scientific data

Check for updates

OPEN

DATA DESCRIPTOR

Assembly and maturation of calf gut microbiome from neonate to post-puberty

Hyun Sik Kim[1,4], Jae-Yun Lee[2,4], Tae Woong Whon[3] ✉ & Jin-Woo Bae[1,2] ✉

With the help of rumen bacteria, ruminants can feed on indigestible plant materials and produce over 70% of their energy as fatty acids. However, during lactation, ruminants exhibit characteristics of monogastric animals due to an undeveloped rumen; therefore, understanding gut microbiome changes in growing calves is essential. Our understanding of the gut microbiome in growing calves remains limited in large populations with the same diet, breed, and period. Here, we describe 16S rRNA gene amplicon sequencing data from 420 faecal samples, 20 rumen contents, 17 small intestine contents, and 18 large intestine contents collected from 57 healthy, antibiotic-free Korean beef cattle from neonatal to post-pubertal age. Eight 16S rRNA gene amplicon datasets from the host diet samples were obtained. Approximately 148 million raw reads, averaging 153,352 ± 96,050 (mean ± SD) reads per sample, and 51,596 unique amplicon sequence variants (381–368 per sample) were identified in the 483 samples. These shareable datasets can be reused by researchers to assess gut microbiome-related functions in growing calves and improve ruminant production and health.

Background & Summary

The beef industry has made significant strides in animal facilities, feeding, welfare, breeding, herd management, and biopharmaceutical use. In addition to being a major source of protein for humans, cattle are economically important animals that efficiently convert plant biomass into nutrients that can be absorbed and utilised by animals to provide milk and meat for humans and play an important role in harnessing solar energy. Bacteria are the predominant microbes in the rumen (>91% of the whole microbiome)[1], and adult ruminants fed indigestible plant materials are larg™ely dependent on their ruminal microbiome for feed digestion[2,3]. Volatile fatty acids and proteins produced by ruminal bacteria provide more than 70% of the required energy and 60% of the non-ammonia nitrogen to dairy cows[4]. A lactating calf has the characteristics of a monogastric animal with an underdeveloped rumen because the only source of energy is the mother's milk. Therefore, the initial assembly and subsequent maturation of the microbiome in the small and large intestines are crucial for nutrient absorption, disease susceptibility, and further growth and development, similar to monogastric animals[5,6]. Given the impact of microbes residing in the gut on host immunity, nutrition, metabolism, and physiology in mammalian infants[7,8], it is essential to understand the microbial signatures in the intestinal tract of calves during the neonatal and growing periods.

In our previous studies, we aimed to determine the effects of hypogonadism following male castration on the gut microbiota, serum metabolites, and host metabolic phenotype and the contribution of the altered microbiota to increased adiposity in hypogonadal cattle models (Korean brown cattle and Holstein Friesian cattle). We revealed that the recovery of the normal gut microbiome assembly through faecal microbiota transplantation, whereby the faeces of healthy calves were injected into diarrhoeal calves, was more effective than traditional antibiotic treatment in neonatal calves. This advanced approach dramatically reduces mortality and contributes to increased farm income. Therefore, understanding the normal gut microbiome assembly from the neonatal calf stage to the developmental stages in cattle is of great importance.

In this report, we present V3–V4 16S ribosomal RNA (rRNA) gene amplicon sequencing data derived from 420 longitudinal faecal samples (256 males and 164 females) from 51 or 57 healthy cattle (Korean brown cattle

[1]Department of Biology, Kyung Hee University, Seoul, 02447, Republic of Korea. [2]Department of Life and Nanopharmaceutical Sciences, Kyung Hee University, Seoul, 02447, Republic of Korea. [3]Kimchi Functionality Research Group, World Institute of Kimchi, Gwangju, 61755, Republic of Korea. [4]These authors contributed equally: Hyun Sik Kim, Jae-Yun Lee. ✉e-mail: twwhon@gmail.com; baejw@khu.ac.kr

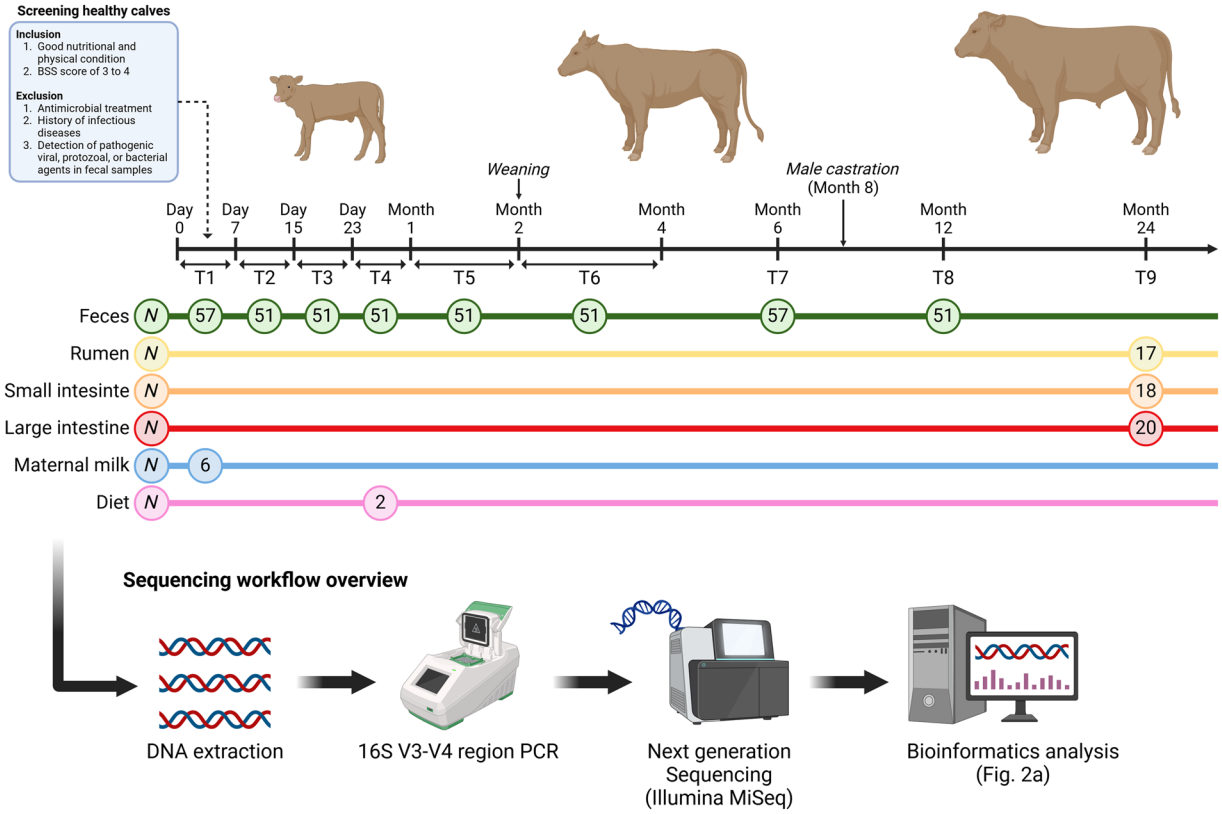

Fig. 1 Overview of the sampling process. Faecal samples of the calves (256 males and 164 females) were sampled longitudinally from birth to 12 months. Sampling timepoints are 1–7 d (T1), 8–15 d (T2), 16–23 d (T3), 24–30 d (T4), 1–2 months (T5), 2–4 months (T6), 6 months (T7), and 12 months of age (T8). On the 24th month (T9), calves were euthanized and rumen and small and large intestines were collected. Maternal milk samples from the six mothers were collected at T1, and the two diet samples were collected at T4. All calves were selected by screening criteria (see Methods for details) on T1 and weaned at 2 months of age. Male calves were castrated at 8 months. The numbers in coloured circles indicate the number of collected samples (*N*) at each timepoint. Collected samples were sequenced according to the illustrated workflow (bottom). BSS, Bristol stool scale; 16S V3-V4, V3-V4 hypervariable region of 16S rRNA gene. This figure is created in BioRender: Bae, L. (2025) https://BioRender.com/q70f878.

Bos taurus coreanae, the most common breed in South Korea) from neonates to post-puberty. The sampling timepoints corresponded to the main developmental stages: 1–7 d (T1), 8–15 d (T2), 16–23 d (T3), 24–30 d (T4), 1–2 months (T5), 2–4 months (T6), 6 months (T7), and 12 months (T8) of age. Additionally, at 24 months (T9), 20 cattle (10 males and 10 females) were randomly selected, and the contents of the rumen, large intestine, and small intestine were collected. Given the crucial role of the host diet in gut microbiome assembly[9,10], we collected maternal milk ($N = 6$) and diet ($N = 2$) samples, which represent the host diet at specific developmental stages at T1 and T4, respectively. The collected samples underwent DNA extraction, V3–V4 16S rRNA gene amplification by PCR, sequencing using the Illumina MiSeq platform, and bioinformatics pipeline analysis. A schematic diagram of the current study is illustrated in Fig. 1, and the detailed sampling method is described in the Methods section.

The data analysis was conducted as depicted in Fig. 2a. The quality of the raw reads was assessed using FastQC and MultiQC[11] (Fig. 2b). After quality assessment, raw reads were processed and analysed according to our workflow (see the Methods section for a detailed description). Faeces (38,932 ± 13,308 [mean ± SD] total amplicon sequence variants [ASVs]), large intestine (54,914 ± 11,782), small intestine (64,821 ± 13,402), rumen (53,864 ± 25,866), and the milk (69,441 ± 16,285) showed similar counts of total ASVs, but the diet (6,272 ± 448) showed relatively low counts of total ASVs (Fig. 2c). The depths along the timepoints and sexes were also depicted in Fig. 2d,e. Alpha rarefaction curve analysis indicated that the sequencing depth of the samples was sufficient to capture most of the diversity present in the dataset (Fig. 2f–h).

Our data encompass a wide range of bacterial taxonomy (Fig. 3). For instance, we present the top 20 most prevalent genera along with their higher taxonomic ranks. Genera belonging to the phyla Bacteroidota and Firmicutes, which are commonly found in intestinal sample of animals, were identified. Additionally, *Escherichia-Shigella*, a genus belonging to the phylum Proteobacteria, that was previously reported in our study to be associated with calf diarrhea[12], was also observed.

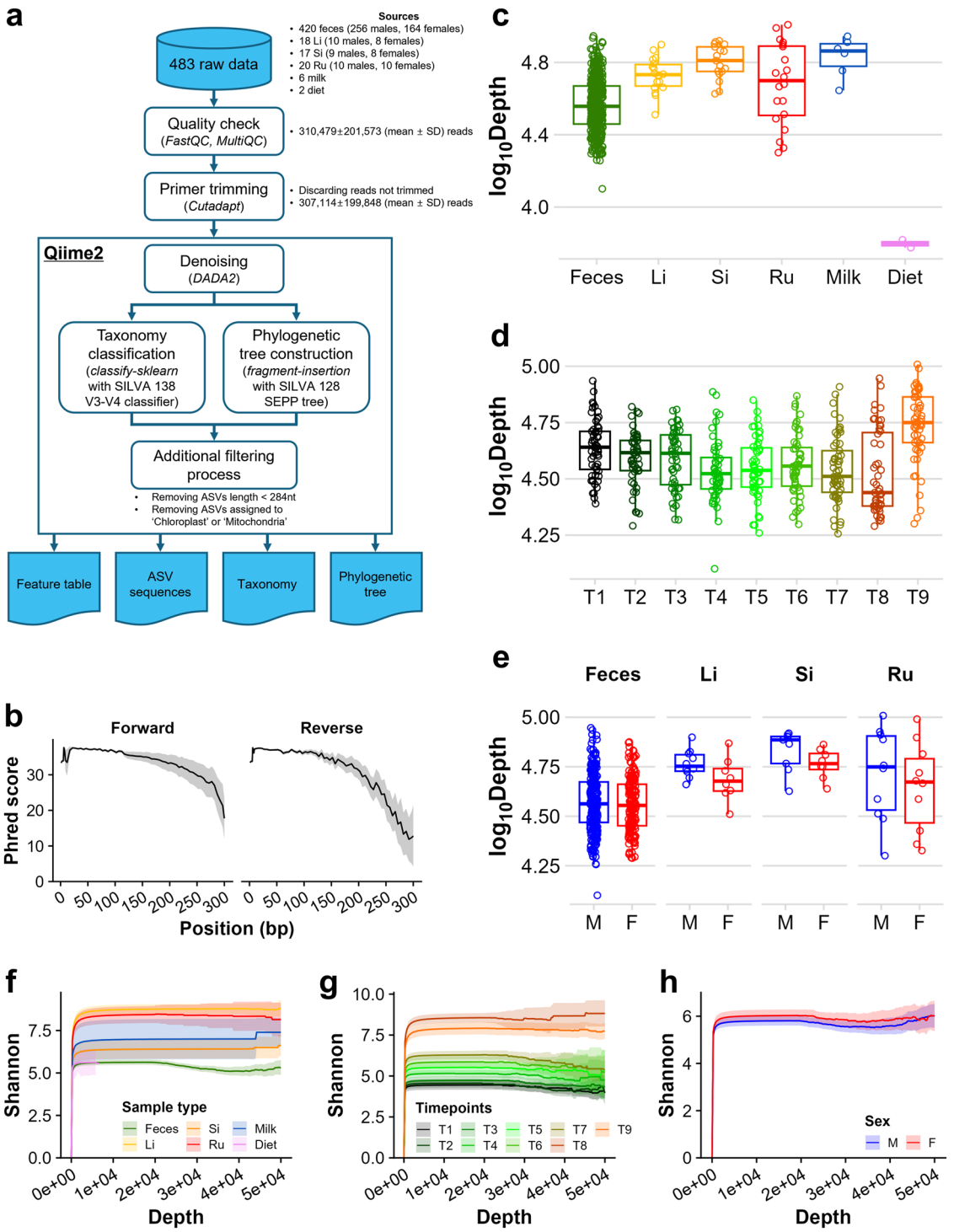

Fig. 2 Assessment of the data quality and analysis workflow. (**a**) Schematic diagrams for 16S amplicon data processing and analysis workflow. The quality of 483 raw sequencing data was assessed and checked for their eligibility for analysis. The sequences of V3–V4 PCR primers were trimmed, and the processed data were imported into the Qiime2 platform for further analysis. Detailed methods are described in the Methods section. (**b**) Quality plot of the 483 raw sequencing data. (**c**–**e**) Depth of each sample after DADA2 denoising, visualised according to (**c**) sample type, (**d**) timepoints, and (**e**) sample type divided by sex. (**f**–**h**) Rarefaction curve of the samples visualised by (**f**) sample type, (**g**) timepoints, and (**h**) sex.

Altogether, we extended the collection of resources related to the development of the gut microbiome in cattle. The datasets described here provide an understanding of the relation between development and the gut microbiome and the characterisation of the normal assembly of poorly understood commensal bacteria

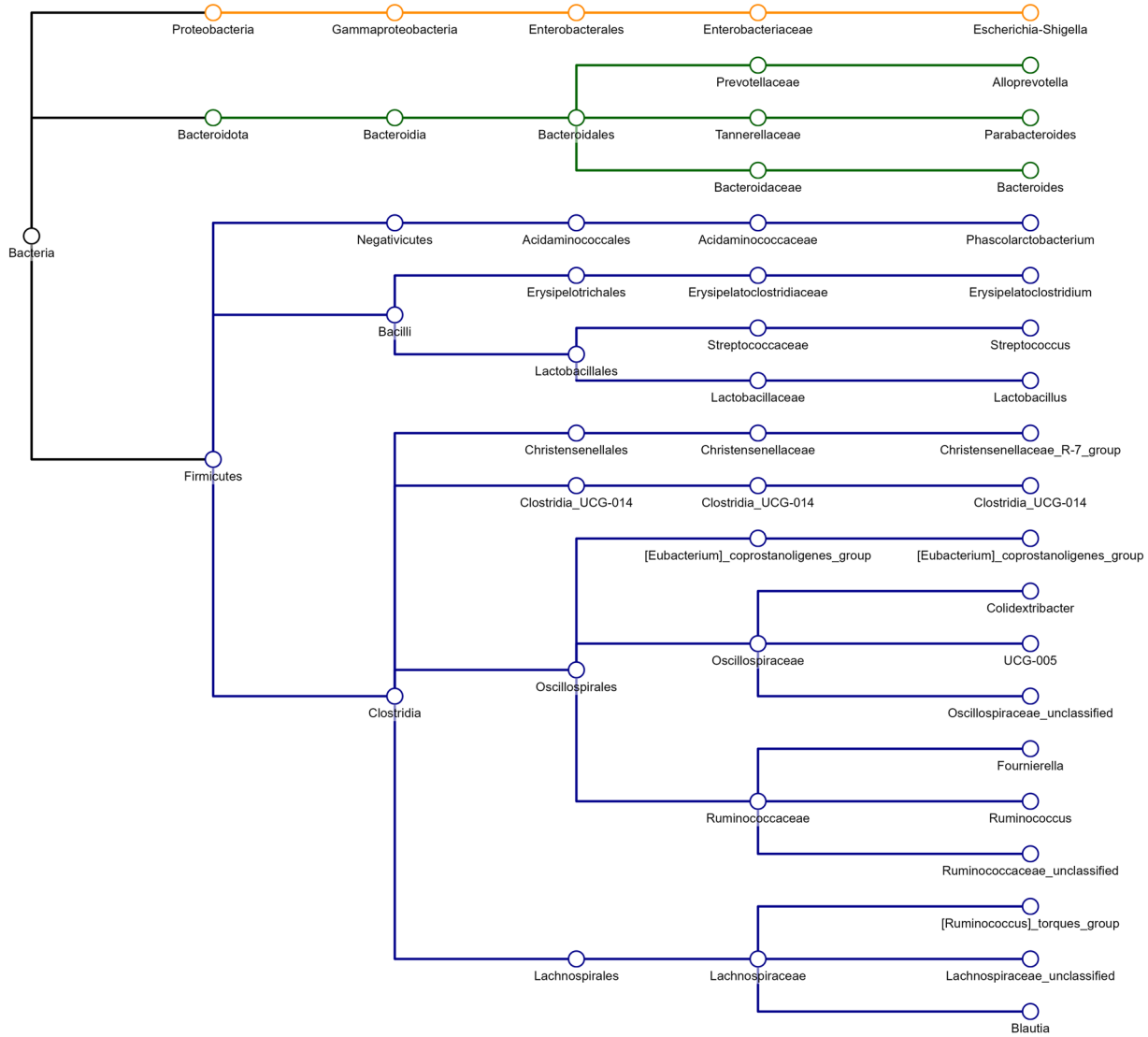

Fig. 3 Top 20 prevalent genera and their taxonomy. From left to right, each circle represents the taxonomy of the top 20 prevalent genera in the order of kingdom, phylum, class, order, family, and genus. The color of the lines following the phylum level indicates the corresponding phylum (orange, Proteobacteria; green, Bacteroidota; blue, Firmicutes). Taxonomic assignment was performed based on the SILVA 138 SSURef NR99 database.

associated with cattle. These data are expected to pave the way for the development of breeding and therapeutic strategies to improve productivity and disease resistance in cattle.

Methods

A substantial portion of the method sections were similar to our previously peer-reviewed study[6] because the experiments were conducted using the same methodology.

Ethics approval. The study protocol was approved by the Institutional Animal Care and Use Committee of Kyung Hee University (KHUASP[SE]-17-026, KHUASP[SE]-17-028, and KHUASP[SE]-17-145) for the cattle study. The experiments were performed in accordance with the ARRIVE guidelines.

Animals. All the cows used in our study were brown Korean cattle (*Bos taurus coreanae*), the most common breed in South Korea. They had free access to food and water, and their mothers nurtured them in individual barns. Before weaning, which was performed 60 days after birth, each calf was housed with its mother. After weaning, the calves were separated from their mothers and housed separately. Although not all calves were kept in the same space, a physical area was created that could be accessed by calves, allowing them to eat their feed and interact with other calves. After weaning, the calves were randomly mixed and housed in stalls containing five calves each. All individuals were housed in stalls and provided with an identical diet at each growth stage to minimize inter-individual variation in intestinal microbiota. All the barns were divided into 3 m × 3 m spaces.

The floor was kept dry, and individual buckets and feed bins were daily cleaned throughout the study. At 8 months of age, the male cattle were orchiectomized by a veterinarian (both testicles were removed). Calves that had been previously administered antimicrobial agents or had a history of infectious diseases were excluded.

Screening for the selection of normal calves. Calves with similar birth dates were selected based on their nutritional status and physical conditions, such as clean nose, ears, mouth, tail, hair, or rump. Calves that had undergone antimicrobial treatment (antibiotics, antifungals, or antivirals) or had a history of infectious diseases (including anthrax, foot-and-mouth disease, contagious bovine pleuropneumonia, or brucellosis) were excluded from the study.

Calves with a Bristol stool scale score of 3 to 4 for faeces collected by rectal enema were considered healthy. A diagnostic multiplex PCR assay was employed to test for specific pathogens in faecal samples, and calves that were positive for any of the following were also excluded: causative viral (group A, B, and C rotaviruses, bovine enteric Nebraska-like calicivirus, bovine coronavirus, bovine norovirus, bovine nebovirus, bovine torovirus, or bovine viral diarrhoea virus), protozoal (*Eimeria zuernii*), or bacterial agents (*Salmonella enterica*, *Salmonella enterica* serovar Typhimurium, *Clostridium perfringens*, shigatoxigenic *Escherichia coli*, enterotoxigenic *Escherichia coli*, or enterohemorrhagic *Escherichia coli*). According to the screening criteria, a total of fifty-seven calves were selected for the current dataset.

Sample collection. Faecal samples were collected using a rectal enema, while the collector wore clean disposable latex gloves. Fifty-one or 57 faecal samples were collected at the following timepoint: 1–7 d (T1), 8–15 d (T2), 16–23 d (T3), 24–30 d (T4), 1–2 months (T5), 2–4 months (T6), 6 months (T7), and 12 months (T8). At 24 months of age (T9), the contents of the small and large intestines and rumen were measured. To obtain additional samples from the calf environment, we also collected feed pellets ($N = 2$) and maternal milk ($N = 6$) from the mothers of randomly selected calves. To prevent the contamination of milk samples with environmental microbes, the udders and teats of the cows were wiped with cotton wool soaked in 70% ethanol, and the initial few streams of milk were discarded before sample collection. The luminal contents of the rumen, ileum (small intestine), and colon (large intestine) were collected from 24-month-old adult cattle at a local slaughterhouse (Gunwi-Gun, South Korea), under the supervision of an official veterinarian. The collected samples were transported to the laboratory on dry ice and stored at −80 °C until use. A schematic of the sampling process is shown in Fig. 1.

DNA extraction and V3–V4 16S rRNA gene amplicon sequencing. Bacterial genomic DNA was extracted from faeces, maternal milk, and diet using a Repeated Bead-Beating plus column (QIAamp DNA Stool Mini Kit; Qiagen). For Illumina MiSeq sequencing, a hypervariable V3–V4 region of 16S rRNA gene was amplified by PCR using 338 F forward primer (5′-TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT ACG GGN GGC WGC AG-3′) and 805 R reverse primer (5′-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGA CTA CHV GGG TAT CTA ATC C-3′) using the C1000 Thermal Cycler (Bio-Rad). The PCR conditions were as follows: initial denaturation at 95 °C for 3 min; followed by 23 cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 30 s; and then a final extension step at 72 °C for 5 min. Products from three individual tubes using the same DNA template were pooled together after PCR and purified using a QIAquick PCR Purification Kit (Qiagen). The amplicon library was prepared using Nextera XT Index (Illumina) and then sequenced through the Illumina MiSeq platform (2 × 300 bp), according to the manufacturer's instructions.

Bioinformatics analysis. The quality of the raw data (310,479 ± 201,573 [mean ± SD] reads) was assessed by FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and MultiQC[11]. The 338 F and 805 R PCR primer sequences in the 5′ end of the raw reads were trimmed using Cutadapt (v4.4)[13] with the '–discard-untrimmed' flag. Trimmed amplicon data were imported into the Qiime2 (v2023.5) platform[14] and subjected to subsequent analyses. The DADA2 pipeline (v1.26.0)[15] was implemented to denoise sequences into ASVs with the "–p-trunc-len-f 255–p-trunc-len-r 195–p-max-ee-f 3–p-max-ee-r 3" flags. For taxonomic classification, we processed SILVA 138 SSURef NR99 DB[16] data and the RESCRIPt plugin[17] and built a classifier using the q2-feature-classifier plugin[18]. A phylogenetic tree was constructed using the SATé-enabled phylogenetic placement[19] method based on the SILVA 128 SATé-enabled phylogenetic placement reference tree using the q2-fragment-insertion plugin[20]. ASVs whose length was below 284 nt or assigned to 'Chloroplast' or 'Mitochondria' were discarded and not used for the downstream analysis.

Visualisation. All graphical figures, except Figs. 1, 2a, were constructed using the R packages ggplot2, cowplot (v1.1.2), ggalluvial (v0.12.5), ggrepel (v0.9.4), ggtext (v0.1.2) and tidygraph (v1.3.1). Data were imported into R and processed using the R packages phyloseq (v1.46.0)[21] and microbiome (v1.24.0)[22] for visualisation. Figure 1 was created using BioRender.

Data Records

Raw sequencing data. The raw sequences of the 16S rRNA genes obtained from the intestinal luminal contents and faecal, milk, and feed samples of cattle were deposited in the NCBI Sequence Read Archive (SRA) under the accession number SRP520455 (http://identifiers.org/insdc.sra:SRP520455)[23].

Sample metadata. Metadata for all 483 sequencing samples is recorded in Supplementary Table 1. It is also deposited under the file name 'Metadata.tsv' in the Zenodo (https://doi.org/10.5281/zenodo.14878589)[24].

The metadata consists of five columns:

1. 'sample_name': The unique identifier for each sample used for the current dataset. These identifiers match the raw sequencing data deposited in the NCBI SRA under accession number SRP520455 (http://identifiers.org/insdc.sra:SRP520455)[23].
2. 'calf_id': The unique identifier for each calf from which the sample was collected. Since multiple samples were collected from the same calf at different time points, the same 'calf_id' appears multiple times.
3. 'timepoints': Indicates the sampling timepoints, range from T1 to T9. The details of each time point are illustrated in Fig. 1.
4. 'sample_origin': Represents the sample type, categorized as *Feces, Large_intestine, Small_intestine, Rumen, Diet*, and *Milk*.
5. 'host_sex': Indicates the sex of the corresponding calf.

Ready-to-use processed data. The following are the final processed data, prepared according to the workflow in Fig. 2a and ready for direct analysis. The data have been deposited to Zenodo (https://doi.org/10.5281/zenodo.14878589)[24].

- 'ASV_table.tsv': ASV abundance table. The columns are sample names, and the rows are ASV names.
- 'Representative_ASV_sequences.fasta': Representative sequences of each ASV in FASTA format.
- 'Phylogenetic_tree.nwk': Newick format phylogenetic tree of ASVs, which can be used for calculating UniFrac distance in beta-diversity analysis.
- 'Taxonomy.tsv': Taxonomic annotation of ASVs. Each column represents taxonomy rank of corresponding ASV.

Technical Validation

A substantial portion of the '*Assessment of potential DNA contaminants*' section were similar to our previously peer-reviewed study[6] because the experiments were conducted using the same methodology.

Assessment of potential DNA contaminants. All reagents used for DNA extraction were conducted PCR for determining possible contaminants. The PCR of the reagents using the 338 F and 805 R primers (30-cycle reaction) revealed no apparent contamination. Additionally, V3–V4 amplicons from two vials of standard mock community (ZymoBIOMICS™ Microbial Community Standard; Zymo), DNA extraction controls, and negative PCR controls (PCR products generated from a template obtained through a sham extraction, in which no fecal sample was added) were included in each sequencing run for quality assurance. Taxonomy of the ASVs from standard mock community was assigned by aligning representative ASV sequences to SILVA 123 QIIME-compatible database.

Sequence quality assessment and bioinformatics pipeline. The sequencing quality of the raw data was evaluated using the FastQC software. FastQC reports were merged using MultiQC[11] and checked for any problems with sequencing quality (Fig. 2b). Most reads had quality scores higher than Q20. The DADA2 pipeline generates an error model from sequencing quality information and performs quality filtering independently[15]. Therefore, quality filtering prior to using the DADA2 pipeline is not recommended. Thus, we removed the PCR primer sequences using Cutadapt[13] with a discard-untrimmed flag to remove reads which did not possess the proper primer sequences. ASVs shorter than 284 nt (2/3 length of target V3–V4 region) and assigned to 'Chloroplast' or 'Mitochondria' were discarded to ensure the quality of the data.

Usage Notes

The ASV abundance table ('ASV_table.tsv'), representative sequences ('Representative_ASV_sequences.fasta'), phylogenetic tree ('Phylogenetic_tree.nwk'), and taxonomy ('Taxonomy.tsv'), generated using our analysis pipeline and the sample metadata ('Metadata.tsv') (see Methods and Data Records sections for detailed explanation) were uploaded to Zenodo (https://doi.org/10.5281/zenodo.14878589)[24]. These data have already been processed into an analysis-ready format and can therefore be directly imported into analysis programs, such as Qiime2 or R. If researchers plan to integrate their data with other datasets for specific purposes (including meta-analysis), we recommend standardising the analysis pipeline to reduce analytical variation owing to different analysis methods and parameters. For this purpose, researchers can download our raw data from the NCBI SRA with the accession number SRP520455 (http://identifiers.org/insdc.sra:SRP520455)[23].

Code availability

All the custom codes used for the current study are available at GitHub (https://github.com/jylee3247/healthy-calf-microbiome-assembly) and Zenodo (https://doi.org/10.5281/zenodo.14878589)[24]. Additionally, all analysis-ready format data, including the ASV abundance table, ASV sequences, ASV taxonomy, ASV phylogenetic tree, and sample metadata, are available.

Received: 29 July 2024; Accepted: 19 February 2025;
Published online: 04 March 2025

References

1. Brulc, J. M. *et al*. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc Natl Acad Sci USA* **106**, 1948–1953 (2009).
2. Flint, H. J., Bayer, E. A., Rincon, M. T., Lamed, R. & White, B. A. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol* **6**, 121–131 (2008).

3. Jami, E., White, B. A. & Mizrahi, I. Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. *PLoS One* **9**, e85423 (2014).
4. Clark, J. H., Klusmeyer, T. H. & Cameron, M. R. Microbial protein synthesis and flows of nitrogen fractions to the duodenum of dairy cows. *J Dairy Sci* **75**, 2304–2323 (1992).
5. Du, Y. *et al*. Colonization and development of the gut microbiome in calves. *J Anim Sci Biotechnol* **14**, 46 (2023).
6. Kim, H. S. *et al*. Longitudinal evaluation of fecal microbiota transplantation for ameliorating calf diarrhea and improving growth performance. *Nat Commun* **12**, 161 (2021).
7. Tamburini, S., Shen, N., Wu, H. C. & Clemente, J. C. The microbiome in early life: implications for health outcomes. *Nat Med* **22**, 713–722 (2016).
8. Ferretti, P. *et al*. Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. *Cell Host Microbe* **24**, 133–145 e135 (2018).
9. Backhed, F. *et al*. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**, 690–703 (2015).
10. Stewart, C. J. *et al*. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583–588 (2018).
11. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
12. Whon, T. W. *et al*. Calf Diarrhea Caused by Prolonged Expansion of Autochthonous Gut Enterobacteriaceae and Their Lytic Bacteriophages. *Msystems* **6** (2021).
13. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, 3 (2011).
14. Bolyen, E. *et al*. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2 *Nat Biotechnol* **37**, 1091–1091 (2019).
15. Callahan, B. J. *et al*. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581-+ (2016).
16. Quast, C. *et al*. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590–D596 (2013).
17. Robeson, M. I. I. *et al*. RESCRIPt: Reproducible sequence taxonomy reference database management. *PLOS Comput Biol* **17** (2021).
18. Bokulich, N. A. *et al*. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6** (2018).
19. Mirarab, S., Nguyen, N. & Warnow, T. SEPP: SATe-Enabled Phylogenetic Placement. *Biocomput-Pac Sym*, 247–258 (2012).
20. Janssen, S. *et al*. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. *Msystems* **3** (2018).
21. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217 (2013).
22. Lahti, L. & Shetty, S. Introduction to the microbiome R package. https://microbiome.github.io/tutorials (2018).
23. *NCBI Sequence Read Archive* http://identifiers.org/insdc.sra:SRP520455 (2024).
24. Kim, H. S., Lee, J. Y., Whon, T. W. & Bae, J. W. jylee3247/healthy-calf-microbiome-assembly: healthy-calf-microbiome-assembly. *Zenodo* https://doi.org/10.5281/zenodo.14878589 (2025).

Acknowledgements

This study was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. RS-2021-NR061913 to H.S.K.), the World Institute of Kimchi (No. KE2501-1-1 to T.W.W.), Mid-Career Researcher Program (No. 2020R1A2C3012797 to J.-W.B.) through the NRF, grant (No. 22213MFDS537 to J.-W.B.) from the Ministry of Food and Drug Safety in 2022, and the NRF funded by the Korean government (MSIT) (No. RS-2023-00227274 to J.-W.B.). We would like to thank Editage (www.editage.co.kr) for English language editing.

Author contributions

H.S.K. and J.Y.L. designed the experiments. H.S.K., J.Y.L. and T.W.W. performed most of the experiments and analysed the data. H.S.K. and J.Y.L. wrote the manuscript. T.W.W. and J.W.B. critically reviewed the manuscript.

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/s41597-025-04677-7.

Correspondence and requests for materials should be addressed to T.W.W. or J.-W.B.