assembled contigs with bowtie2<sup>34</sup>, and their depths calculated using the *jgi\_summarize\_bam\_contig\_depth* script in MetaBAT2<sup>36</sup>. The metagenome binning process in metaBAT2 tool<sup>36</sup>, used the following options: *-minCV 0.5, -minContig 2,000, -minCVSum 0.5, -maxP 92, -maxEdges 150*. From the constructed bins, low quality bins (< 50% completeness, > 5% contamination) were filtered out using CheckM v1.1.3<sup>37</sup>. Filtered bins were then aggregated and dereplicated using dRep v3.2.0<sup>38</sup> with a similarity threshold set at 95% (-sa 0.95).

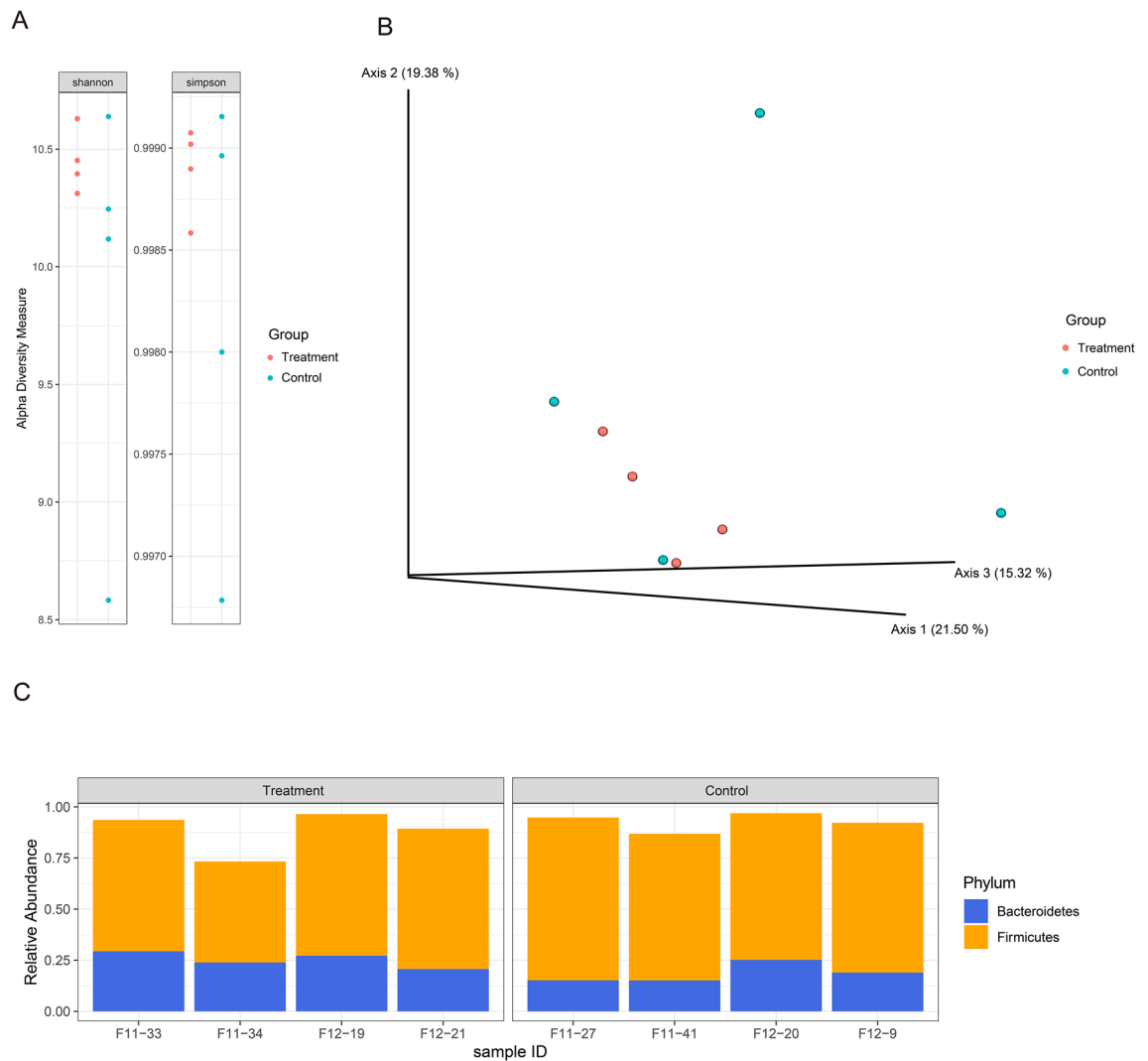
For taxonomic classification of MAGs, the high quality and dereplicated 360 bins were classified by utilizing the GTDB-Tk v1.5.0 software and basing on the Genome Taxonomy Database (GTDB R06-RS202)<sup>39</sup>. The GTDB-Tk tool uses average nucleotide Identity (ANI), phylogenetic tree topology and relative evolutionary divergence (RED) to assign genome-based taxonomies. For downstream analysis in TaxonKit v0.8.0<sup>40</sup> and kraken2 v2.1.1<sup>41</sup> the GTDB-based taxonomies were annotated to NCBI taxonomies.

To detect antibiotic-resistance genes within our metagenome samples, we used the ABRicate 1.0.1 tool (<https://github.com/tseemann/abricate>). We screened our per-sample assembled contigs (metaSPAdes output) based on NCBI AMRFinderPlus database<sup>42</sup>.

**Bioinformatic analysis of read-based taxonomies.** A read-based taxonomic classification was conducted using kraken2 v2.1.1 software<sup>41</sup>. We used both the in-built Standard Kraken2 database, and a custom host database to classify the paired end reads from the Jeju Black pig colonic microbial environment. The custom host database was built based on the default database and included the National Center for Biotechnology Information's (NCBI) RefSeq genomes as well as our classified MAGs. For comparison, reads were aligned to ChocoPhlAn database (2019.01v) using MetaPhlAn3 (v3.0.4), and UniRef90 database using HuManN3 (v3.0.0)<sup>43</sup>.

Following classification in Kraken, and with a normalized read length parameter set at 150, the relative taxonomic abundances of microbes in the samples were re-estimated using bracken v2.6.2 software<sup>44</sup>. Alpha and beta diversity analyses were calculated and visualized using the R package, phyloSeq (v1.34.0)<sup>45</sup>. Differential abundance of the microbes in the samples was statistically analyzed using the edgeR package<sup>32</sup> in R. Only species with a Log-transformed count-per-million value > 1 and a False Discovery rate (FDR) < 0.01 were considered differentially abundant with statistical significance.

**Gene-based metagenome assembly.** For gene-based differential abundance analysis, PLASS assembly (release 4-687d7)<sup>46</sup> was employed to obtain protein sequences for each sample. Assembled protein sequences were merged, dereplicated and clustered using the *easy-linlinclust* option in MMseqs2 (version 13.45111)<sup>47</sup> with sequence identity parameter set at 0.5. Sequence searches were then performed using DIAMOND 2.0.11 software<sup>48</sup> and with the option "*blastx -evalue 0.001 -fast -max-target-seqs 1*", gene cluster abundances were estimated. Clustered representative sequences were then functionally annotated using emapper-2.1.5<sup>49</sup> basing on eggno orthology data<sup>50</sup>. The R package, edgeR<sup>32</sup> was then used to estimate group specific changes in relative



**Figure 1.** Microbiome profiling results from 16S rRNA gene amplicon sequencing. **(A)** Alpha diversity of the microbial communities within samples in the study groups. **(B)** Principal coordinate analysis (PCoA) plot of the beta diversity based on the Bray–Curtis dissimilarity index between microbial community samples in the study groups. **(C)** Relative composition of the two dominant phyla based on taxonomic analysis of the 16S rRNA amplicon sequence data.

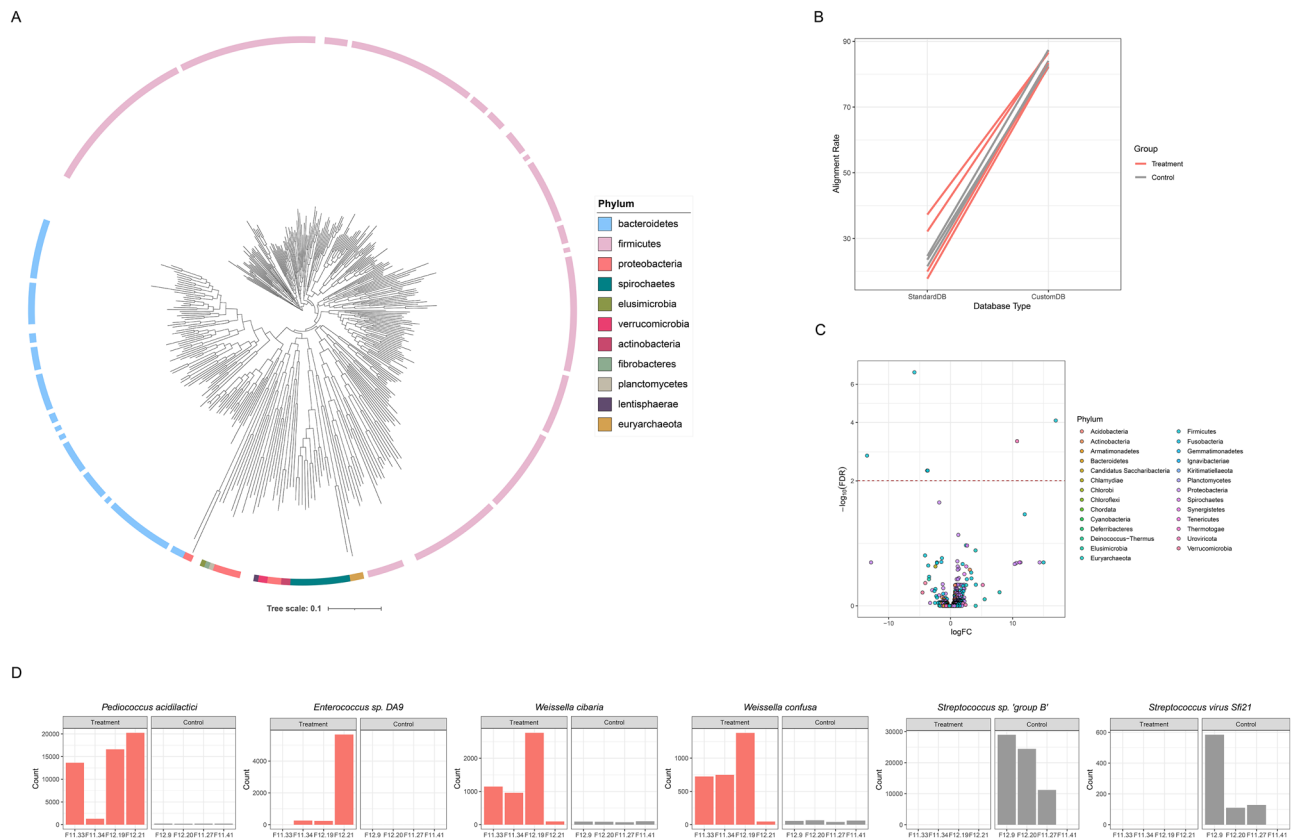
abundances of genes within the colonic microbiomes of the study animals. Statistical significance was considered at  $FDR < 0.01$ .

**Ethical approval.** All experiments were performed in compliance with the ARRIVE guidelines<sup>51</sup>.

## Results

**16S rRNA metagenome sequencing.** To investigate possible difference in taxonomic composition between the groups, V3–V4 hypervariable regions of the 16S rRNA gene were sequenced from microbial communities in the proximal colon of the study animals. Paired-end sequences were quality-filtered, merged and denoised leaving > 10,000 high quality reads. We assessed possible difference in alpha diversity between the two groups based on 16S rRNA metagenome sequences. Alpha diversity analysis using Shannon and Simpson indices indicated that there was no significance in within-group diversity between our symbiotic-fed animal (treatment group) and controls. The mean number of Shannon index in the treatment group was  $10.447 \pm 0.135$ , and  $9.896 \pm 0.903$  for control group. Also, Simpson index was  $0.9989 \pm 0.0002$  and  $0.9982 \pm 0.0011$  for treatment group and control group, respectively (Fig. 1A). Analysis of beta diversity distances using Bray–Curtis dissimilarity showed that symbiotic-fed group is well clustered while control group showed sparse distribution (Fig. 1B).

Taxonomic analysis indicated that *Firmicutes* and *Bacteroidetes* were the most abundant phyla found in the samples. On average, *firmicutes* had a relatively lower abundance in the treatment group ( $62.8\% \pm 3.764$ ) compared to the control group ( $74.1\% \pm 3.764$ ). Meanwhile, members of the phylum, *Bacteroidetes* had a higher relative abundance ( $25.3\% \pm 3.815$ ) in the treatment group compared to controls ( $18.6\% \pm 4.728$ ) (Fig. 1C). However,



**Figure 2.** Analysis of the Metagenome-Assembled Genomes (MAGs) obtained using a host-customized database. **(A)** Phylogenetic tree of MAGs **(B)** rate of classified reads using standard database and custom database in kraken2. **(C)** Volcano plot for differentially abundant taxonomies. Each point represents the species-level taxonomies colored with phylum taxa. The taxa are presented with  $\log_2$  fold change (X-axis) and  $\log_{10}$ (FDR-corrected P-value) (Y-axis). Horizontal line represents P-value cut off (FDR P-value 0.01). **(D)** Differentially abundant taxonomies (FDR < 0.01) between groups. Red bars represent composition in samples from the treatment group while gray bars, composition in samples from the control group.

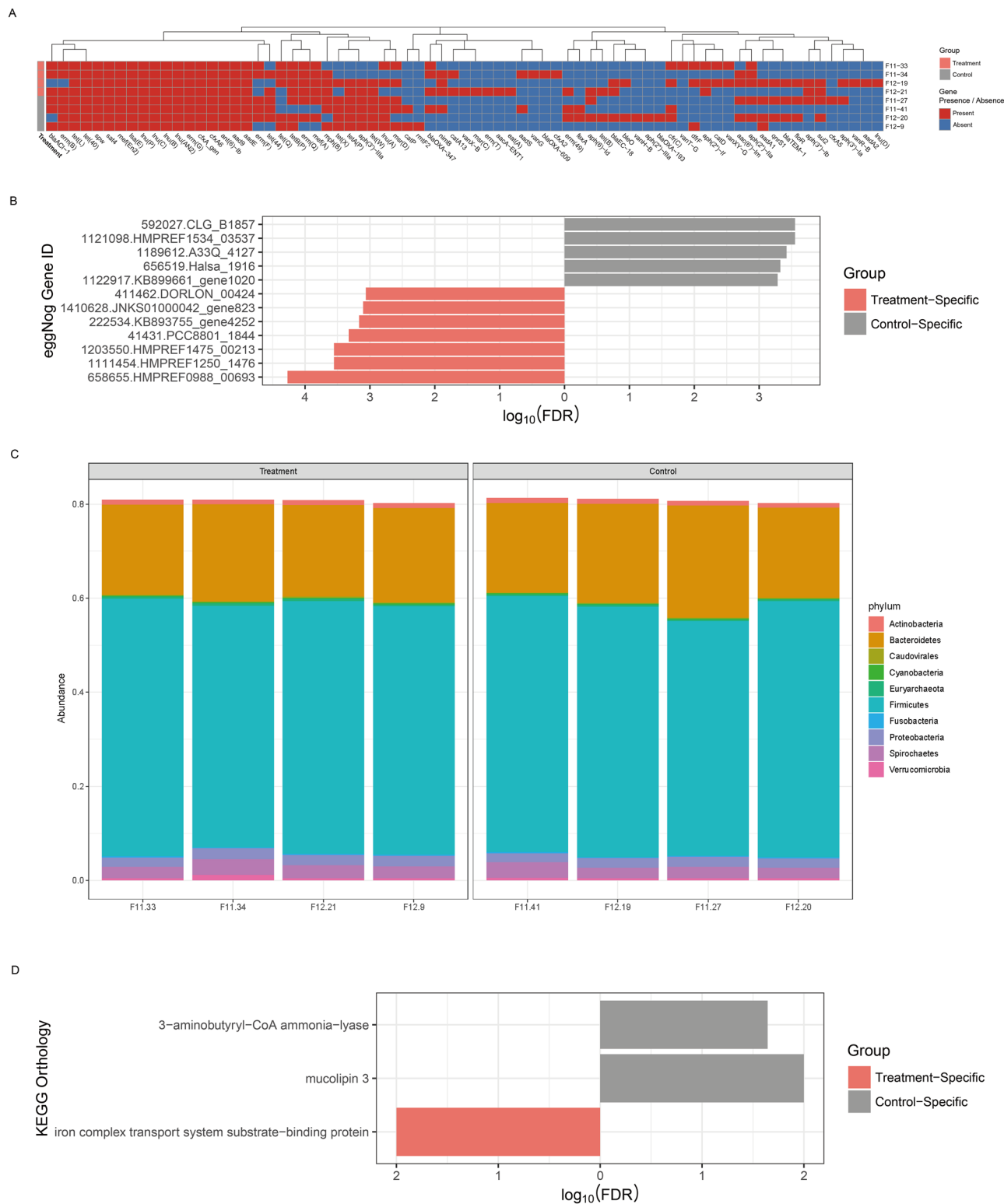
differential abundance in genus level testing between groups found no statistical significance (FDR p-value < 0.05) in the observed differences.

**Shotgun metagenome sequencing.** Sequencing on MGISEQ generated 116 Giga Bases with mean quality scores of 34.8 from our samples. After quality trimming and removal of host-derived sequences, about 12.7 Giga Bases (a mean of 43 million read pairs) per sample remained. Single-sample assembly and binning of resulting contigs produced 1,603 bins, which were then quality checked in CheckM (Fig. S2). Concatenation, filtration, and dereplication of all the bins resulted in a total of 360 high-quality metagenome-assembled genomes (MAGs). Taxonomies assigned by a GTDB-based classification of the MAGs as well as the phylogenetic relationships between the MAGs are presented in Fig. 2A. The 360 MAGs were dominated by members of the phyla, Firmicutes (214 MAGs or 59.44%) and Bacteroidetes (73 MAGs making up 20.28%) (Figs. 2A and 3C).

Read-based classification of the shotgun sequences based on the Standard Kraken2 Database yielded a low average alignment (25.07%) (Fig. 2B). To improve classification rate, we built a custom host database which put together our 360 MAGs and the Standard Kraken DB. With this, we were able to achieve a higher average read alignment rate (84.49%). Also, Metaphlan3/Humann3 read classification analysis were conducted and compared with Kraken2 results, showing that almost all of reads were not aligned in nucleotide alignment and about 50 ~ 60% of reads were not aligned in translated alignment (Table S1).

With the read-based taxonomies, we estimated the alpha- and beta-diversities of the microbiota within the proximal colons of the Jeju black pig gilts. By estimating colonic microbial diversities using our shotgun dataset, we obtained lower alpha diversities in the treatment group relative to controls (Fig. S3A). This is in contrast to earlier results from diversity analysis using the 16S rRNA dataset which showed the reverse; higher alpha diversities in the treatment group compared to the controls group (Fig. 1A). However, data from both sequencing methods showed larger deviations in the alpha diversity values within the treatment group compared to the control group. Analysis of Beta diversity distances found no group specific clustering (Fig. S3B).

**Differential abundance analysis.** Considering the indistinguishability in abundance and diversity between groups at community level, further analysis was conducted to explore differential abundance at the spe-



**Figure 3.** Functional Profiling from shotgun metagenome sequencing. **(A)** Antibiotic resistance genes identified from each sample. **(B)** Bar chart for differentially abundant genes between groups. FDR P-value 0.0005 was used as threshold. **(C)** Taxonomic composition profiling at gene-level. **(D)** Clusters of Orthologous Groups (COG) functional category profiling results. **(E)** Bar chart for differentially abundant KEGG Orthology functional category. Used FDR P-value was 0.05.

eggNOG orthology ID	logFC	FDR	Organism	COG	COG description
HMPREF0988_00693	- 7.14133	5.36E-05	<i>Lachnospiraceae bacterium</i>	S	Unknown
HMPREF1475_00213	- 6.10137	2.80E-04	<i>Prevotella oralis</i>	S	Unknown
HMPREF1250_1476	- 13.3631	2.80E-04	<i>Megasphaera</i> sp. BV3C16-1	Q	Secondary metabolites biosynthesis, transport, and catabolism
HMPREF1534_03537	5.414188	2.80E-04	<i>Bacteroides massiliensis</i>	S	Unknown
CLG_B1857	7.929562	2.80E-04	<i>Clostridium botulinum</i>	H	Coenzyme transport and metabolism
A33Q_4127	7.210967	3.78E-04	<i>Indibacter alkaliphilus</i> LW1	K	Transcription
PCC8801_1844	- 12.8701	4.71E-04	<i>Cyanothece</i> sp. PCC 8801	T	Signal transduction mechanisms
Halsa_1916	7.255372	4.71E-04	<i>Halanaerobium hydrogeniformans</i>	J	Translation, ribosomal structure and biogenesis
KB899661_gene1020	6.949026	5.19E-04	<i>Paenibacillus daejeonensis</i>	G	Carbohydrate transport and metabolism
KB893755_gene4252	- 12.4924	6.81E-04	<i>Frankia</i> sp.	L	Replication, recombination and repair
JNKS01000042_gene823	- 5.89862	7.85E-04	<i>Lachnospiraceae bacterium</i>	S	Unknown
DORLON_00424	- 7.82383	8.65E-04	<i>Dorea longicatena</i>	S	Unknown

**Table 2.** Differentially abundant genes from protein-level assembly.

cies level (Fig. 2C). In total, 6 species were identified as statistically different between the treatment group and the control group. Among these species, 5 were bacteria belonging to the firmicutes phylum, while the other was a bacteriophage (*Streptococcus virus sf21*) (Fig. 2D). On close inspection, the taxa that were overrepresented in the treatment group (treatment-specific taxa) belonged to 2 families; *Lactobacillaceae* family (*Weissella cibaria*, *W. confuse* & *Pediococcus acidilactici*) and *Enterococcaceae* family (*Enterococcus* sp. DA9). On the other hand, the control-specific taxa (*Streptococcus* sp. Group B) belonged to the *Streptococcaceae* family. Also, a *Streptococcus* specific bacteriophage, (*Streptococcus virus sf21*) was detected only in the control group.

**Functional profiling of pig microbiome.** To study the antimicrobial resistome of the proximal colonic samples from our Jeju black pig gilts, we scanned for antibiotic resistance genes within the metagenome assembled contigs. From the samples, we found a total of 73 antibiotic resistance genes (Fig. 3A). The distribution of these AMR genes varied among individuals and showed no differential abundances between groups.

We then investigated the functional role and diversity of proteins encoded by microbial genes associated with colonic microbiomes of the pig samples. To get insights into the gene cluster abundances in metagenomes, gene-based assembly was conducted, and 90,555,907 non-redundant coding sequences were generated. We then assigned functional categories to the genes by mapping them to Clusters of Orthologous Groups (COG) categories. Overall, results showed very little variation in the number of COGs between groups, and samples had coterminous COG profiles (Table S2). Results also showed that 7 COG categories of known functions had average abundances above 5.0% within the samples and together accounted for 53.56% of the total gene abundances. These categories included the COG L (replication & repair), COG G (carbohydrate metabolism and transport), COG M (Cell wall/membrane/envelop biogenesis), COG E (amino acid metabolism and transport), COG J (Translation), COG K (Transcription) and COG C (Energy production and conversion). Differential analysis uncovered 12 differentially abundant genes from 8 COG groups that passed the statistical significance cut-off of FDR corrected p-values < 0.0005 (Table 2, Fig. 3B).

Gene-based assignment of taxonomies yielded phylum-level profiles presented in (Fig. 3C). The proportion of the phylum, *Bacteroidetes* was 20.37% ( $\pm 0.016$ ), and that of *Firmicutes* was 53.18% ( $\pm 0.017$ ). These proportions were in agreement with the 16S rRNA dataset. Protein-level profiles were nearly the same between samples.

We also mapped our data to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database to yield KEGG orthologues. Analysis of the KEGG orthologs revealed a statistically significant increase in the 'Iron complex transport system substrate-binding protein' within the treatment group. While the relative increase of the 'mucolipin 3' and '3-aminobutyryl-CoA ammonia-lyase' groups in the control group was statistically significant. We however, found no statistically significant differences based on other functional categories (CAZyme, KEGG pathway and PFAM).

## Discussion

In this study, we looked to improve our understanding of the effect of a synbiotic on the colonic microbiome of Korean native black pigs that are traditionally reared on the Jeju Island. We combined shotgun metagenomic sequencing and 16s rRNA marker gene sequencing techniques to study the impact on the microbial composition in the proximal colon of these pigs. The results indicate that the synbiotic causes a subtle shift in the functional profile of the proximal colonic microbiome but no detectable shift in phylogenetic diversity. This is also the first study into the colonic microbiome of this native breed of pigs that are raised under the unique environmental conditions of Jeju Island.

During analysis of the shotgun metagenome sequence data, we encountered a low rate of alignment in the existing reference databases and tools (Fig. 2B, Table S1). One of the reason for this could be the limitation in coverage of a database that is still under curation, considering the quantity of data required to meticulously record all complete bacterial, archaeal, and viral genomes in all environments<sup>52</sup>. Secondly, we suspect that due to the spatial



isolation of this native breed of pigs on Jeju Island, their gut environments might represent unique, previously unstudied microbiomes which are not included in the kraken database. To overcome this limitation, we built a custom database for our data through identifying novel MAGs. Consequently 360 novel and high-quality MAGs were identified from the proximal colonic luminal content samples in this study (Fig. 2A). Classification based on the customized database resulted in high-resolution taxonomic composition, with fewer unclassified reads. Although 10 ~ 20% of the reads remained unclassified, we were able to explain more of the unknown reads from the sequence data. In this way, the analysis managed overcome the limitations imposed by a reference database that could not reflect many of the rarely researched microbiota of Native Black Pig.

We found that dietary supplementation with a synbiotic formula had no statistically significant effect on the microbial evenness, alpha- and beta-diversity in the proximal colon of Korean native pigs. The age of the piglets might be one of the reasons why the synbiotic did not have a detectable impact on the community-level microbial diversity<sup>8</sup>. The developing gut of young piglets is more susceptible to significant shifts and becomes less so as it develops into a richer, more mature and stable environment. Our study animals were weaner-growers or sub-adults which have a relatively stable gut microbial composition compared to younger piglets in the peri-weaning phases of growth. Although there was no obvious clustering according to study group, we found some unique/ differentially abundant taxa between the groups. There was an increased representation of some genera of the *Lactobacillaceae* family (*Weissella* spp., *Pediococcus* spp.,) and *Enterococcaceae* (*Enterococcus* spp.) probably due to the stimulatory effect of the probiotic *Lactobacilli*<sup>53</sup> or the fructo-oligosaccharides in the synbiotic supplement<sup>54</sup>. There was also a reduction in some species of the *Streptococcaceae* family (*Streptococcus* spp.) as well as a *streptococcus* targeting bacteriophage (*Streptococcus virus sfi21*).

In the functional profiling of the proximal colonic microbiomes, we used both shotgun sequence and 16S rRNA marker gene sequence data (Fig. 3). We were able to detect some statistically significant effects of the synbiotic treatment on the distribution of microbial functional gene groups using shotgun sequence data but not with the 16S rRNA sequence data. This was not surprising, as it has been reported in several previous studies that the shotgun sequence data yields more genes and functional metagenomes than the 16S rRNA sequence data<sup>55,56</sup>. During functional profiling, the shotgun sequence data enabled identification of antimicrobial resistance (AMR) genes in the proximal colonic microbiomes of our study animals. We found a total of 73 AMR genes in the proximal colonic microbiome of the study animals, and their abundance did not exhibit group-specific shifts in this study (Fig. 3A). Genes responsible for resistance against tetracyclines, macrolides, aminoglycosides, lincosamides and beta lactams were found in the majority if not all of the samples. This finding in Jeju Black pigs studied here, is similar to findings from several studies that have reported occurrence of resistance genes among commercial pig breeds<sup>57,58</sup>. Pollock and colleagues<sup>57</sup>, found that occurrence and abundance of these AMR genes, however, did not always correspond to antimicrobial exposure. In our study, the animals were not exposed to antibiotics throughout the experimental period, however exposure prior to inclusion into the study was not investigated. Nonetheless, AMR genes are known to be ubiquitous among microbial communities, having coevolved with bacterial strains as a strategy to counter toxic chemicals within their niches<sup>59</sup>.

In addition to identifying antimicrobial resistance genes within this microbial community, the metagenomic analysis provided insights into synbiotic-associated shifts in the functional capacity of the proximal colonic microbiomes (Fig. 3B). The analysis revealed 12 statistically significant differentially distributed gene (belonging to 8 COG groups) between the study groups. Synbiotic treated pigs were enriched in gene categories involved in secondary metabolites biosynthesis, transport, and catabolism (COG Q), Signal transduction mechanisms (COG T) and Replication, recombination, and repair (COG L). The COG Q category includes genes encoding Isochorismatase, an ether hydrolase that catalyzes the synthesis of iron-binding molecules called siderophores. Siderophores have been reported to act as virulence factors<sup>60,61</sup>, and in some instances to mediate microbial interactions that result in suppression of potential gut pathogens<sup>62,63</sup>. Genes in the COG T enable microorganism to sense, and respond to a number of intracellular and extracellular stimuli<sup>64</sup>. Therefore, an abundance of gene families involved in signal transduction could indicate a higher capacity for the bacteria in this group to respond to changes in their environment including nutrient changes, antibiotics, pH among others<sup>65</sup>. While the COG L family of genes is involved in DNA transposition in bacteria. An abundance of this gene family might indicate a high potential for genetic recombination<sup>66</sup> and/or lateral gene transfer in the microbial community within the proximal colon of synbiotic-fed pigs<sup>67</sup>. Interestingly, some of the relatively increased genes in the treatment group, originated from *Lachnospiraceae* bacterium and *Dorea longicatena*, which are stimulated by fructo-oligosaccharides<sup>68,69</sup>.

We also found that gene families involved in coenzyme transport and metabolism (COG H), transcription (COG K), translation, ribosomal structure, and biogenesis (COG J) as well as carbohydrate transport and metabolism (COG G) were significantly lower in the synbiotic-fed pigs compared to controls. Genes in the COG H category are involved in coenzyme transport and metabolism. Coenzymes are involved in a wide range of intracellular reactions. They play important roles in energy metabolism, cellular response to stress, intracellular redox homeostasis, nucleic acid repair, xenobiotic metabolism, and sometimes, modulation of microbial virulence<sup>70-75</sup>. Genes belonging to the COG K category code for RNA polymerase sigma factors and are involved in initiation of bacterial transcription during exponential growth and gene transcription in response to environmental stressors as well as during morphological changes<sup>76,77</sup>. Fewer COG K genes in the proximal colonic microbiome of synbiotic treated pigs might imply a relative reduction in bacteria with the capacity for rapid multiplication and appropriate response to new environmental stimuli. The COG J category of genes codes for the ribosomal L22 protein which is a structural component of the 50S ribosomal subunit where it binds with multiple domains in the 23S rRNA and facilitates ribosomal assembly<sup>78</sup>. The L22 protein forms part of the peptide exit tunnel of bacterial ribosomes and is therefore thought to regulate translation<sup>79</sup>. The protein also interacts with and mediates susceptibility to macrolides which work by inhibiting translation<sup>80</sup>. COG G is a family of genes that encodes extracellular solute-binding proteins which facilitate bacterial uptake and utilization of carbohydrates<sup>81,82</sup>. Their

lower representation could mean a reduced ability to utilize carbohydrates in their environment including the prebiotic studied here<sup>83</sup>. It is worth noting that some of the significantly increased genes within the control groups, originated from alkaliphilic bacteria (*Indibacter alkaliphilus*, *Halanaerobium hydrogeniformans*, *Pae-nibacillus daejeonensis*). A possible reason for this might be that relative increase of these genes in the control group microbiomes could be a result of an actual reduction of these genes in the microbiomes of the treatment group and not necessarily an increase in the controls. This is due to the acidic conditions created by the lactic acid producing bacteria in the synbiotic and the relative proliferation of lactobacillus producing bacteria that is stimulated in the treatment group. Overall, the mechanistic association between the synbiotic and the significant differentially expressed genes between the groups is unclear. However, the pattern of differential distribution suggests a shift of functional capacity away from stress tolerance within the synbiotic fed group. This reduced capacity/need for stress tolerance, in turn, could indicate a stabilizing effect of the synbiotic on the environment conditions within the proximal colons of the treated pigs.

A major limitation of this study was the small sample size within study groups. Nonetheless, we hope that when interpreted within these limits, the study provides insights into this previously unexplored microbial environment of an unstudied Korean native breed of pigs.

## Conclusion

In this study, we used metagenomic techniques to study the impact of a synbiotic on the proximal colonic microbiomes of Korean native, Jeju black pig gilts. Our results show that the synbiotic, containing *Lactobacillus buchmeri* NLRI-1201, *L. plantarum* NLRI-101, *L. casei* DK128 & fructo-oligosaccharide does not cause significant shifts in the microbial abundances, evenness, and diversity at community level. However, the synbiotic increases abundance of members of the family, Lactobacillaceae which have been found to offer health promoting effects to their hosts. Further functional analysis uncovered an effect of the synbiotic on genes belonging to 12 COGs within the colonic microbiome of the Jeju black pigs implying changes in functional capacity of the colonic microbiome. This work improves our understanding of the effect of synbiotics on the microbiome of rare pig breeds such as the Korean black pigs that are native to Jeju Island. We can also infer an understanding of the synbiotics' potential to promote productivity and industrialization of Korean black pigs.

## Data availability

Samples are available from Sequence Read Archive (SRA) with the Bioproject accession number PRJNA815969 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA815969>).

Received: 18 March 2022; Accepted: 12 August 2022

Published online: 15 September 2022

## References

1. Cho, I.-C. *et al.* The robust phylogeny of Korean wild boar (*Sus scrofa coreanus*) using partial D-loop sequence of mtDNA. *Mol. Cells* **28**, 423–430 (2009).
2. Ko, K., Kim, K. & Huh, J. Variations of wind speed in time on Jeju Island, Korea. *Energy* **35**, 3381–3387 (2010).
3. Cho, I. *et al.* QTL analysis of white blood cell, platelet and red blood cell-related traits in an F2 intercross between Landrace and Korean native pigs. *Anim. Genet.* **42**, 621–626 (2011).
4. Sodhi, S. S. *et al.* Comparative transcriptomic analysis by RNA-seq to discern differential expression of genes in liver and muscle tissues of adult Berkshire and Jeju Native Pig. *Gene* **546**, 233–242 (2014).
5. Oliveira, G. & González-Molero, I. An update on probiotics, prebiotics and symbiotics in clinical nutrition. *Endocrinol. Nutr.* **63**, 482–494 (2016).
6. Hill, C. *et al.* The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat. Rev. Gastroenterol. Hepatol.* **11**, 506–514 (2014).
7. Gibson, G. R. & Roberfroid, M. B. Dietary modulation of the human colonic microbiota: Introducing the concept of prebiotics. *J. Nutr.* **125**, 1401–1412 (1995).
8. Yang, F., Hou, C., Zeng, X. & Qiao, S. The use of lactic acid bacteria as a probiotic in swine diets. *Pathogens* **4**, 34–45 (2015).
9. Valeriano, V., Balolong, M. & Kang, D. K. Probiotic roles of *Lactobacillus* sp. in swine: Insights from gut microbiota. *J. Appl. Microbiol.* **122**, 554–567 (2017).
10. Kenny, M., Smidt, H., Mengheri, E. & Miller, B. Probiotics—do they have a role in the pig industry?. *Animal* **5**, 462–470 (2011).
11. Leser, T. D. *et al.* Culture-independent analysis of gut bacteria: The pig gastrointestinal tract microbiota revisited. *Appl. Environ. Microbiol.* **68**, 673–690 (2002).
12. Konstantinov, S. R. *et al.* Feeding of *Lactobacillus sobrius* reduces *Escherichia coli* F4 levels in the gut and promotes growth of infected piglets. *FEMS Microbiol. Ecol.* **66**, 599–607 (2008).
13. Alakomi, H.-L. *et al.* Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane. *Appl. Environ. Microbiol.* **66**, 2001–2005 (2000).
14. Cotter, P. D., Hill, C. & Ross, R. P. Bacteriocins: Developing innate immunity for food. *Nat. Rev. Microbiol.* **3**, 777–788 (2005).
15. Wang, A., Yu, H., Gao, X., Li, X. & Qiao, S. Influence of *Lactobacillus fermentum* I5007 on the intestinal and systemic immune responses of healthy and *E. coli* challenged piglets. *Antonie Van Leeuwenhoek* **96**, 89–98 (2009).
16. Kim, E.-Y. *et al.* Selection of *Lactobacillus* sp. PSC101 that produces active dietary enzymes such as amylase, lipase, phytase and protease in pigs. *J. Gen. Appl. Microbiol.* **53**, 111–117 (2007).
17. Markowiak, P. & Śliżewska, K. The role of probiotics, prebiotics and synbiotics in animal nutrition. *Gut Pathog.* **10**, 1–20 (2018).
18. Bomba, A. *et al.* Improvement of the probiotic effect of micro-organisms by their combination with maltodextrins, fructo-oligosaccharides and polyunsaturated fatty acids. *Br. J. Nutr.* **88**, S95–S99 (2002).
19. Mosenthin, R. Physiology of small and large intestine of swine: Review. *Asian Australas. J. Anim. Sci.* **11**, 608–619 (1998).
20. Wang, H., Shen, J., Pi, Y., Gao, K. & Zhu, W. Low-protein diets supplemented with casein hydrolysate favor the microbiota and enhance the mucosal humoral immunity in the colon of pigs. *J. Anim. Sci. Biotechnol.* **10**, 1–13 (2019).
21. Durazzi, F. *et al.* Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. *Sci. Rep.* **11**, 1–10 (2021).
22. Weinstock, G. M. Genomic approaches to studying the human microbiota. *Nature* **489**, 250–256 (2012).

23. Langille, M. G. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).
24. Herlemann, D. P. *et al.* Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* **5**, 1571–1579 (2011).
25. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
26. Bokulich, N. A. *et al.* Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 1–17 (2018).
27. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
28. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
29. Computing, R. R. *A Language and Environment for Statistical Computing* (R Core Team, 2013).
30. Wickham, H. ggplot2. *Wiley Interdiscip. Rev.* **3**, 180–185 (2011).
31. Thorsen, J. *et al.* Large-scale benchmarking reveals false discoveries and count transformation sensitivity in 16S rRNA gene amplicon data analysis methods used in microbiome studies. *Microbiome* **4**, 1–14 (2016).
32. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
33. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
34. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
35. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: A new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
36. Kang, D. D. *et al.* MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).
37. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
38. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
39. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925–1927. <https://doi.org/10.1093/bioinformatics/btz848> (2019).
40. Shen, W. & Ren, H. TaxonKit: A practical and efficient NCBI taxonomy toolkit. *J. Genet. Genom.* **48**, 844–850 (2021).
41. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 1–13 (2019).
42. Feldgarden, M. *et al.* Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob. Agents Chemother.* **63**, e00483-00419 (2019).
43. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* **10**, e65088 (2021).
44. Lu, J., Breitwieser, F. P., Thielen, P. & Salzberg, S. L. Bracken: Estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **3**, e104 (2017).
45. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
46. Steinegger, M., Mirdita, M. & Söding, J. Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nat. Methods* **16**, 603–606 (2019).
47. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
48. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
49. Cantalapiedra, C. P., 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.* **38**, 5825–5829 (2021).
50. Huerta-Cepas, J. *et al.* eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
51. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLoS Biol.* **8**, e1000412 (2010).
52. Yen, S. & Johnson, J. S. Metagenomics: A path to understanding the gut microbiome. *Mammal. Genome* **32**, 282–296 (2021).
53. Ohashi, Y., Tokunaga, M., Taketomo, N. & Ushida, K. Stimulation of indigenous lactobacilli by fermented milk prepared with probiotic bacterium, *Lactobacillus delbrueckii* subsp. *bulgaricus* strain 2038, in the pigs. *J. Nutr. Sci. Vitaminol.* **53**, 82–86 (2007).
54. Doo, E.-H., Schwab, C., Chassard, C. & Lacroix, C. Cumulative effect of yeast extract and fructooligosaccharide supplementation on composition and metabolic activity of elderly colonic microbiota in vitro. *J. Funct. Foods* **52**, 43–53 (2019).
55. Tessler, M. *et al.* Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Sci. Rep.* **7**, 1–14 (2017).
56. Brumfield, K. D., Huq, A., Colwell, R. R., Olds, J. L. & Leddy, M. B. Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. *PLoS ONE* **15**, e0228899 (2020).
57. Pollock, J. *et al.* Resistance to change: AMR gene dynamics on a commercial pig farm with high antimicrobial usage. *Sci. Rep.* **10**, 1–10 (2020).
58. Wang, C. *et al.* Characterization of the pig gut microbiome and antibiotic resistome in industrialized feedlots in China. *Msystems* **4**, e00206-00219 (2019).
59. Finley, R. L. *et al.* The scourge of antibiotic resistance: The important role of the environment. *Clin. Infect. Dis.* **57**, 704–710 (2013).
60. Manos-Turvey, A. *et al.* Inhibition studies of *Mycobacterium tuberculosis* salicylate synthase (MbtI). *ChemMedChem* **5**, 1067–1079 (2010).
61. Lawlor, M. S., O'connor, C. & Miller, V. L. Yersiniabactin is a virulence factor for *Klebsiella pneumoniae* during pulmonary infection. *Infect. Immun.* **75**, 1463–1472 (2007).
62. Sassone-Corsi, M. *et al.* Microcins mediate competition among Enterobacteriaceae in the inflamed gut. *Nature* **540**, 280–283 (2016).
63. Deriu, E. *et al.* Probiotic bacteria reduce salmonella typhimurium intestinal colonization by competing for iron. *Cell Host Microbe* **14**, 26–37 (2013).
64. Skerker, J. M., Prasol, M. S., Perchuk, B. S., Biondi, E. G. & Laub, M. T. Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: A system-level analysis. *PLoS Biol.* **3**, e334 (2005).
65. Wolanin, P. M., Thomason, P. A. & Stock, J. B. Histidine protein kinases: Key signal transducers outside the animal kingdom. *Genome Biol.* **3**, 1–8 (2002).
66. Vigil-Stenman, T., Larsson, J., Nylander, J. A. & Bergman, B. Local hopping mobile DNA implicated in pseudogene formation and reductive evolution in an obligate cyanobacteria-plant symbiosis. *BMC Genom.* **16**, 1–18 (2015).
67. Vigil-Stenman, T., Ininbergs, K., Bergman, B. & Ekman, M. High abundance and expression of transposases in bacteria from the Baltic Sea. *ISME J.* **11**, 2611–2623 (2017).

68. Shang, Y., Kumar, S., Thippareddi, H. & Kim, W. K. Effect of dietary fructooligosaccharide (FOS) supplementation on ileal microbiota in broiler chickens. *Poult. Sci.* **97**, 3622–3634 (2018).
69. Respondek, F. *et al.* Short-chain fructo-oligosaccharides modulate intestinal microbiota and metabolic parameters of humanized gnotobiotic diet induced obesity mice. *PLoS ONE* **8**, e71026 (2013).
70. Xie, N. *et al.* NAD<sup>+</sup> metabolism: Pathophysiological mechanisms and therapeutic potential. *Signal Transduct. Target. Ther.* **5**, 1–37 (2020).
71. Green, J. & Paget, M. S. Bacterial redox sensors. *Nat. Rev. Microbiol.* **2**, 954–966 (2004).
72. Domergue, R. *et al.* Nicotinic acid limitation regulates silencing of *Candida* adhesins during UTI. *Science* **308**, 866–870 (2005).
73. Chen, X., Li, S. & Liu, L. Engineering redox balance through cofactor systems. *Trends Biotechnol.* **32**, 337–343 (2014).
74. Cahová, H., Winz, M.-L., Höfer, K., Nübel, G. & Jäschke, A. NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* **519**, 374–377 (2015).
75. Benedetti, M. S. *Annales Pharmaceutiques Françaises* 45–52 (Elsevier, 2020).
76. Gruber, T. M. & Gross, C. A. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**, 441 (2003).
77. Fang, F. C. Sigma cascades in prokaryotic regulatory networks. *Proc. Natl. Acad. Sci. USA* **102**, 4933–4934 (2005).
78. Wilson, D. N. & Nierhaus, K. H. Ribosomal proteins in the spotlight. *Crit. Rev. Biochem. Mol. Biol.* **40**, 243–267 (2005).
79. Nakatogawa, H. & Ito, K. The ribosomal exit tunnel functions as a discriminating gate. *Cell* **108**, 629–636 (2002).
80. Davydova, N., Streltsov, V., Wilce, M., Liljas, A. & Garber, M. L22 ribosomal protein and effect of its mutation on ribosome resistance to erythromycin. *J. Mol. Biol.* **322**, 635–644 (2002).
81. Tam, R. & Saier, M. H. Jr. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* **57**, 320–346 (1993).
82. Berntsson, R.P.-A., Smits, S. H., Schmitt, L., Slotboom, D.-J. & Poolman, B. A structural classification of substrate-binding proteins. *FEBS Lett.* **584**, 2606–2617 (2010).
83. Yoshida, K. *et al.* Bifidobacterium response to lactulose ingestion in the gut relies on a solute-binding protein-dependent ABC transporter. *Commun. Biol.* **4**, 1–8 (2021).

### Author contributions

Conceptualization, J.J., K.S., H.L., Y.S., D.L., and J.H. Experimentation, M.L., Y.C., K.S., Y.S., D.L., S.L., and Y.-J.S. Data analysis and interpretation, J.J., A.W.B., and J.H. Writing-original draft, J.J., A.W.B., and J.H.

### Funding

This research was supported by grants from the Next-Generation BioGreen 21 Program (PJ01322302, PJ01316701), Rural Development Administration, Republic of Korea.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-18503-2>.

**Correspondence** and requests for materials should be addressed to J.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022