BIODIVERSITAS
Volume 26, Number 9, September 2025
Pages: 4522-4533

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d260924

Gut microbiota composition and variation in Baduy infants living traditional lifestyles in Banten, Indonesia

DERICO HITIPEUW[1,2], RAISHA NURANINDITA[2,3], BAGOES WIDJANARKO[4], FAUZI MUH[2,3,♥]

[1]Department of Environmental Health, Faculty of Public Health, Universitas Diponegoro. Jl. Prof. Jacub Rais, Semarang 50275, Central Java, Indonesia
[2]Department of Epidemiology and Tropical Diseases, Faculty of Public Health, Universitas Diponegoro. Jl. Prof. Jacub Rais, Semarang 50275, Central Java, Indonesia. Tel.: +62-24-7460044, ♥email: fauzimuh010@gmail.com
[3]Public Health Genomics Lab, Faculty of Public Health, Universitas Diponegoro. Jl. Prof. Jacub Rais, Semarang 50275, Central Java, Indonesia
[4]Department of Health Promotion and Health Behavior, Faculty of Public Health, Universitas Diponegoro. Jl. Prof. Jacub Rais, Semarang 50275, Central Java, Indonesia

Manuscript received: 3 June 2025. Revision accepted: 20 September 2025.

Abstract. *Hitipeuw D, Nuranindita R, Widjanarko B, Muh F. 2025. Gut microbiota composition and variation in Baduy infants living traditional lifestyles in Banten, Indonesia. Biodiversitas 26: 4522-4533.* Early-life gut microbiota development shapes immediate and long-term health. Studying infants in traditional, pre-industrial populations provides insights into natural microbial colonization, yet data on microbiota assembly in minimally medically exposed indigenous communities remain scarce. This study aimed to characterize the gut microbiota profiles of infants from the Baduy community of Indonesia, an indigenous population that maintains traditional practices with minimal exposure to industrialization. Fecal samples were collected from two Baduy infants aged 12-13 months (one male and one female) living under traditional conditions. Total genomic DNA was extracted using standardized protocols, and full-length 16S rRNA gene sequencing (V1-V9 regions, 1484 bp) was performed using Oxford Nanopore technology to achieve species-level taxonomic resolution. Alpha diversity metrics, including the Shannon diversity index, Simpson index, and species richness estimators (Chao1 and ACE), were calculated to assess within-sample microbial diversity. Female infants (13BL_P) exhibited higher Shannon diversity (3.44 vs 3.03) and were dominated by beneficial taxa, including *Faecalibacterium* (Firmicutes), *Bifidobacterium* (Actinobacteria), and *Anaerobutyricum* (Firmicutes). Conversely, the male infant (12BL_L) demonstrated higher species richness (543 vs. 511 observed species) but was dominated by potentially pathogenic genera, including *Enterococcus* and *Streptococcus* (both Firmicutes), alongside *Bifidobacterium*. Preliminary findings indicate significant microbiota variability within culturally similar traditional populations, suggesting that host factors affect microbial colonization. This study provides baseline microbiota data for the Baduy population and offers frameworks for investigating gut diversity in indigenous communities.

Keywords: 16S rRNA gene, Baduy community, indigenous populations, infant gut microbiota, traditional lifestyles

INTRODUCTION

The human gut microbiota undergoes critical assembly during the first 1000 days of life, establishing microbial communities that influence immune system maturation, metabolic programming, and lifelong health trajectories (Bäckhed et al. 2015; Robertson et al. 2019). Early colonization patterns determine susceptibility to chronic diseases, including obesity, allergies, inflammatory bowel disease, and neurodevelopmental disorders, making infant microbiota development a fundamental determinant of human health (Tamburini et al. 2016; Stiemsma and Michels 2018).

Infant gut development follows distinct temporal phases: an initial colonization phase (0-3 months) with rapid bacterial establishment, an expansion phase (3-14 months) with increasing diversity, and eventual stabilization toward adult-like communities by 2-3 years (Stewart et al. 2018). However, trajectories vary across populations, with traditional communities showing assembly patterns distinct from industrialized societies (Clemente et al. 2015; Smits et al. 2017; Stewart et al. 2018).

Traditional populations harbor exceptional gut microbiota diversity, reflecting ancestral human-microbe relationships that are largely absent in industrialized nations. Hadza hunter-gatherers demonstrate seasonal microbiota cycling, with bacterial taxa appearing and disappearing based on diet, a flexibility lost in Western populations (Smits et al. 2017). Yanomami Amerindians exhibit the highest recorded human gut bacterial diversity, nearly double that of Americans, with unique antimicrobial resistance genes and protective bacteria, including *Oxalobacter*, which prevents kidney stones (Clemente et al. 2015). These communities maintain fiber-degrading bacteria, including *Prevotella*, *Treponema*, and *Spirochaetes*, which support carbohydrate metabolism and immune regulation (Sonnenburg and Sonnenburg 2014).

Co-evolution between the human diet and gut microbiota has been demonstrated in comparative studies. Traditional populations consuming high-fiber diets (100-150 g daily) support different microbial communities than industrialized populations consuming processed foods with minimal fiber (15 g daily) (Smits et al. 2017). Fermented food consumption provides microbial diversity through indigenous lactic acid bacteria, enhancing short-chain fatty acid production and immunomodulation (Tamang et al. 2016). These dietary patterns maintain microbial gene diversity for metabolizing

plant polysaccharides, with traditional populations harboring 40-77% of unknown bacterial species representing therapeutic potential (Pasolli et al. 2019).

Birth mode and feeding practices shape microbiota development. Vaginal delivery promotes maternal-infant microbial transmission, fostering beneficial bacteria like Bacteroidales and Lachnospiraceae, while cesarean section leads to colonization by opportunistic taxa such as *Enterococcus* and *Klebsiella* (Dominguez-Bello et al. 2016; Shao et al. 2019). Exclusive breastfeeding supports *Bifidobacterium* growth, influenced by maternal genetics (Lewis et al. 2015). Traditional populations show enhanced preservation of milk oligosaccharide-degrading bacterial capabilities, with studies showing higher *Bifidobacterium infantis* prevalence in non-industrialized communities than in Western infants, where this species is largely absent (Tamburini et al. 2016; Carter et al. 2023).

Despite their critical role in understanding the development of the natural human microbiota, traditional populations are underrepresented in research. Southeast Asia accounts for 26% of the global population but only 2% of microbiome samples, whereas North America and Europe, 14% of the population, comprise 71% of the samples (Blake 2024). This underrepresentation limits our understanding of healthy microbiota patterns and their therapeutic applications. Studies of indigenous Malaysians have shown that gut bacteria are mostly uncultured species, highlighting knowledge gaps in traditional community microbiota (Tee et al. 2022).

Indonesian indigenous communities, especially Baduy of West Java, are critically understudied in the literature. They maintain pre-industrial lifestyles that are ideal for studying natural microbiota. The Baduy, approximately 12,000 people, practice strict traditions prohibiting modernization, including sustainable agriculture, traditional foods, and medicine (Iskandar and Iskandar 2017; Iskandar et al. 2018). They have minimal exposure to antibiotics, processed foods, and pollutants that are typical of industrialized societies.

These research gaps necessitate the urgent documentation of microbiota in traditional populations before globalization eliminates these microbial communities. Over 124 bacterial species have vanished from industrialized populations, with traditional communities serving as irreplaceable microbial heritage reservoirs (Carter et al. 2023). Understanding microbiota assembly in traditional Indonesian populations could help identify beneficial bacteria for population-specific interventions while preserving indigenous knowledge.

This study aimed to provide the first characterization of gut microbiota of Baduy infants as an initial investigation into the microbial profiles of this traditional Indonesian population. Examining the microbial community in infants from a population maintaining pre-industrial practices contributes foundational data for understanding the natural development of human microbiota.

MATERIALS AND METHODS

Study area

This study was conducted in Baduy Luar, Kanekes Village, Banten, Indonesia, which represents an indigenous Sundanese community that maintains traditional subsistence practices and exhibits minimal exposure to modern dietary interventions or pharmaceutical treatments. The Baduy Luar settlement constitutes the outer community of the broader Baduy traditional territory, where inhabitants practice sustainable agriculture, predominantly consume plant-based traditional foods, and maintain cultural practices that have remained largely unchanged for centuries (Figure 1).

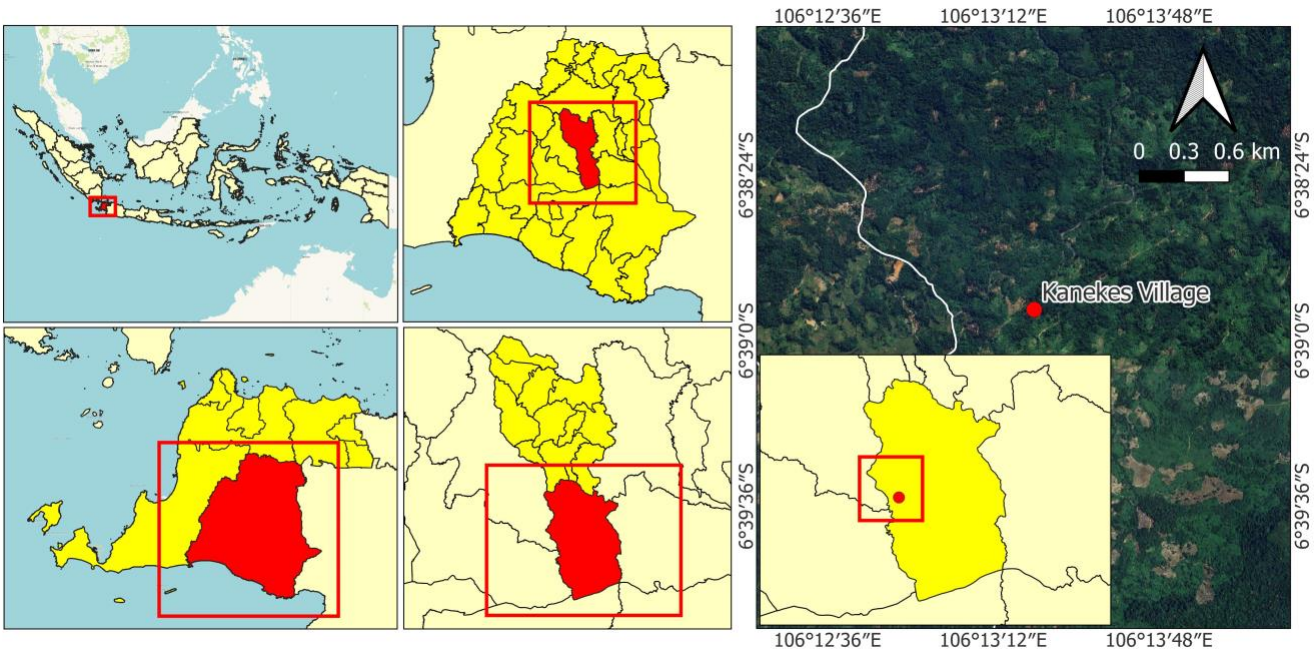

Figure 1. Study site location in Baduy Luar, Kanekes Village, Banten Province, Indonesia. The red dot indicates the sampling site location

Ethics statement

This study was approved by the Health Research Ethics Committee of the Faculty of Public Health, Universitas Diponegoro (No: 319.03/EA/KEPK-FKM/2023). At the beginning of the field study, we reported and asked permission for conducting the research. Written informed consent was obtained from the parents, as explained in a language that is easy to understand, including consent for the collection and publication of information related to the gender and age of participants.

Participant selection criteria

Participant recruitment followed systematically defined inclusion and exclusion criteria developed through culturally responsive methodological frameworks that accommodate traditional community knowledge systems and indigenous privacy protocols within the Baduy research context. The inclusion criteria were as follows: (i) infant age range of 12-15 months representing the transitional microbiota developmental phase; (ii) continuous residence within Baduy Luar territory since birth, ensuring consistent exposure to traditional environmental and dietary factors; and (iii) multi-generational family residence within Baduy territory, indicating sustained traditional lifestyle maintenance. Exclusion criteria included: (i) extended residence outside Baduy territory (>7 consecutive days) within 60 days of sampling and (ii) inability to obtain adequate fecal sample volume due to practical constraints.

Participant demographic and perinatal data collection was constrained by cultural protocols and community privacy considerations within Baduy traditional society. While basic demographic information (age and sex) was documented with parental consent, detailed perinatal variables, including delivery mode, gestational age, birth weight, early feeding practices, maternal health status, and antibiotic exposure history, were not systematically collected. These methodological constraints represent acknowledged limitations that may influence the interpretation of microbiota composition.

Sample collection

Fecal specimens were collected in the morning during natural defecation events using sterile collection methods to prevent contamination from toilet water. Fresh fecal samples (approximately 1 g) were immediately transferred into DNA/RNA Shield™ Fecal Collection Tubes (Zymo Research, Irvine, CA, USA) using a sterile spoon, according to the manufacturer's instructions. Each collection tube containing 9 mL of DNA/RNA Shield reagent was tightly capped and inverted 10 times to create a homogeneous suspension, ensuring complete sample stabilization. The preserved samples were maintained at ambient temperature during transport and subsequently stored at -20°C prior to DNA extraction.

DNA extraction and quality assessment

Total genomic DNA was extracted from fecal samples using the ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research, Irvine, CA, USA), following the manufacturer's protocol. DNA concentration and purity were determined using dual quantification methods: a NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA) for initial assessment and a Qubit fluorometer (Thermo Fisher Scientific) for precise quantification. DNA integrity was evaluated by agarose gel electrophoresis to confirm that high-molecular-weight genomic DNA is suitable for long-read sequencing applications.

Library preparation and nanopore sequencing

Full-length 16S ribosomal RNA gene amplification (1484 bp, V1-V9 regions) was performed using universal bacterial primers 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1492R (5′-TACGGYTACCTTGTTACGACTT-3′), which represent optimized variants with enhanced taxonomic coverage (Isenbarger et al. 2008) based on the foundational primer design framework of Weisburg et al. (1991). PCR reactions were conducted using MyTaq HS Red Mix, 2X (Bioline, BIO-25048, London, UK) according to the manufacturer's specifications, with reaction volumes optimized for full-length 16S rRNA gene amplification. Each 25 µL reaction contained 12.5 µL MyTaq HS Red Mix (2X), 1 µL each of 10 pmol μL^{-1} forward and reverse primers, 5.5 µL nuclease-free water, and 5 µL template DNA (100-250 ng μL^{-1}). Thermal cycling conditions were as follows: initial denaturation at 95°C for 3 min, followed by 30 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 90 s, and a final extension at 72°C for 5 min (Weisburg et al. 1991; Frank et al. 2008).

PCR amplification products were systematically evaluated by agarose gel electrophoresis using a 1% agarose gel matrix prepared in TBE buffer (Bioline, London, UK) supplemented with SYBR™ Safe DNA Gel Stain (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) at manufacturer-recommended concentrations for optimal DNA visualization under blue light transillumination. Electrophoretic separation was conducted with 2 µL of PCR products alongside a 1 kb DNA ladder (Promega, Madison, WI, USA) loaded at 2.5 µL volume as a molecular weight standard to confirm the successful amplification of the expected 1484 bp fragment. Non-Template Control (NTC) reactions were systematically included in each electrophoretic run to verify the absence of contamination artifacts during the PCR amplification. Amplicons demonstrating appropriate molecular weight correspondence and fluorescent intensity were subsequently purified using standard purification protocols and quantified spectrophotometrically prior to library preparation using commercial kits manufactured by Oxford Nanopore Technologies (ONT, Oxford, UK). Library preparation encompassed comprehensive end-repair, barcode ligation, and sequencing adapter attachment procedures, following the manufacturer's optimized protocol for full-length 16S rRNA gene sequencing applications, ensuring maximal sequencing efficiency and taxonomic resolution.

Nanopore sequencing was performed using the MinKNOW software version 23.04.5 on R9.4 flow cells with a GridION sequencer (Oxford Nanopore Technologies). Real-time base-calling was conducted using Guppy version

6.5.7 with high-accuracy models to maximize read quality and taxonomic resolution (Wick et al. 2019).

Bioinformatics analysis

Raw electrical signals (FAST5 format) were converted to FASTQ files through base-calling using Guppy v6.5.7 with High-Accuracy (HAC) models. Quality control and summary statistics were assessed using NanoPlot v1.40.0 (De Coster et al. 2018). Reads were then filtered using NanoFilt v2.8.0 (De Coster et al. 2018) with the following optimized parameters: Phred quality score Q≥12 and read length 1200-1800 bp. This filtering removed low-quality reads and adapter sequences, yielding datasets enriched for full-length 16S rRNA gene amplicons.

Taxonomic classification of filtered reads was performed using Centrifuge classifier v1.0.4 (Kim et al. 2017) against a comprehensive bacteria and archaea index constructed from the NCBI 16S RefSeq database (https://www.ncbi.nlm.nih.gov/refseq/targetedloci/). Classification results were processed using score-based filtering with a minimum alignment score of ≥300 and minimum match length of ≥40% of the read length to ensure taxonomic accuracy. Reads that failed to meet these criteria were classified as unassigned.

Downstream analysis and visualization were conducted using multiple complementary platforms: Pavian v1.0 (https://github.com/fbreitwieser/pavian) for interactive metagenomics analysis, Krona Tools v2.8.1 (https://github.com/marbl/Krona) for hierarchical taxonomic visualization, and R v4.2.3 within RStudio (https://www.R-project.org/) for diversity metric calculations and custom visualization.

Data analysis

Alpha diversity metrics including observed species richness, Chao1 estimator, Shannon diversity index, Simpson index, Inverse Simpson index, Abundance-based Coverage Estimator (ACE), and Fisher's alpha were calculated to assess within-sample microbial diversity following established ecological principles for microbiota analysis.

Taxonomic abundance data were analyzed at multiple hierarchical levels (phylum, class, order, family, and genus) to provide a comprehensive community structure assessment. Relative abundance calculations were performed after normalization to account for variations in the sequencing depth between samples. Due to the exploratory pilot study design and limited sample size (n = 2), formal statistical hypothesis testing was not performed. Alpha diversity metrics and taxonomic abundance patterns were presented as descriptive comparisons to establish the baseline microbiota characteristics within this traditional population.

Data visualization included stacked bar plots for taxonomic composition, Krona charts for hierarchical abundance representation, and Sankey diagrams for quantitative taxonomic flow visualization. All analyses were performed using appropriate quality control measures and analytical validation protocols to ensure the reproducibility and accuracy of the results.

RESULTS AND DISCUSSION

Electrophoresis result and data quality information

PCR amplification of the full-length 16S rRNA gene (approximately 1484 bp) from fecal samples of two Baduy infants yielded distinct single bands at the expected size (Figure 2). Both samples (lanes 1 and 2) showed clean, specific amplification, while no band was observed in the NTC, confirming the absence of contamination and validating the specificity of the PCR reactions.

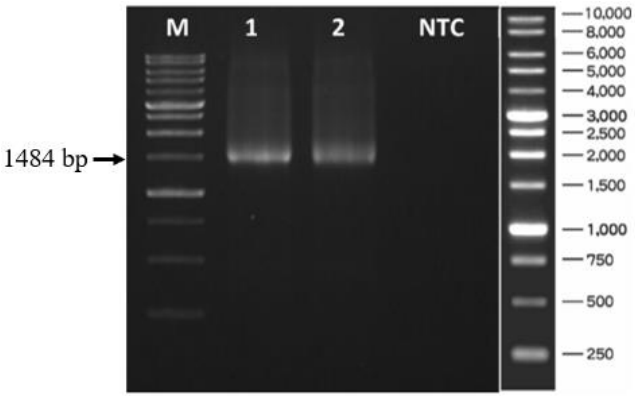

Figure 2. Agarose-gel electrophoresis of PCR amplification products of the full-length 16S rRNA gene (V1-V9, approximately 1484 bp) from fecal samples of two Baduy infants. M: 1-kb DNA ladder, band sizes indicated at right). 1: 13BL_P (13-month-old female), 2: 12BL_L (12-month-old male); NTC: non-template control

Table 1. Nanopore sequencing quality metrics and preprocessing outcomes for full-length 16S rRNA gene amplicons from Baduy infant fecal samples

Sample ID	Raw dataset metrics				Quality-filtered dataset metrics				Processing efficiency
	Total reads (reads)	Total bases (bp)	Mean length (bp)	Mean Q-score (Phred)	Retained reads (reads)	Retained bases (bp)	Mean length (bp)	Mean Q-score (Phred)	Retention rate (%)
12BL_L	100,000	149,905,445	1499	14.7	84,635	135,879,667	1606	15.1	84.6
13BL_P	97,673	143,260,378	1467	14.6	80,562	127,787,416	1586	15.0	82.5

Note: Q-scores >14.0 indicate >97.5% base-calling accuracy, and retention rates >80% indicate optimal sequencing performance. Both samples achieved sufficient sequencing depth (>80,000 high-quality reads) for robust taxonomic classification and species-level resolution of the data

Nanopore sequencing of these PCR amplicons generated high-quality long-read datasets suitable for microbial community analysis. The 12BL_L sample yielded 100,000 raw reads (mean length 1499 bp; mean quality score 14.7), corresponding to 149.9 Mb. After quality filtering with NanoFilt (Q≥12; read length 1200-1800 bp), 84,635 reads (approximately 85%) were retained, with a mean length of 1606 bp (approximately 1.6 kb), a mean quality score of 15.1, and a total of 135.9 Mb. Similarly, the 13BL_P sample produced 97,673 raw reads (mean length 1467 bp; mean quality score 14.6; total 143.3 Mb). After filtering, 80,562 reads (approximately 82%) remained, with a mean length of 1586 bp (approximately 1.6 kb), a mean quality score of 15.0, and a total of 127.8 Mb.

Overall, approximately 82-85% of the reads were retained after filtering, and the filtered datasets exhibited mean read lengths of approximately 1.6 kb, consistent with the expected size of full-length 16S rRNA gene amplicons. Sequencing depths exceeding 125 Mb per sample indicated sufficient coverage to support downstream microbial diversity and taxonomic analyses. A detailed summary of sequencing statistics for each sample is provided in Table 1. The high-quality sequencing datasets obtained from both infants provided a solid basis for subsequent analyses of microbial diversity and taxonomic composition.

Diversity and composition of microbiota communities

Alpha diversity analysis was used to quantify bacterial richness and diversity patterns within each sample. The two infant samples, 13BL_P (13-month-old female) and 12BL_L (12-month-old male), showed different patterns. The richness metrics revealed that sample 12BL_L contained more bacterial species (543 observed species) than compared to sample 13BL_P (511 observed species) (Figure 3). Estimators confirmed this pattern, with 12BL_L showing Chao1 values of 940 (±77) and ACE values of 890 (±17), whereas 13BL_P exhibited lower values of 755 (±49) and 761 (±15), respectively (Figure 3). Diversity indices demonstrated contrasting patterns, with 13BL_P displaying higher Shannon diversity (3.44 versus 3.03), InvSimpson values (15.9 versus 8.4), and Simpson index values (0.94 versus 0.88). For Fisher's alpha values, the 12BL_L sample was relatively higher than the 13BL_P sample (77.6 versus 72.9) (Figure 3).

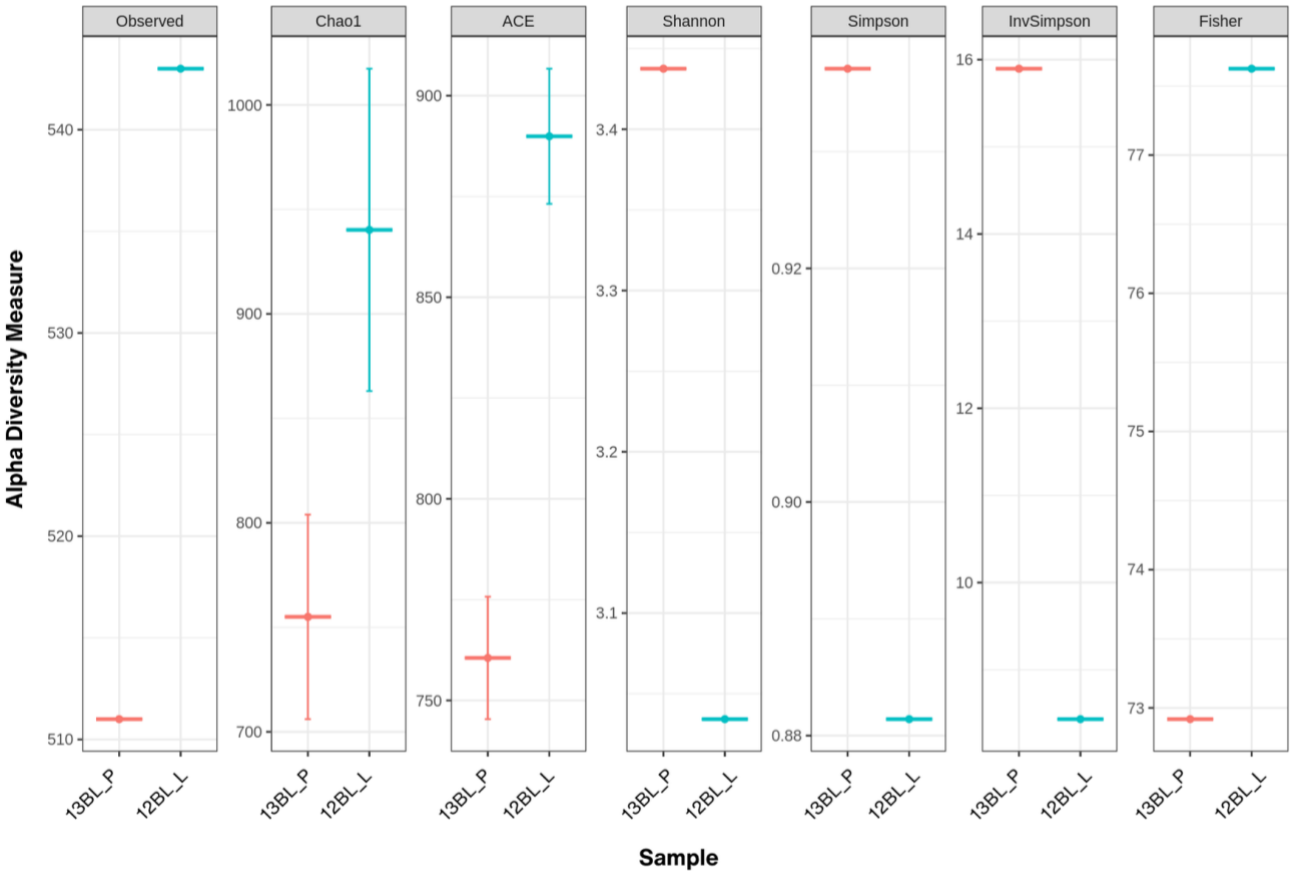

Figure 3. Alpha diversity metrics for fecal microbiota samples from two Baduy infants. Sample identification: 13BL_P (13-month-old female, red) and 12BL_L (12-month-old male, blue). Seven diversity indices were displayed across separate panels: Species richness estimators (Observed, Chao1, ACE), diversity indices (Shannon, Simpson, and InvSimpson), and Fisher's alpha. Error bars represent standard errors for the Chao1 and ACE richness estimators

Building on these diversity analyses, genus-level taxonomic profiles revealed clear differences between the two infant samples (Figure 4). At 13 months, sample 13BL_P exhibited higher taxonomic heterogeneity and evenness, dominated by *Faecalibacterium*, followed by *Anaerobutyricum*, *Bifidobacterium*, *Bacteroides*, and *Veillonella*, with minor contributions from *Mediterraneibacter* and *Limosilactobacillus*. At 12 months, sample 12BL_L, in contrast, displayed lower diversity and was dominated by *Enterococcus* and *Bifidobacterium*, with a substantial representation of *Streptococcus*. These two profiles highlight marked differences in taxonomic composition: one exhibits greater evenness, while the other is dominated by a few taxa (Figure 4).

To further support these observations, hierarchical abundance relationships among bacterial genera were assessed using heatmap visualization (Figure 5). This analysis demonstrated pronounced compositional differences across the full spectrum of detected taxa, with abundance gradients spanning four orders of magnitude (1-4096 sequence counts). Distinct clustering patterns were observed, reflecting systematic segregation of bacterial groups between the two infant samples.

Further hierarchical visualization using Krona charts and Sankey diagrams provided enhanced taxonomic resolution, revealing the diverse composition of the female profile at the species level (Figure 6.A). Sample 13BL_P demonstrated substantial proportions of *Anaerobutyricum hallii* (4%), *Anaerobutyricum soehngenii* (5%), and *Fusicatenibacter saccharivorans* (5%), with notable presence of *Veillonella ratti* (4%) and *Sellimonas intestinalis* (4%) (Figure 6.A). Quantitative analysis through Sankey visualization confirmed greater *Bacteroides* representation (4.49k, Figure 6.A), alongside substantial populations within Oscillospiraceae (23.6k, Figure 6.A), a family that includes taxa which have been associated with intestinal inflammation in some studies. Consistent with these observations, the female infants displayed elevated relative abundances of *Faecalibacterium butyricigenerans* (5%), *Blautia* (2%), and *Dorea longicatena* (4%) (Figure 6.A), which are potentially involved in inflammatory modulation, collectively representing a more balanced distribution among constituent taxa.

Comparative visualization methodologies consistently demonstrated a pronounced predominance of potentially pathogenic bacteria in the male sample (Figure 6.B). Sample 12BL_L exhibited marked abundance of *Bifidobacterium longum* (29% versus 11% in female) alongside a substantial enterococcal burden, with significantly elevated concentrations of multiple *Enterococcus* species: *E. faecium* (9.56k/11%), *E. faecalis* (5.42k/6%), *E. avium* (2.55k/3%), and *E. durans* (2.11k/2%) (Figure 6.B). The male microbiota displayed considerable streptococcal diversity through both visualization techniques, with *Streptococcus lactarius* (2%), *S. thermophilus* (2%), and *S. salivarius* (3%), represented by Krona visualization (Figure 6.B, left panel radial sectors), corresponding to quantitative measurements of *S. lactarius* (1.55k), *S. thermophilus* (1.39k), and *S. salivarius* (2.43k) in the Sankey diagram (Figure 6.B, right panel numerical annotations). Notably, this sample uniquely exhibited substantial presence of Enterobacteriaceae (6.06k), with distinct *Escherichia* (1.60k) and *Shigella* (409) components established enteric pathogens largely absent in the female profile (Figure 6.B, Sankey diagram numerical annotations; compare with Figure 6.A). These integrated visualization approaches collectively substantiate the distinct colonization patterns of opportunistic pathogens between the two subjects, with implications for understanding developmental microbiota dynamics. These taxonomic observations provide compositional insights, although functional activity and metabolic consequences require validation through complementary approaches, including metatranscriptomics, metabolomics, and cultivation-based methods.

Discussion

This exploratory study of two Baduy infants provides preliminary insights into gut microbiota patterns within an understudied traditional Indonesian population, although the findings require cautious interpretation given the limited sample size. The two contrasting microbiota profiles, one dominated by beneficial commensals and the other by potentially pathogenic taxa, illuminate the fundamental principles of microbial ecosystem assembly in traditional populations, while highlighting the complex interplay of factors influencing early life microbiota development. Recent large-scale studies consistently demonstrate that the age-related developmental stage exerts a substantially stronger influence on infant gut microbiota composition than sex, with even a one-month age difference producing measurable changes (Wu et al. 2024; Fahur Bottino et al. 2025). Consistent with these findings, the differences observed between the two subjects in this study are more plausibly attributable to age-related microbial maturation rather than sex.

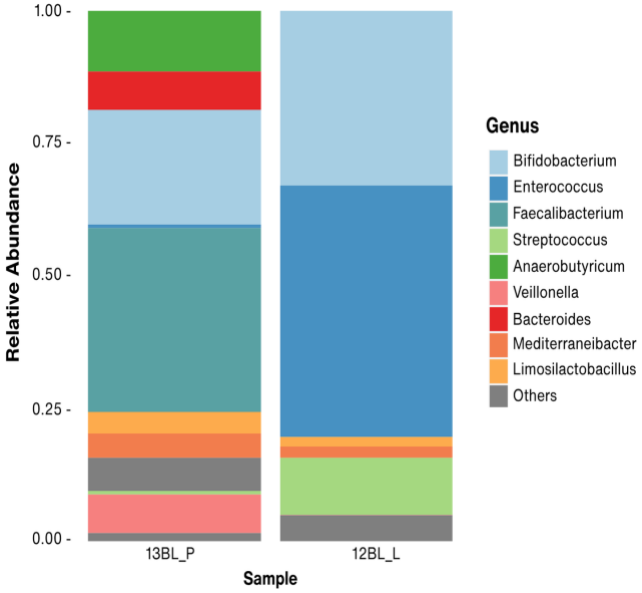

Figure 4. Taxonomic composition at the genus level in fecal microbiota samples from two Baduy infants. 13BL_P: 13-month-old female infant, 12BL_L: 12-month-old male infant

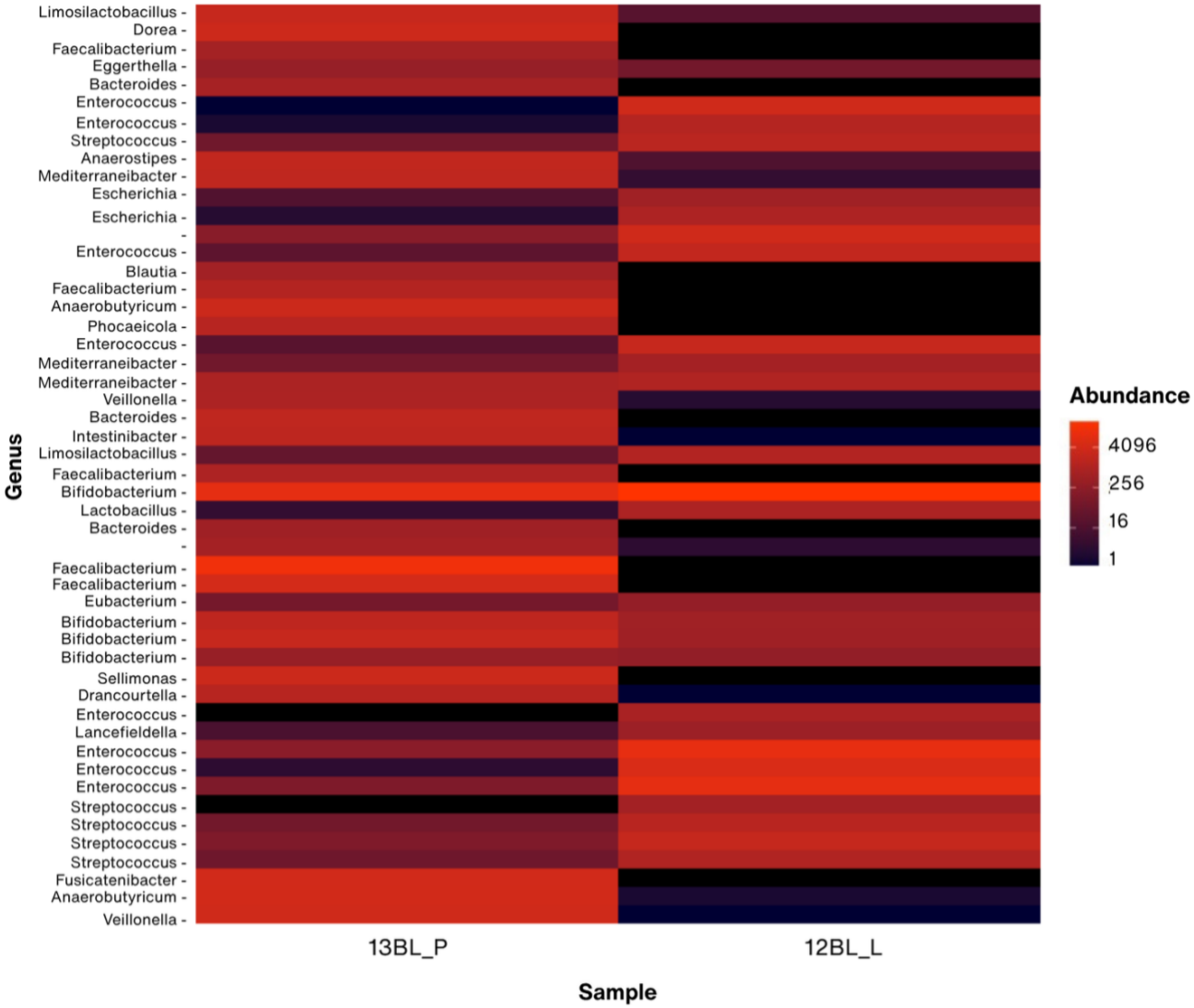

Figure 5. Genus-level abundance heatmap of fecal microbiota in Baduy infants. Hierarchical visualization displaying the relative abundance patterns of predominant bacterial genera across two infant fecal samples: 13BL_P (13-month-old female) and 12BL_L (12-month-old male). Color intensity reflects relative abundance patterns using logarithmic scaling, with red indicating higher abundance and dark blue to black representing lower abundance (scale: 1-4096 sequence counts represent the empirical range of raw sequence counts before logarithmic transformation and color mapping)

Baduy infant microbiota patterns can be contextualized within the broader literature on traditional populations, although direct comparisons require caution given the methodological differences. Studies of traditional populations have revealed distinctive microbiota characteristics that differ markedly from those of industrialized societies. Hadza hunter-gatherers demonstrate exceptional microbial diversity, with over 20% of detected genomes representing novel species absent from global databases and the complete absence of *Bifidobacterium* throughout development (Schnorr et al. 2014; Olm et al. 2022).

Yanomami populations exhibited the highest recorded bacterial diversity in humans, with Firmicutes/Bacteroidetes ratios of 0.24 reflecting traditional lifestyle patterns (Clemente et al. 2015). Mexican Me´phaa indigenous children show high *Prevotella* abundance and enrichment of VANISH (Volatile and/or Associated Negatively with Industrialized Societies of Humans) bacterial groups, including *Treponema* and *Eubacterium* species (Sánchez-Quinto et al. 2020). In comparison, the microbiota composition of female Baduy infants, characterized by *Faecalibacterium* dominance and substantial *Bifidobacterium* representation, exhibited features distinct from those of documented traditional populations. The presence of appreciable *Bifidobacterium* levels contrasts with Hadza patterns, but aligns more closely with rural populations that maintain some traditional practices while having limited exposure to industrialization. Elevated *Enterococcus* and *Streptococcus* abundances in male infants represent patterns less commonly reported in traditional population studies.

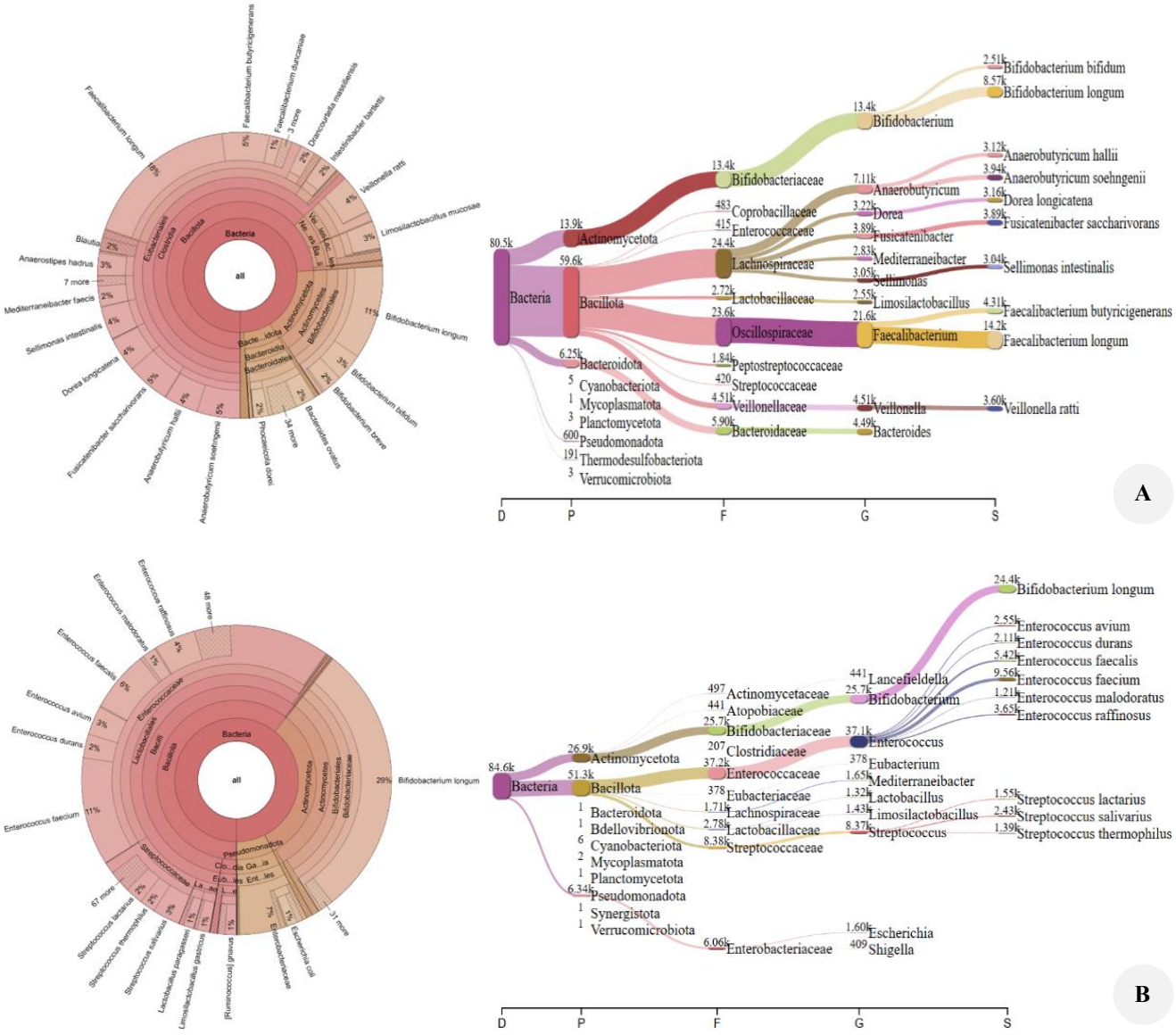

Figure 6. Integrated visualization approaches for metagenomic profiling of infant fecal microbiota samples using complementary analytical platforms. Krona charts (left panels) display hierarchical taxonomic organization with proportional abundance representation from the domain to species level, with sector size corresponding to relative abundance and colors distinguishing taxonomic classifications. Sankey diagrams (right panels) quantitatively illustrate taxonomic abundances across hierarchical classifications (D: Domain, P: Phylum, F: Family, G: Genus, S: Species), with bandwidth proportional to absolute read counts and flow direction representing taxonomic relationships from higher to lower classification levels. Panel A: Sample 13BL_P (13-month-old female infant), Panel B: Sample 12BL_L (12-month-old male infant). The numerical values within the Sankey diagrams represent the absolute sequence counts for each taxonomic classification

Rural-urban comparative studies help explain how Baduy's traditional lifestyle might influence their infants' microbiota, relative to more modernized populations. Nigerian rural Bassa infants demonstrated significantly higher microbial diversity and accelerated adult-like maturation compared to their urban counterparts, with environmental factors, including untreated water consumption and extensive soil contact, promoting microbial dispersal throughout communities (Ayeni et al. 2018). Vietnamese studies revealed that rural Ha Long Bay infants maintained higher beneficial bacterial concentrations while urban Ha Noi infants exhibited increased potentially harmful bacteria including *Klebsiella* and *Citrobacter* (Kortman et al. 2023). The urban infant microbiota consistently demonstrates reduced diversity, delayed maturation, and altered taxonomic composition across multiple geographic regions, correlating with higher rates of cesarean delivery, increased antibiotic exposure, and reduced environmental microbial contact (Morandini et al. 2023). These rural-urban differences demonstrate that traditional lifestyle factors preserve beneficial microbiota patterns, providing a framework for understanding the distinct microbiota profiles of Baduy infants as products of their traditional environmental exposure.

The observed microbiota differences between the two infant samples may reflect the different environmental microbial exposure characteristics of traditional Baduy agricultural practices. The Baduy maintains a zero-tillage swidden agriculture system (*huma*) with complete prohibition of chemical fertilizers and pesticides (Iskandar et al. 2018), preserving natural soil microbial communities that could influence human gut colonization through multiple pathways (Blum et al. 2019). Their traditional farming methods involve direct soil contact during rice cultivation and forest foraging activities (Iskandar and Iskandar 2017), potentially exposing individuals to diverse soil microbiota that are absent in industrialized populations. The 5-10-year fallow period practiced in Baduy agriculture allows natural microbial succession in soil ecosystems, creating temporal and spatial variations in environmental microbial exposure (Blum et al. 2019). The contrasting microbiota compositions may also reflect varying degrees of forest microbiota exposure, as approximately 48.85% of Baduy territory consists of protected forests providing regular microbial contact through traditional livelihood activities (Saleh et al. 2020). The Baduy practice extensive forest foraging for honey, medicinal plants, and seasonal foods (Iskandar and Iskandar 2017), exposing individuals to diverse environmental bacteria that can seed the gut microbiota (Mhuireach et al. 2023). Additionally, the distinct microbiota profiles observed may reflect variations in traditional Baduy dietary exposure, particularly regarding fermented foods and traditional food preparation methods. The complete prohibition of modern food additives, preservatives, and processed foods (Lindawati et al. 2024) means that gut microbiota development occurs exclusively through traditional food sources, potentially supporting distinct bacterial populations compared with industrialized populations (Li et al. 2025).

The bacterial composition observed in female infants includes genera with well-documented beneficial functions, although extrapolating health implications from taxonomic data alone requires caution. *Faecalibacterium prausnitzii* produces butyrate, which supports intestinal barrier function and exhibits anti-inflammatory properties by reducing inflammatory signaling pathways (Rios-Covian et al. 2015). *Bifidobacterium* species demonstrate a specialized capacity for human milk oligosaccharide utilization and produce metabolites that support immune development (Rivière et al. 2016; Laursen et al. 2021). *Anaerobutyricum* species contribute to butyrate production via lactate conversion mechanisms, potentially supporting metabolic cross-feeding networks (Shetty et al. 2020). The presence of these taxa may indicate optimal environmental microbial exposure, thereby supporting beneficial fermentation processes. Conversely, the taxonomic profile of male infants includes bacterial groups with documented pathogenic potential in clinical contexts, although their significance within healthy traditional populations remains unclear. *Enterococcus* species, particularly *E. faecium* and *E. faecalis*, exhibit intrinsic antibiotic resistance and have been associated with healthcare-associated infections (Prieto et al. 2016). Certain *Streptococcus* species can cause neonatal infections, although clinical outcomes depend on multiple factors including strain characteristics, host immunity, and environmental context. Enterobacteriaceae expansion has been linked to intestinal inflammation in some populations (Litvak et al. 2017), although these clinical associations are primarily derived from studies in industrialized populations and may not apply directly to traditional communities.

Methodological considerations necessarily temper interpretation. The absence of critical perinatal metadata, including delivery mode, gestational age, early feeding practices, and maternal microbiota composition, represents a significant constraint, as these variables constitute primary determinants of infant gut colonization patterns (Milani et al. 2017; Wampach et al. 2018). The cross-sectional design and limited sample size precluded determination of temporal stability or causal relationships. Additionally, although 16S rRNA sequencing provides robust taxonomic resolution, it cannot directly assess functional gene expression or strain-level virulence determinants. This exploratory study identified notable microbiota heterogeneity between two infants within a culturally homogeneous traditional population, suggesting that environmental homogeneity may not uniformly determine the microbial assembly patterns. While these preliminary observations cannot establish causality given the limited sample size, they indicate the potential roles of intrinsic host factors in shaping colonization trajectories.

In conclusion, this investigation established the first microbiota baseline for the Baduy indigenous community, demonstrating that traditional populations harbor irreplaceable microbial diversity critical for global microbiota conservation efforts. The observed heterogeneity among infants, particularly the contrasting beneficial versus potentially pathogenic profiles, reveals urgent research priorities for understanding the protective mechanisms against modern microbiota depletion. The documented microbiota patterns suggest potential resilience mechanisms that confer resistance to modern disorders prevalent in industrialized populations, including inflammatory diseases, metabolic dysfunction, and immune-mediated conditions. The enrichment of butyrate-producing taxa (*Faecalibacterium*, *Anaerobutyricum*) and beneficial commensals (*Bifidobacterium*) in female infants represents microbial configurations associated with enhanced barrier function, anti-inflammatory responses, and optimal immune programming that may provide natural protection against obesity, autoimmune disorders, and chronic inflammatory conditions characteristic of Western populations.

Future investigations should prioritize longitudinal cohort studies that incorporate comprehensive environmental sampling, maternal microbiota analysis, and detailed documentation of traditional practices, including fermented food preparation and forest interaction protocols. Such research would elucidate the mechanisms of microbial inheritance within traditional communities and inform therapeutic strategies for restoring the beneficial microbiota in industrialized populations. Understanding how traditional populations maintain protective microbial communities represents a critical window for developing culturally appropriate interventions, while preserving indigenous

knowledge systems and respecting community sovereignty over biological resources.

ACKNOWLEDGEMENTS

The authors express their sincere gratitude to the Faculty of Public Health, Universitas Diponegoro, Semarang, Indonesia, for providing facilities and resources essential for the successful completion of this study. We extend our appreciation to the Baduy community, particularly to the families who participated in this study and shared their invaluable traditional knowledge. We also appreciate the local public health center for facilitating this study.

REFERENCES

Ayeni FA, Biagi E, Rampelli S, Fiori J, Soverini M, Audu HJ, Cristino S, Caporali L, Schnorr SL, Carelli V, Brigidi P, Candela M, Turroni S. 2018. Infant and adult gut microbiome and metabolome in rural Bassa and urban settlers from Nigeria. Cell Rep 23 (10): 3056-3067. DOI: 10.1016/j.celrep.2018.05.018.

Bäckhed F, Roswall J, Peng Y et al. 2015. Dynamics and stabilization of the human gut microbiome during the first year of life. Cell Host Microbe 17 (5): 690-703. DOI: 10.1016/j.chom.2015.04.004.

Blake KS. 2024. Missing microbiomes: Global underrepresentation restricts who research will benefit. J Clin Invest 134 (14): e183884. DOI: 10.1172/JCI183884.

Blum WEH, Zechmeister-Boltenstern S, Keiblinger KM. 2019. Does soil contribute to the human gut microbiome? Microorganisms 7 (9): 287. DOI: 10.3390/microorganisms7090287.

Carter MM, Olm MR, Merrill BD, Dahan D, Tripathi S, Spencer SP, Yu FB, Jain S, Neff N, Jha AR, Sonnenburg ED, Sonnenburg JL. 2023. Ultra-deep sequencing of Hadza hunter-gatherers recovers vanishing gut microbes. Cell 186: 3111-3124.e13. DOI: 10.1016/j.cell.2023.05.046.

Clemente JC, Pehrsson EC, Blaser MJ, Sandhu K, Gao Z, Wang B, Magris M, Hidalgo G, Contreras M, Noya-Alarcón O, Lander O, McDonald J, Cox M, Walter J, Oh PL, Ruiz JF, Rodriguez S, Shen N, Song SJ, Metcalf J, Knight R, Dantas G, Dominguez-Bello MG. 2015. The microbiome of uncontacted Amerindians. Sci Adv 1 (3): e1500183. DOI: 10.1126/sciadv.1500183.

De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: Visualizing and processing long-read sequencing data. Bioinformatics 34: 2666-2669. DOI: 10.1093/bioinformatics/bty149.

Dominguez-Bello MG, De Jesus-Laboy KM, Shen N, Cox LM, Amir A, Gonzalez A, Bokulich NA, Song SJ, Hoashi M, Rivera-Vinas JI, Mendez K, Knight R, Clemente JC. 2016. Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. Nat Med 22 (3): 250-253. DOI: 10.1038/nm.4039.

Fahur Bottino G, Bonham KS, Patel F, McCann S, Zieff M, Naspolini N, Ho D, Portlock T, Joos R, Midani FS, Schüroff P, Das A, Shennon I, Wilson BC, O'Sullivan JM, Britton RA, Murray DM, Kiely ME, Taddei CR, Beltrão-Braga PCB, Campos AC, Polanczyk GV, Huttenhower C, Donald KA, Klepac-Ceraj V. 2025. Early life microbial succession in the gut follows common patterns in humans across the globe. Nat Commun 16 (1): 660. DOI: 10.1038/s41467-025-56072-w.

Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol 74 (8): 2461-2470. DOI: 10.1128/aem.02272-07.

Isenbarger TA, Finney M, Ríos-Velázquez C, Handelsman J, Ruvkun G. 2008. Miniprimer PCR, a new lens for viewing the microbial world. Appl Environ Microbiol 74 (3): 840-849. DOI: 10.1128/aem.01933-07.

Iskandar BS, Iskandar J, Partasasmita R. 2018. Strategy of the Outer Baduy community of South Banten (Indonesia) to sustain their swidden farming traditions by temporary migration to non-Baduy areas. Biodiversitas 19 (2): 453-464. DOI: 10.13057/biodiv/d190212.

Iskandar J, Iskandar BS. 2017. Various plants of traditional rituals: Ethnobotanical research among the Baduy community. Biosaintifika 9 (1): 114-125. DOI: 10.15294/biosaintifika.v9i1.8117.

Kim D, Song L, Breitwieser FP, Salzberg SL. 2017. Centrifuge: Rapid and sensitive classification of metagenomic sequences. Genome Res 26 (12): 1721-1729. DOI: 10.1101/gr.210641.116.

Kortman GAM, Timmerman HM, Schaafsma A, Stoutjesdijk E, Muskiet FAJ, Nhien NV, van Hoffen E, Boekhorst J, Nauta A. 2023. Mothers' breast milk composition and their respective infant's gut microbiota differ between five distinct rural and urban regions in Vietnam. Nutrients 15 (22): 4802. DOI: 10.3390/nu15224802.

Laursen MF, Sakanaka M, von Burg N, Mörbe U, Andersen D, Moll JM, Pekmez CT, Rivollier A, Michaelsen KF, Mølgaard C, Lind MV, Dragsted LO, Katayama T, Frandsen HL, Vinggaard AM, Bahl MI, Brix S, Agace W, Licht TR, Roager HM. 2021. *Bifidobacterium* species associated with breastfeeding produce aromatic lactic acids in the infant gut. Nat Microbiol 6 (11): 1367-1382. DOI: 10.1038/s41564-021-00970-4.

Lewis ZT, Totten SM, Smilowitz JT, Popovic M, Parker E, Lemay DG, Van Tassell ML, Miller MJ, Jin Y-S, German JB, Lebrilla CB, Mills DA. 2015. Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants. Microbiome 3: 13. DOI: 10.1186/s40168-015-0071-z.

Li F, Armet AM, Korpela K et al. 2025. Cardiometabolic benefits of a non-industrialized-type diet are linked to gut microbiome modulation. Cell 188 (5): 1226-1247.e18. DOI: 10.1016/j.cell.2024.12.034.

Lindawati L, Anugrah IS, Tarigan H, Indraningsih KS, Purwantini TB, Hariyanto W. 2024. *Huma* rice management model and food security system of Baduy indigenous communities, Indonesia. Bio Web Conf 119: 02004. DOI: 10.1051/bioconf/202411902004.

Litvak Y, Byndloss MX, Tsolis RM, Bäumler AJ. 2017. Dysbiotic Proteobacteria expansion: A microbial signature of epithelial dysfunction. Curr Opin Microbiol 39: 1-6. DOI: 10.1016/j.mib.2017.07.003.

Mhuireach GÁ, Van Den Wymelenberg KG, Langellotto GA. 2023. Garden soil bacteria transiently colonize gardeners' skin after direct soil contact. Urban Agric Region Food Syst 8 (1): e20035. DOI: 10.1002/uar2.20035.

Milani C, Duranti S, Bottacini F, Casey E, Turroni F, Mahony J, Belzer C, Palacio SD, Montes SA, Mancabelli L, Lugli GA, Rodriguez JM, Bode L, de Vos W, Gueimonde M, Margolles A, van Sinderen D, Ventura M. 2017. The first microbial colonizers of the human gut: Composition, activities, and health implications of the infant gut microbiota. Microbiol Mol Biol Rev 81 (4): e00036-17. DOI: 10.1128/mmbr.00036-17.

Morandini F, Perez K, Brot L, Seck SM, Tibère L, Grill J-P, Macia E, Seksik P. 2023. Urbanization associates with restricted gut microbiome diversity and delayed maturation in infants. iScience 26 (11): 108136. DOI: 10.1016/j.isci.2023.108136.

Olm MR, Dahan D, Carter MM, Merrill BD, Yu FB, Jain S, Meng X, Tripathi S, Wastyk H, Neff N, Holmes S, Sonnenburg ED, Jha AR, Sonnenburg JL. 2022. Robust variation in infant gut microbiome assembly across a spectrum of lifestyles. Science 376 (6598): 1220-1223. DOI: 10.1126/science.abj2972.

Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P, Collado MC, Rice BL, DuLong C, Morgan XC, Golden CD, Quince C, Huttenhower C, Segata N. 2019. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. Cell 176 (3): 649-662.e20. DOI: 10.1016/j.cell.2019.01.001.

Prieto AMG, van Schaik W, Rogers MRC, Coque TM, Baquero F, Corander J, Willems RJL. 2016. Global emergence and dissemination of enterococci as nosocomial pathogens: Attack of the clones? Front Microbiol 7: 788. DOI: 10.3389/fmicb.2016.00788.

Rios-Covian D, Gueimonde M, Duncan SH, Flint HJ, de los Reyes-Gavilan CG. 2015. Enhanced butyrate formation by cross-feeding between *Faecalibacterium prausnitzii* and *Bifidobacterium adolescentis*. FEMS Microbiol Lett 362 (21): fnv176. DOI: 10.1093/femsle/fnv176.

Rivière A, Selak M, Lantin D, Leroy F, De Vuyst L. 2016. Bifidobacteria and butyrate-producing colon bacteria: Importance and strategies for their stimulation in the human gut. Front Microbiol 7: 979. DOI: 10.3389/fmicb.2016.00979.

Robertson RC, Manges AR, Finlay BB, Prendergast AJ. 2019. The human microbiome and child growth - first 1000 days and beyond. Trends Microbiol 27 (2): 131-147. DOI: 10.1016/j.tim.2018.09.008.

Saleh I, Atmaja ISW, Syahadat RM. 2020. Prohibition in Baduy Dalam Community: Soil and water conservation perspective. In: Proceedings

of the International Conference on Agriculture, Social Sciences, Education, Technology and Health (ICASSETH 2019). DOI: 10.2991/assehr.k.200402.040.

Sánchez-Quinto A, Cerqueda-García D, Falcón LI, Gaona O, Martínez-Correa S, Nieto J, G-Santoyo I. 2020. Gut microbiome in children from indigenous and urban communities in México: Different subsistence models, different microbiomes. Microorganisms 8 (10): 1592. DOI: 10.3390/microorganisms8101592.

Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, Turroni S, Biagi E, Peano C, Severgnini M, Fiori J, Gotti R, Bellis GD, Luiselli D, Brigidi P, Mabulla A, Marlowe F, Henry AG, Crittenden AN. 2014. Gut microbiome of the Hadza hunter-gatherers. Nat Commun 5: 3654. DOI: 10.1038/ncomms4654.

Shao Y, Forster SC, Tsaliki E, Vervier K, Strang A, Simpson N, Kumar N, Stares MD, Rodger A, Brocklehurst P, Field N, Lawley TD. 2019. Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. Nature 574 (7776): 117-121. DOI: 10.1038/s41586-019-1560-1.

Shetty SA, Boeren S, Bui TPN, Smidt H, de Vos WM. 2020. Unravelling lactate-acetate and sugar conversion into butyrate by intestinal Anaerobutyricum and *Anaerostipes* species by comparative proteogenomics. Environ Microbiol 22 (11): 4863-4875. DOI: 10.1111/1462-2920.15269.

Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, Knight R, Manjurano A, Changalucha J, Elias JE, Dominguez-Bello MG, Sonnenburg JL. 2017. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. Science 357 (6353): 802-806. DOI: 10.1126/science.aan4834.

Sonnenburg ED, Sonnenburg JL. 2014. Starving our microbial self: The deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. Cell Metab 20 (5): 779-786. DOI: 10.1016/j.cmet.2014.07.003.

Stewart CJ, Ajami NJ, O'Brien JL et al. 2018. Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature 562 (7728): 583-588. DOI: 10.1038/s41586-018-0617-x.

Stiemsma LT, Michels KB. 2018. The role of the microbiome in the developmental origins of health and disease. Pediatrics 141 (4): e20172437. DOI: 10.1542/peds.2017-2437.

Tamang JP, Shin D-H, Jung S-J, Chae S-W. 2016. Functional properties of microorganisms in fermented foods. Front Microbiol 7: 578. DOI: 10.3389/fmicb.2016.00578.

Tamburini S, Shen N, Wu HC, Clemente JC. 2016. The microbiome in early life: Implications for health outcomes. Nat Med 22 (7): 713-722. DOI: 10.1038/nm.4142.

Tee MZ, Er YX, Easton AV, Yap NJ, Lee IL, Devlin J, Chen Z, Ng KS, Subramanian P, Angelova A, Oyesola O, Sargsian S, Ngui R, Beiting DP, Boey CCM, Chua KH, Cadwell K, Lim YAL, Loke P, Lee LC. 2022. Gut microbiome of helminth-infected indigenous Malaysians is context dependent. Microbiome 10 (1): 214. DOI: 10.1186/s40168-022-01385-x.

Wampach L, Heintz-Buschart A, Fritz JV, Ramiro-Garcia J, Habier J, Herold M, Narayanasamy S, Kaysen A, Hogan AH, Bindl L, Bottu J, Halder R, Sjöqvist C, May P, F Andersson A, de Beaufort C, Wilmes P. 2018. Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential. Nat Commun 9 (1): 5091. DOI: 10.1038/s41467-018-07631-x.

Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 1991. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173 (2): 697-703. DOI: 10.1128/jb.173.2.697-703.1991.

Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome Biol 20 (1): 129. DOI: 10.1186/s13059-019-1727-y.

Wu J, Shen H, Lv Y, He J, Xie X, Xu Z, Yang P, Qian W, Bai T, Hou X. 2024. Age over sex: Evaluating gut microbiota differences in healthy Chinese populations. Front Microbiol 15: 1412991. DOI: 10.3389/fmicb.2024.1412991.