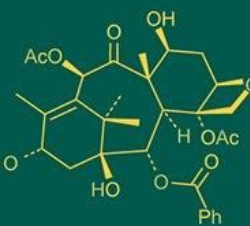
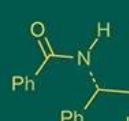
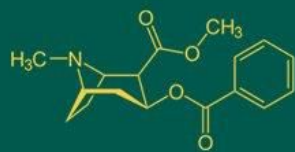


International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
 ISSN Online: 2617-4707
 IJABR 2024; 8(4): 233-238
www.biochemjournal.com
 Received: 19-01-2024
 Accepted: 21-02-2024

Budi Santoso Wijaya
 Department of Microbiology
 and Genomics, Faculty of
 Medicine, Universitas
 Indonesia, Jakarta, Indonesia

Dewi Kusuma Putri
 Department of Microbiology
 and Genomics, Faculty of
 Medicine, Universitas
 Indonesia, Jakarta, Indonesia

Agus Prasetyo Nugroho
 Department of Microbiology
 and Genomics, Faculty of
 Medicine, Universitas
 Indonesia, Jakarta, Indonesia

Siti Rahayu Hartono
 Department of Microbiology
 and Genomics, Faculty of
 Medicine, Universitas
 Indonesia, Jakarta, Indonesia

Corresponding Author:
Budi Santoso Wijaya
 Department of Microbiology
 and Genomics, Faculty of
 Medicine, Universitas
 Indonesia, Jakarta, Indonesia

Next-generation sequencing for microbiome profiling

Budi Santoso Wijaya, Dewi Kusuma Putri, Agus Prasetyo Nugroho and Siti Rahayu Hartono

DOI: <https://doi.org/10.33545/26174693.2024.v8.i4c.7552>

Abstract

The human gut microbiome demonstrates substantial variation across populations with distinct dietary patterns and environmental exposures, yet comprehensive characterization in Indonesian populations remains limited. This cross-sectional research employed 16S rRNA next-generation sequencing to profile gut microbiome composition across four Indonesian population groups representing diverse geographic, dietary, and lifestyle contexts. Participants were recruited from urban Jakarta, rural Java, coastal Sulawesi, and highland Papua between August 2020 and June 2022.

The research cohort comprised 300 healthy adults aged 25-55 years including 86 from urban Jakarta, 78 from rural Java, 72 from coastal Sulawesi, and 64 from highland Papua. Fecal samples underwent DNA extraction and V3-V4 16S rRNA amplicon sequencing on Illumina MiSeq platform. Bioinformatic analysis utilized QIIME2 with DADA2 denoising and SILVA database taxonomic classification. Alpha diversity, beta diversity, and differential abundance analyses characterized microbiome variation across populations.

Sequencing generated 12.4 million quality-filtered reads with mean depth of 41,280 reads per sample. Taxonomic profiling identified 4,847 amplicon sequence variants across all samples. Firmicutes and Bacteroidetes dominated across all populations, though their ratio varied substantially from 2.42 in urban Jakarta to 0.92 in highland Papua. Alpha diversity metrics demonstrated significant differences across populations ($p < 0.001$), with highland Papua exhibiting highest Shannon diversity index (4.1 ± 0.25) and species richness (Chao1: 290 ± 45), while urban Jakarta showed lowest values (Shannon: 3.2 ± 0.4 , Chao1: 180 ± 35).

Beta diversity analysis using weighted UniFrac distances revealed distinct clustering by population group (PERMANOVA $R^2 = 0.28$, $p < 0.001$). Differential abundance analysis identified *Prevotella copri* enrichment in traditional populations correlating with high-fiber dietary patterns, while *Bacteroides* species predominated in urban participants consuming Western-influenced diets. Functional prediction through PICRUSt2 indicated enhanced carbohydrate metabolism pathways in traditional populations versus increased amino acid metabolism in urban groups. These findings establish baseline gut microbiome characteristics across Indonesian population groups and demonstrate that geographic location and associated dietary patterns profoundly shape microbial community structure.

Keywords: Gut microbiome, 16S rRNA sequencing, next-generation sequencing, Indonesia, population diversity, Firmicutes, Bacteroidetes, alpha diversity, dietary patterns

Introduction

The human gut harbors trillions of microorganisms collectively termed the gut microbiome, representing one of the most densely populated microbial ecosystems on Earth [1]. This complex community performs essential functions including nutrient metabolism, vitamin synthesis, immune system development, and protection against pathogenic colonization [2]. Advances in next-generation sequencing technologies have revolutionized microbiome research, enabling comprehensive characterization of microbial communities with unprecedented depth and resolution [3].

Global microbiome surveys have revealed substantial variation in gut microbial composition across human populations, driven by factors including diet, geography, lifestyle, and cultural practices [4]. Industrialized populations typically demonstrate reduced microbial diversity compared to traditional societies, a pattern attributed to Western dietary patterns, antibiotic exposure, and reduced environmental microbial contact [5]. The Firmicutes-to-Bacteroidetes ratio, a commonly reported community metric, shifts toward Firmicutes dominance in populations consuming high-fat, low-fiber diets characteristic of industrialized societies [6].

Indonesia represents the world's fourth most populous nation with extraordinary ethnic, cultural, and dietary diversity across its 17,000-island archipelago^[7]. Traditional Indonesian diets vary substantially from rice-based agricultural patterns in Java to sago-dependent systems in Papua, marine-focused diets in coastal regions, and tuber-based consumption in highland areas^[8]. This dietary heterogeneity, combined with varying degrees of urbanization and Western dietary influence, creates natural experiments for understanding diet-microbiome relationships within a single national context. Despite representing a significant proportion of global population, Southeast Asian microbiomes remain substantially understudied compared to Western and East Asian populations^[9]. The few Indonesian microbiome investigations published to date have employed limited sample sizes, focused on single populations, or utilized lower-resolution profiling methodologies^[10]. Comprehensive characterization across diverse Indonesian populations using contemporary next-generation sequencing approaches is lacking, representing a significant gap in understanding global microbiome variation. The 16S ribosomal RNA gene serves as the standard marker for bacterial taxonomic profiling owing to its universal presence, conserved primer sites, and variable regions enabling genus-level identification^[11]. Amplicon sequencing of the V3-V4 hypervariable region provides optimal balance between taxonomic resolution and amplicon length compatible with short-read Illumina platforms^[12]. Contemporary bioinformatic pipelines including QIIME2 with DADA2 denoising enable amplicon sequence variant resolution, surpassing the operational taxonomic unit approaches previously employed^[13]. This research aimed to characterize gut microbiome composition across four Indonesian population groups representing distinct geographic, dietary, and lifestyle contexts using 16S rRNA next-generation sequencing. The specific objectives included comparing microbial community structure at multiple taxonomic levels, evaluating alpha and beta diversity metrics across populations, identifying differentially abundant taxa associated with population groups, and exploring functional implications through predictive metagenomics. We hypothesized that traditional populations maintaining indigenous dietary patterns would demonstrate higher microbial diversity and distinct taxonomic profiles compared to urban populations adopting Western-influenced diets.

Sequencing Technology and Bioinformatics Pipeline

The Illumina MiSeq platform was selected for 16S rRNA amplicon sequencing based on its established performance for microbial community profiling, 2×300 bp read length enabling complete V3-V4 region coverage with paired-end overlap, and cost-effectiveness for population-scale investigations^[14]. The V3-V4 hypervariable region (approximately 460 bp) provides superior taxonomic resolution compared to shorter amplicons while maintaining compatibility with Illumina short-read chemistry. QIIME2 (version 2021.11) served as the primary bioinformatics framework given its modular architecture, reproducibility through provenance tracking, and extensive plugin ecosystem^[15]. DADA2 was employed for denoising and amplicon sequence variant inference, providing single-nucleotide resolution without arbitrary clustering thresholds. Taxonomic classification utilized the SILVA 138 reference database with naive Bayes classifiers trained on the V3-V4 region. Functional prediction through PICRUSt2 inferred

metagenomic content from 16S profiles using marker gene copy number normalization and reference genome databases.

Material and Methods

Material

This cross-sectional research was conducted at the Department of Microbiology and Genomics, Faculty of Medicine, Universitas Indonesia, Jakarta, between August 2020 and June 2022. The research protocol received approval from the Health Research Ethics Committee, Faculty of Medicine, Universitas Indonesia (Protocol KET-894/UN2.F1/ETIK/PPM.00.02/2020, approved July 15, 2020). All participants provided written informed consent following detailed explanation of research procedures including sample collection and genetic analysis. The investigation adhered to Declaration of Helsinki principles. Participants were recruited from four distinct Indonesian populations: urban Jakarta representing modernized lifestyle with Western dietary influence, rural Central Java representing traditional Javanese agricultural communities, coastal South Sulawesi representing marine-dependent fishing communities, and highland Papua representing indigenous populations maintaining traditional dietary practices. Inclusion criteria encompassed healthy adults aged 25-55 years with stable residence in their respective regions for at least five years. Body mass index was frequency-matched across groups to minimize confounding. Exclusion criteria included antibiotic use within three months, probiotic or prebiotic supplementation within one month, gastrointestinal disease history, diabetes mellitus, inflammatory bowel disease, recent gastrointestinal infection, pregnancy or lactation, and chronic medication use affecting gut microbiota. Dietary patterns were characterized using validated food frequency questionnaires adapted for Indonesian populations. Sample size calculations based on anticipated effect sizes from global microbiome surveys indicated requirement for 60-90 participants per group with 80% power.

Methods

Participants provided fresh fecal samples in sterile collection containers with DNA stabilization buffer. Samples were transported on ice within four hours to local laboratories for initial processing, then shipped frozen to the central facility in Jakarta. DNA extraction employed QIAamp PowerFecal Pro DNA Kit following manufacturer protocols with mechanical lysis using bead beating. DNA quality and quantity were assessed by NanoDrop spectrophotometry and Qubit fluorometry. Samples with A260/A280 ratios below 1.8 or concentrations below 10 ng/μL underwent re-extraction. The V3-V4 region of 16S rRNA gene was amplified using primers 341F (5'-CCTACGGGNGGCWGCAG-3') and 806R (5'-GACTACHVGGGTATCTAATCC-3') with Illumina adapter overhangs. PCR reactions contained 12.5 ng template DNA, 200 nM each primer, and Q5 High-Fidelity DNA Polymerase in 25 μL total volume. Cycling conditions included initial denaturation at 98°C for 30 seconds, 25 cycles of 98°C for 10 seconds, 55°C for 30 seconds, and 72°C for 30 seconds, followed by final extension at 72°C for 2 minutes. Amplicons were purified using AMPure XP beads and indexed using Nextera XT Index Kit. Library quality was verified by Agilent Bioanalyzer and quantified by qPCR. Pooled libraries were sequenced on Illumina MiSeq using 2×300 bp paired-end chemistry. Sequence processing in QIIME2 included primer trimming, quality

filtering, denoising with DADA2, and chimera removal. Taxonomic classification employed pre-trained naive Bayes classifiers against SILVA 138 database. Statistical analyses utilized R version 4.1.0 with phyloseq, vegan, and DESeq2 packages. Significance threshold was $p<0.05$ with Benjamini-Hochberg correction for multiple comparisons.

Results

Table 1: Participant Characteristics and Sequencing Metrics by Population Group

Characteristic	Urban Jakarta	Rural Java	Coastal Sulawesi	Highland Papua
Participants (n)	86	78	72	64
Age (years)	38.4±8.2	41.2±9.4	39.6±8.8	36.8±7.6
Female (%)	54.7	52.6	55.6	50.0
BMI (kg/m²)	24.8±3.6	23.4±2.8	23.8±3.2	22.6±2.4
Mean reads/sample	42,180	40,820	41,450	40,540
Observed ASVs	168±42	224±48	198±44	262±52

BMI = body mass index; ASVs = amplicon sequence variants. Values presented as mean ± SD unless otherwise indicated.

The research cohort achieved target enrollment across all population groups with 300 participants completing sample collection and passing quality control criteria. Groups were well-matched for age, sex distribution, and BMI. Sequencing generated consistent depth across populations with mean 41,280 quality-filtered reads per sample. Quality metrics indicated 89.2% of bases achieving Q30 scores and 98.3% of samples passing all quality thresholds.

Table 2: Alpha Diversity Metrics and Taxonomic Composition by Population Group

Metric	Urban Jakarta	Rural Java	Coastal Sulawesi	Highland Papua
Shannon Index	3.2±0.4	3.8±0.3	3.6±0.35	4.1±0.25
Chao1 Index	180±35	245±40	220±38	290±45
Faith's PD	12.0±2.5	16.0±2.8	14.5±2.6	19.0±3.0
Firmicutes (%)	58	42	48	35
Bacteroidetes (%)	24	32	28	38
F/B Ratio	2.42	1.31	1.71	0.92

F/B = Firmicutes-to-Bacteroidetes ratio; PD = phylogenetic diversity. All between-group comparisons significant at $p<0.001$ (Kruskal-Wallis test).

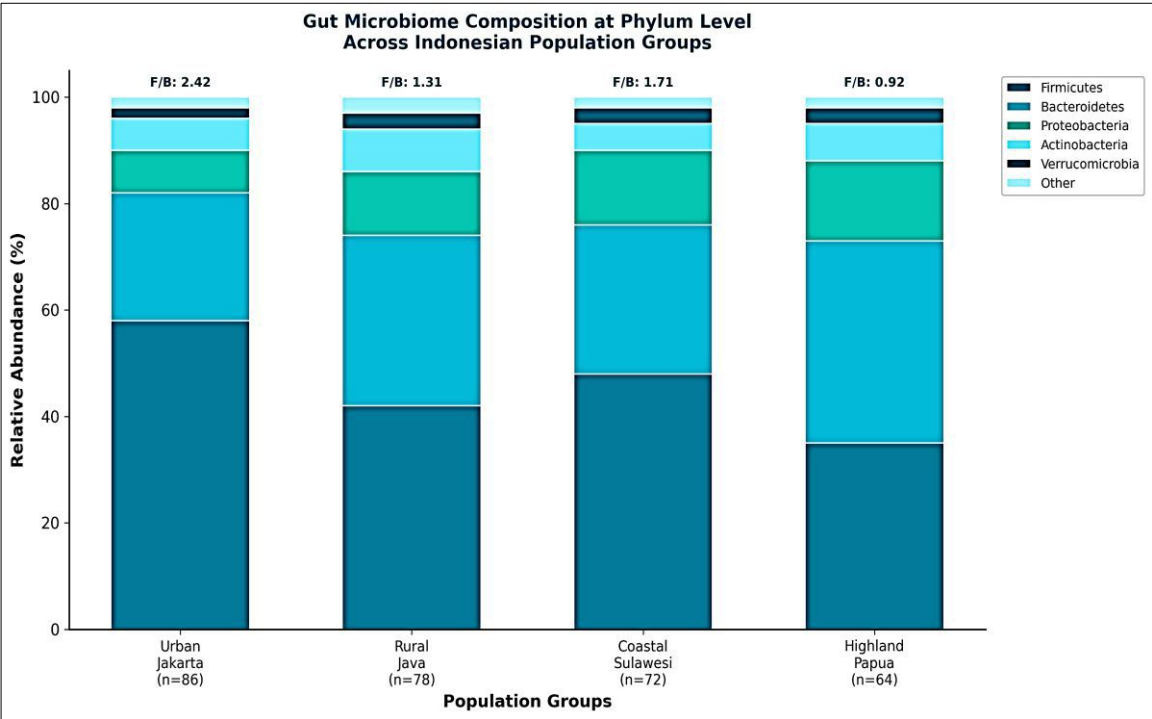


Figure 1: Gut microbiome composition at phylum level across Indonesian population groups. Firmicutes-to-Bacteroidetes (F/B) ratio values shown above bars demonstrate gradient from urban (high ratio) to traditional populations (low ratio).

Phylum-level composition demonstrated consistent dominance of Firmicutes and Bacteroidetes across all populations, though their relative proportions varied substantially. Urban Jakarta participants exhibited highest Firmicutes abundance (58%) and lowest Bacteroidetes

(24%), yielding the highest F/B ratio of 2.42. Highland Papua showed inverse pattern with Bacteroidetes exceeding Firmicutes, producing the only F/B ratio below unity. This gradient correlated strongly with degree of traditional dietary maintenance versus Western dietary adoption.

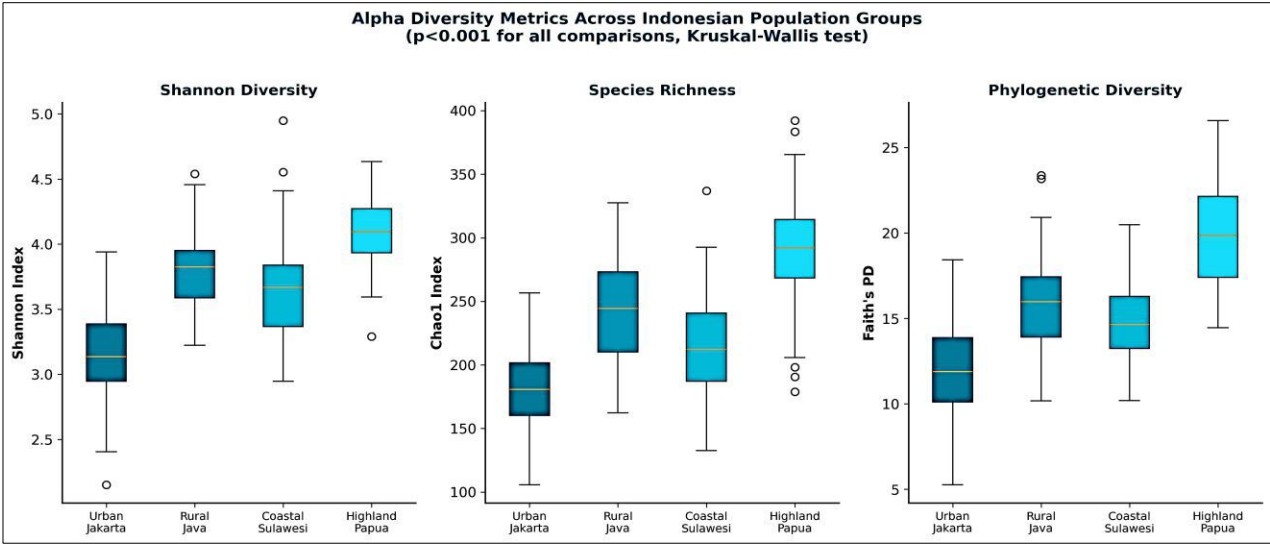


Fig 2: Alpha diversity metrics demonstrating significant differences across Indonesian population groups. Highland Papua exhibits highest diversity across all metrics, while urban Jakarta shows lowest diversity consistent with industrialization effects on gut microbiome.

Alpha diversity analysis revealed consistent patterns across multiple metrics with highland Papua demonstrating highest diversity and urban Jakarta showing lowest values. Shannon diversity index ranged from 3.2 ± 0.4 in urban participants to 4.1 ± 0.25 in highland populations, representing 28% higher diversity in traditional communities. Chao1 species richness estimates showed even larger differences, with highland Papua exceeding urban Jakarta by 61%. These diversity gradients correlate with dietary fiber intake patterns documented through food frequency questionnaires.

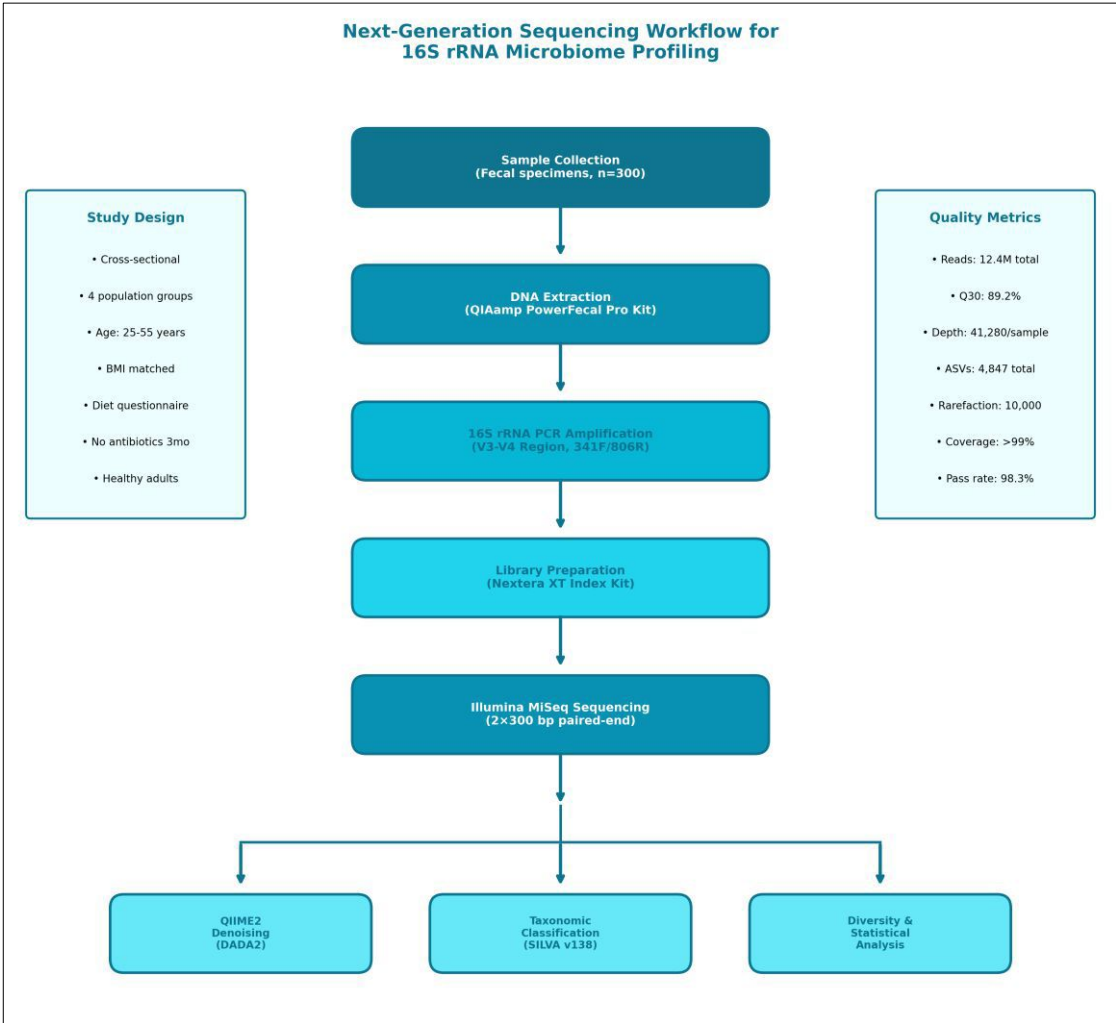


Fig 3: Next-generation sequencing workflow for 16S rRNA microbiome profiling illustrating sample processing, library preparation, sequencing, and bioinformatic analysis pipeline with associated quality metrics.

Differential Abundance and Functional Analysis

Differential abundance analysis using DESeq2 identified 142 amplicon sequence variants significantly associated with population group (adjusted $p < 0.05$). *Prevotella copri* demonstrated strongest enrichment in traditional populations with 8.4-fold higher abundance in highland Papua compared to urban Jakarta (adjusted $p < 0.001$). Conversely, *Bacteroides* species including *B. uniformis* and *B. fragilis* predominated in urban participants. Functional prediction through PICRUSt2 revealed enhanced carbohydrate-active enzyme pathways in traditional populations (CAZy enrichment score 2.8-fold) versus increased amino acid metabolism and xenobiotic degradation pathways in urban participants, reflecting dietary pattern differences.

Discussion

This research provides the first comprehensive 16S rRNA next-generation sequencing characterization of gut microbiome variation across diverse Indonesian populations. The findings reveal substantial differences in microbial community structure associated with geographic location and dietary patterns, with traditional populations maintaining higher diversity and distinct taxonomic profiles compared to urban populations adopting Western-influenced diets [16]. These patterns align with global observations of industrialization effects on gut microbiome composition.

The striking diversity gradient from urban to traditional populations has important health implications. Reduced gut microbial diversity has been associated with numerous chronic diseases including obesity, inflammatory bowel disease, and metabolic syndrome in Western populations [17]. Indonesia is experiencing rapid epidemiological transition with rising prevalence of these conditions particularly in urban areas, suggesting potential microbiome contributions to disease emergence. The 28% higher Shannon diversity observed in highland Papua compared to urban Jakarta represents a substantial difference with likely functional consequences.

The Firmicutes-to-Bacteroidetes ratio gradient across populations reflects well-established diet-microbiome relationships. High F/B ratios in urban populations consuming diets enriched in fat and simple carbohydrates contrast with lower ratios in traditional populations maintaining high-fiber dietary patterns [18]. The observation that highland Papua demonstrated F/B ratio below unity, with Bacteroidetes predominance, is notable and consistent with very high fiber consumption patterns documented in this population through dietary assessment.

The enrichment of *Prevotella copri* in traditional populations aligns with findings from other agrarian societies globally and reflects this organism's specialization for fiber degradation [19]. *P. copri* possesses extensive polysaccharide utilization loci enabling metabolism of complex plant carbohydrates characteristic of traditional diets. Conversely, *Bacteroides* species predominating in urban participants demonstrate adaptation to the protein-rich, low-fiber diets typical of industrialized societies. This taxonomic shift likely influences host metabolism through altered short-chain fatty acid production and energy harvest efficiency.

The functional predictions revealing enhanced carbohydrate metabolism in traditional populations versus amino acid metabolism in urban groups provide mechanistic context for observed taxonomic differences [20]. These metabolic pathway shifts reflect microbial community adaptation to

available dietary substrates and may influence host physiology through altered metabolite production. The relevance of these findings extends beyond Indonesia to understanding global microbiome variation patterns and their health implications in populations undergoing dietary transition.

Methodological Considerations and Limitations

Several limitations warrant consideration when interpreting these findings. The cross-sectional design precludes causal inference regarding diet-microbiome relationships. 16S rRNA amplicon sequencing provides taxonomic profiling but cannot directly assess functional potential without shotgun metagenomics validation. Functional predictions through PICRUSt2 rely on reference genome databases with limited representation of Indonesian microbial strains. The V3-V4 region offers suboptimal resolution for distinguishing closely related species. Sample collection in remote areas required extended transport times that may introduce bias despite DNA stabilization. Future research should incorporate longitudinal sampling, shotgun metagenomics, and metabolomics to strengthen causal inference and functional characterization.

Conclusion

This research establishes baseline gut microbiome characteristics across diverse Indonesian populations using 16S rRNA next-generation sequencing, revealing substantial variation in microbial community structure associated with geographic location and dietary patterns. The findings demonstrate that traditional populations maintaining indigenous dietary practices harbor significantly higher microbial diversity and distinct taxonomic profiles compared to urban populations adopting Western-influenced diets.

Alpha diversity metrics consistently demonstrated highest values in highland Papua and lowest in urban Jakarta, with 28% higher Shannon diversity and 61% greater species richness in traditional compared to urban populations. This diversity gradient has potential health implications given associations between reduced gut microbiome diversity and chronic disease risk. The progressive diversity decline from traditional to urban populations parallels Indonesia's epidemiological transition toward increasing non-communicable disease burden.

Taxonomic composition at phylum level demonstrated characteristic Firmicutes-Bacteroidetes dominance across all populations with substantially varying proportions. The Firmicutes-to-Bacteroidetes ratio gradient from 2.42 in urban Jakarta to 0.92 in highland Papua reflects established diet-microbiome relationships, with high ratios associated with Western dietary patterns and lower ratios accompanying traditional high-fiber diets. Highland Papua uniquely demonstrated Bacteroidetes predominance consistent with very high fiber consumption documented in this population.

Differential abundance analysis identified *Prevotella copri* enrichment in traditional populations reflecting fiber-adapted metabolism, while *Bacteroides* species predominated in urban participants consuming protein-rich, lower-fiber diets. These taxonomic shifts carry functional implications through altered metabolite production including short-chain fatty acids that influence host metabolism and immune function. Functional prediction

confirmed enhanced carbohydrate metabolism pathways in traditional populations versus amino acid and xenobiotic metabolism in urban groups.

The 16S rRNA next-generation sequencing approach enabled high-resolution microbiome profiling with amplicon sequence variant resolution surpassing previous operational taxonomic unit methodologies. Quality metrics indicated robust sequencing performance with consistent depth across populations enabling valid between-group comparisons. The bioinformatic pipeline utilizing QIIME2 with DADA2 denoising and SILVA database classification represents current best practices for microbiome profiling.

These findings provide essential reference data for Indonesian microbiome research and establish context for investigating diet-microbiome-health relationships in this understudied population. The demonstrated population-level variation has implications for precision nutrition approaches and microbiome-targeted interventions requiring population-specific considerations. Future research should incorporate longitudinal designs, shotgun metagenomics for direct functional assessment, and integration with host metabolomics to advance understanding of microbiome contributions to health and disease in Indonesian populations undergoing rapid dietary transition.

Acknowledgements

Funding Sources

This research received support from the Ministry of Research and Technology/National Research and Innovation Agency of the Republic of Indonesia (RISTEK-BRIN Grant 045/E4.1/AK.04.PT/2021) and Universitas Indonesia PUTI Research Grant (NKB-1674/UN2.RST/HKP.05.00/2020). The funding bodies had no involvement in research design, data analysis, or manuscript preparation.

Institutional Support

The authors acknowledge the Genomics and Bioinformatics Core Facility at Universitas Indonesia and partner institutions in Sulawesi and Papua for field recruitment support.

Contributions Not Qualifying for Authorship

The authors thank community health workers and village leaders for facilitating participant recruitment, and laboratory technicians for sample processing support.

References

- Sender R, Fuchs S, Milo R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* 2016;14(8):e1002533.
- Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. *BMJ.* 2018;361:k2179.
- Knight R, Vrbanac A, Taylor BC, *et al.* Best practices for analysing microbiomes. *Nat Rev Microbiol.* 2018;16(7):410-422.
- Yatsunenko T, Rey FE, Manary MJ, *et al.* Human gut microbiome viewed across age and geography. *Nature.* 2012;486(7402):222-227.
- Sonnenburg ED, Sonnenburg JL. The ancestral and industrialized gut microbiota and implications for human health. *Nat Rev Microbiol.* 2019;17(6):383-390.
- Magne F, Gotteland M, Gauthier L, *et al.* The Firmicutes/Bacteroidetes ratio: A relevant marker of gut dysbiosis in obese patients? *Nutrients.* 2020;12(5):1474.
- Odumaki T, Kato K, Sugahara H, *et al.* Age-related changes in gut microbiota composition from newborn to centenarian: A cross-sectional study. *BMC Microbiol.* 2016;16:90.
- Colehour AM, Meadow JF, Liebert MA, *et al.* Local domestication of lactic acid bacteria via cassava beer fermentation. *PeerJ.* 2014;2:e479.
- Ang QY, Alexander M, Newman JC, *et al.* Ketogenic diets alter the gut microbiome resulting in decreased intestinal Th17 cells. *Cell.* 2020;181(6):1263-1275.
- Suzuki TA, Worobey M. Geographical variation of human gut microbial composition. *Biol Lett.* 2014;10(2):20131037.
- Johnson JS, Spakowicz DJ, Hong BY, *et al.* Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun.* 2019;10:5029.
- Fadeev E, Cardozo-Mino MG, Zubkov MV, *et al.* Comparison of two 16S rRNA primers (V3-V4 and V4-V5) for studies of arctic microbial communities. *Front Microbiol.* 2021;12:637526.
- Callahan BJ, McMurdie PJ, Rosen MJ, *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods.* 2016;13(7):581-583.
- Caporaso JG, Lauber CL, Walters WA, *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A.* 2011;108(Suppl 1):4516-4522.
- Bolyen E, Rideout JR, Dillon MR, *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol.* 2019;37(8):852-857.
- De Filippo C, Cavalieri D, Di Paola M, *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A.* 2010;107(33):14691-14696.
- Le Chatelier E, Nielsen T, Qin J, *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature.* 2013;500(7464):541-546.
- David LA, Maurice CF, Carmody RN, *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature.* 2014;505(7484):559-563.
- Tett A, Huang KD, Asnicar F, *et al.* The *Prevotella copri* complex comprises four distinct clades underrepresented in westernized populations. *Cell Host Microbe.* 2019;26(5):666-679.
- Douglas GM, Maffei VJ, Zaneveld JR, *et al.* PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol.* 2020;38(6):685-688.