



OPEN Gut microbiota profiles of peninsular Malaysian populations are associated with urbanization and lifestyle

Nurul Fauzani Jamaluddin^{1,11,12}, Olga Brovkina^{2,3,12}, Nurshamimi Nor Rashid^{1,4}, Anis Rageh Al-Maleki⁵, Yvonne Ai-Lian Lim^{1,6,11}, Maw Pin Tan^{1,4,11}, Soo Ching Lee^{1,6,11}, Claire Duvallet⁷, Christopher H. Corzett^{8,11}, Eric Alm^{7,11}, Mathieu Groussin^{2,7,11}, Mathilde Poyet^{7,9,11}✉ & Fatimah Ibrahim^{1,10,11}✉

There is increasing evidence of distinct gut microbiome compositions between populations living industrialized and non-industrialized lifestyles worldwide. However, whether populations of Malaysia exhibit variations in their microbiome, and to what extent host lifestyle correlates with these variations, remains unclear. Malaysia's extensive geographical and sociocultural diversity provides a unique opportunity to explore how lifestyle and environmental exposures are associated with the human gut microbiome. Here, we characterized the gut microbiome of three populations in peninsular Malaysia, each representing different lifestyle contexts, and identified host factors associated with microbiome variation. Our findings suggest that lifestyle-related factors are strongly associated with differences in microbial community composition across populations. In particular, urban and rural individuals harbor gut microbiota with distinct community structures. We further identified specific taxa as potential microbial signatures of host lifestyle, with the genera *Prevotella* and *Cryptobacteroides* enriched in rural populations, while *Phocaeicola*, *Vescimonas*, and *Megasphaera* were more prevalent among urban individuals. In addition to lifestyle, demographic factors such as age, sex, and BMI were also associated with variation in the gut microbiome. This study highlights the influence of urbanization, lifestyle, and diet on the gut microbiome landscape of Malaysian populations and underscores the importance of considering sociocultural context in future microbiome research.

Keywords Gut microbiota, Urbanization, Industrialization, Indigenous people, 16s rRNA sequencing, Alpha diversity, Beta diversity, Lifestyle variations, Malaysia

The human gut microbiota refers to a group of microorganisms that reside in the gastrointestinal tract of humans, mainly in the colon. These microorganisms, including bacteria, viruses, fungi, and other microbes, play an essential role in various physiological processes, contributing to overall human health. Studies have shown that the diversity and composition of the gut microbiota can vary between urban and rural populations, potentially influenced by factors such as diet, lifestyle, environmental exposure, and access to healthcare^{1–7}. For instance, investigations of rural and urban communities in Peru⁸, South Africa⁴ and Cameroon⁹ showed reduced diversity and shifts in gut microbial profiles of urban vs. rural host communities.

¹Center for Innovation in Medical Engineering, Faculty of Engineering, Universiti Malaya, 50603 Kuala Lumpur, Malaysia. ²Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany. ³Bioinformatic Platform, Kiel University, Kiel, Germany. ⁴Department of Molecular Medicine, Faculty of Medicine, Universiti Malaya, 50603 Kuala Lumpur, Malaysia. ⁵Department of Medical Microbiology, Faculty of Medicine, Universiti Malaya, 50603 Kuala Lumpur, Malaysia. ⁶Department of Parasitology, Faculty of Medicine, Universiti Malaya, 50603 Kuala Lumpur, Malaysia. ⁷Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, USA. ⁸University of Southern California, Los Angeles, USA. ⁹Institute of Experimental Medicine, Kiel University, Kiel, Germany. ¹⁰Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, 50603 Kuala Lumpur, Malaysia. ¹¹Global Microbiome Conservancy Consortium, <http://microbiomeconservancy.org>. ¹²Nurul Fauzani and Olga Brovkina have contributed equally as co-first authors. ✉email: m.poyet@iem.uni-kiel.de; fatimah@um.edu.my

Malaysia is an ideal location for conducting human microbiome studies. Its ethnic and lifestyle diversity, rural–urban gradient and traditional practices and diet offer unique opportunities to investigate the breadth and factors of natural variations in the gut microbiome. As per 2020, the majority of the population consists of Malay, Chinese, and Indian, with about 11% comprising indigenous people including Sabah and Sarawak. The indigenous people in Peninsular Malaysia, also known as Orang Asli, are the native inhabitants of the Malay Peninsular region. They have a rich cultural heritage, unique customs, and subsistence strategies that are closely tied to the land and forests. They aim to preserve the land and culture that they have inherited. They have historically been categorized into three main groups, i.e. Semang, Senoi, and Proto Malay, and consist of 18 ethnicities^{7,10,11}. These historic categories should be treated with caution. These categories were defined during British colonial time as administrative tools that do not reflect the social, cultural and linguistic complexity of local indigenous populations, and are not always recognized by the Orang Asli themselves (Colin¹²). According to the Department of Orang Asli Development (JAKOA), the Orang Asli community comprises 215,215 people out of a population of about 33.7 million people in Malaysia, in 2023. Each ethnicity has unique languages, knowledge systems, and beliefs¹³. Their lifestyle and means of subsistence vary depending on community culture and location⁷. Indigenous communities residing in coastal and lake areas predominantly engage in fishing as their primary occupation, whereas those living in close proximity to or within forested areas rely on hill rice cultivation, hunting, and gathering for sustenance. Some of the indigenous people practice permanent agriculture and manage their own farms, and some practice commercial trade of petai, durian, rattan, and resin products to earn income¹³.

So far, only a few studies characterized the gut microbiome of indigenous populations in Malaysia^{14–17}. Chong et al.¹⁸ demonstrated a close relationship between ethnicity-related socioeconomic disparities and gut microbial compositions in Malaysian children. Dwiyanto et al.¹⁹ found that ethnicity significantly influences the gut microbiota composition of individuals, even among those residing in the same geographical location. This variation is attributed to the distinct dietary lifestyles adopted by each ethnic group in Malaysia. Recently, Yeo et al.⁷ investigated the gut microbiota and cardiometabolic health of Orang Asli and showed that the urban indigenous people from ethnic Temuan suffered from poorer cardiometabolic health. They also found that the gut microbiota in semi-nomadic Jahai people had a high alpha diversity compared to the urban Temuan community. Their data also indicated that there was a distinct gut microbiota composition between the rural and urban indigenous people. Tee et al.¹⁷ conducted a comparative analysis of the gut microbiome in indigenous populations residing in urban versus village environments, identifying a higher prevalence of helminth infections among village residents. This disparity likely reflects differences in environmental exposure and lifestyle. This observation is consistent with the hygiene hypothesis, originally proposed by Strachan²⁰, which suggests that reduced exposure to diverse microbes in early life—due to improved hygiene and sanitation—may increase the risk of allergic and autoimmune diseases. These findings highlight how industrialized lifestyles may indirectly shape the gut microbiome through decreased microbial exposures. Yet, it is worth noting that, compared to urban and/or industrialized populations, indigenous and/or rural populations show a higher prevalence of certain diseases that are also known to be related to the microbiome, particularly for cases of soil-transmitted helminths (STH) and malnutrition. In addition, as indigenous communities change lifestyle and socioeconomic status, they are also increasingly affected by non-communicable diseases¹¹. In this context, our current knowledge on the diversity of gut microbiota between rural and urban populations in Malaysia remains limited, and warrants further investigation.

Here, we performed a multi-ethnic investigation of the diversity and composition of the gut microbiome of Malaysian populations living diverse lifestyles, in three geographic locations (Tasik Banding, Gua Musang and Kuala Lumpur). Our study reveals strong differences in microbiome features between rural and urban communities in Malaysia. We also observed signatures of potentially-ongoing microbiome perturbations in a rural population as a result of recent changes in lifestyle.

Results

Participant demographics and lifestyles

A total of 87 participants were recruited for this study as part of the Global Microbiome Conservancy (<http://microbiomeconservancy.org>) across three localities: Gua Musang (n = 20), Tasik Banding (n = 30), and Kuala Lumpur (n = 37) (Fig. 1). Gua Musang is an isolated rural location, situated in the deeper parts of forested areas, where communities can live nomadically, engaging in hunting and foraging subsistence activities, and having little exposure to industrialized products^{21,22}. Participants from Gua Musang (age = 32.5 ± 11.8 SD (standard deviation), body mass index (BMI) = 19.4 ± 1.77 SD) self-declare as from the Batek and Mendrik ethnic groups (n = 18 and n = 1, respectively). Statistical comparisons in Table 1 were adjusted for multiple testing using the Benjamini–Hochberg procedure to control the false discovery rate (FDR).

Tasik Banding is an isolated rural area that is part of the Belum-Temengor forest reserves. Human communities live semi-nomadically, and are mostly dependent on natural resources for their sustenance and livelihoods, engaging in farming, hunting, fishing, and gathering activities. They also receive food support from the Malaysian government including rice meal once or twice a week. The presence of an eco-tourism industry in Tasik Banding also adds a unique dynamic to this rural setting. Tasik Banding participants (age = 33.6 ± 11.6 SD, BMI = 22.8 ± 4.04 SD) self-declare to be from Jahai (n = 28) and Temiar (n = 3) ethnicities (Table 1).

In Kuala Lumpur, the urban communities are exposed to industrialization and have adopted lifestyles heavily relying on processed and mass-produced food products associated with industrial food production methods. Kuala Lumpur participants (age = 23.8 ± 2.85 SD, BMI = 21.4 ± 4.04 SD) are from Chinese (n = 21), Indian (n = 9), Malay (n = 5), Dusun (n = 1) and Thai (n = 1) ethnicities (Table 1).

Apart from the data collected from three different locations with different industrialization backgrounds, Table 1 also highlights the differences in the access to electricity and drinking water, showing that no participants



Fig. 1. Geographical and lifestyle context of the study populations. The map shows the three study locations in Peninsular Malaysia: Kuala Lumpur (urban), Gua Musang (rural), and Tasik Banded (rural). Photographs depict general lifestyle scenes from the three study locations. The top right image represents Tasik Banded, where indigenous Orang Asli communities live in a rural environment with subsistence practices such as traditional hunting. The bottom left image depicts Kuala Lumpur, a highly urbanized setting characterized by dense infrastructure and access to diverse food markets. The bottom right image represents Gua Musang, a rural locality where Orang Asli communities continue traditional foraging and semi-nomadic lifestyles. The map of Peninsular Malaysia was created using BioRender.com under an academic license.

from Gua Musang and Tasik Banded have access to electricity. A substantial proportion of participants from Gua Musang (42.1%) and Tasik Banded (80.7%) relied on untreated and unfiltered water sources, while a smaller fraction transitioned to filtered water (26.3% and 12.9%, respectively). In contrast, most participants from Kuala Lumpur had access to treated and filtered water (67.6%), with only a minority still relying on untreated water (5.4%). This study also shows that some of participants from Gua Musang reside in homes with various floor types, including dirt (26.3%), wooden (42.1%), and concrete (31.6%) floors, while none reported living in houses with covered flooring.

Level of industrialization of lifestyle impacts gut microbiome diversity and composition

We first examined the association of location, urbanism, and sex factors to microbiome alpha diversity (Fig. 2).

We found that rural populations exhibited significantly higher phylogenetic diversity compared to urban populations (Fig. 2A, Faith's PD index, Wilcoxon test, $p=4.5e-3$). However, no significant difference was detected when using the Shannon diversity index (Fig. 2D, Wilcoxon test, $p=0.89$), suggesting that while overall species richness and evenness remain comparable, rural populations harbor a more phylogenetically diverse microbiome. Further comparisons across localities revealed significant differences in alpha diversity (Fig. 2B, Faith's PD, Wilcoxon test, $p=1.9e-3$). The gut microbiome of individuals from Gua Musang exhibited significantly higher phylogenetic diversity than Kuala Lumpur. Individuals from Tasik Banded also had higher diversity than those in Kuala Lumpur, though the effect was less pronounced. When analyzing Shannon diversity, no significant differences were observed across localities (Fig. 2E, Wilcoxon test, $p>0.05$), reinforcing that phylogenetic diversity, rather than species richness, distinguishes microbiomes across urban and rural settings. Finally, we found no significant differences in microbiome diversity between males and females for Shannon diversity (Fig. 2C, Wilcoxon test, $p=0.13$ and Fig. 2F, Wilcoxon test, $p=0.092$), suggesting that sex does not strongly influence alpha diversity in this cohort.

Next, we investigated differences in microbial compositions between lifestyles and localities (Fig. 3).

We found significant differences in microbiome compositions between rural and urban communities (Fig. 3A, Adonis, p -value <0.001 , Supplementary Table 1). These compositional differences may be attributed to dietary habits, environmental exposures, and other lifestyle-related factors. Significant differences in beta

		Gua Musang (n = 19)	Tasik Banding (n = 31)	Kuala Lumpur (n = 37)	Total (n = 87)	Adj. p-value
Urbanism, n (%)	Rural	19 (100.0)	31 (100.0)	0 (0.0)	50 (57.5)	
	Urban	0 (0.0)	0 (0.0)	37 (100.0)	37 (42.5)	
Access to Electricity	Yes	0 (0.0)	0 (0.0)	37 (100.0)	37 (42.5)	
	No	19 (100.0)	31 (100.0)	0 (0.0)	50 (57.5)	
Subsistence, n (%)	Hunter-gather	19 (100.0)	0 (0.0)	0 (0.0)	19 (21.8)	
	Hunter-gather to farmer	0 (0.0)	31 (100.0)	0 (0.0)	31 (35.6)	
	Industrialized	0 (0.0)	0 (0.0)	37 (100.0)	37 (42.5)	
Ethnicity, n (%)	Batek	18 (94.7)	0 (0.0)	0 (0.0)	18 (20.7)	5.35e-28
	Chinese	0 (0.0)	0 (0.0)	21 (56.8)	21 (24.1)	
	Dusun	0 (0.0)	0 (0.0)	1 (2.7)	1 (1.1)	
	Indian	0 (0.0)	0 (0.0)	9 (24.3)	9 (10.3)	
	Jahai	0 (0.0)	28 (90.3)	0 (0.0)	28 (32.1)	
	Malay	0 (0.0)	0 (0.0)	5 (13.5)	5 (5.7)	
	Mendriq	1 (5.3)	0 (0.0)	0 (0.0)	1 (1.1)	
	Temiar	0 (0.0)	3 (9.7)	0 (0.0)	3 (3.5)	
Sex, n (%)	Male	11 (57.9)	18 (58.1)	21 (56.8)	50 (57.5)	1
	Female	8 (42.1)	13 (41.9)	16 (43.2)	37 (42.5)	
Water Source, n (%)	Untreated & Unfiltered	8 (42.1)	25 (80.7)	0 (0.0)	33 (37.9)	4.48e-12
	Untreated & Unfiltered to Untreated & Filtered	5 (26.3)	4 (12.9)	0 (0.0)	9 (10.3)	
	Untreated & Filtered	6 (31.6)	2 (6.5)	25 (67.6)	33 (37.9)	
	Untreated & Filtered to Treated & Filtered	0 (0.0)	0 (0.0)	10 (27.0)	10 (11.5)	
	Treated & Filtered	0 (0.0)	0 (0.0)	2 (5.4)	2 (2.3)	
Floor Type, n (%)	Dirt	5 (26.3)	0 (0.0)	0 (0.0)	5 (5.7)	3.11e-15
	Wooden	8 (42.1)	30 (96.8)	2 (5.4)	40 (46.0)	
	Concrete	6 (31.6)	1 (3.2)	25 (67.6)	32 (36.8)	
	Covered	0 (0.0)	0 (0.0)	10 (27.0)	10 (11.5)	
Bristol stool scale, n (%)	Type 1	2 (10.5)	0 (0.0)	1 (2.7)	3 (3.4)	7.13e-03
	Type 2	5 (26.3)	0 (0.0)	2 (5.4)	7 (8.0)	
	Type 3	5 (26.3)	4 (12.9)	4 (10.8)	13 (14.9)	
	Type 4	5 (26.3)	11 (35.4)	17 (46.0)	33 (37.9)	
	Type 5	1 (5.3)	4 (12.9)	7 (18.9)	12 (13.8)	
	Type 6	1 (5.3)	6 (19.4)	4 (10.8)	11 (12.6)	
	Type 7	0 (0.0)	6 (19.4)	2 (5.4)	8 (9.1)	
Age, mean(SD)		32.5(±11.8)	33.6(±11.6)	23.8(±2.85)		1.43e-3
BMI, mean(SD)		19.4(±1.77)	22.8(±4.04)	21.4(±4.04)		5.24e-3

Table 1. Demographic characteristics of study participants.

diversity compositions were also found when comparing cohorts by locality (Fig. 3B, Adonis, p-value = 0.001, Supplementary Table 2) and when comparing the two rural cohorts together (Fig. 3B), suggesting that environmental and host characteristics specific to each locality drive microbial compositions beyond the effect of population density (urban vs. rural).

Building on these findings, we next examined how host and environmental factors associate with microbiome compositions (Fig. 4), while accounting for confounders. For this, we used redundancy analysis (RDA) and a stepwise forward model selection approach, and we found that locality is the primary predictor of microbial variations (adjusted $R^2 = 0.056$), followed by age ($R^2 = 0.012$), Bristol stool scale ($R^2 = 0.008$), and BMI ($R^2 = 0.003$). Locality was found to be significantly associated with microbiota compositions when also considering these other factors ($p = 0.002$). However, the other factors did not significantly improve model performance once locality had already been accounted for. The final RDA model constrained 7.75% of the total variation in gut microbiome composition, indicating that while locality and host-related factors influence microbial diversity, a substantial proportion of variation remains unexplained. Additionally, the RDA (Fig. 4) shows an inverse association between Kuala Lumpur and age, reflecting the demographic structure of our dataset, where individuals from Kuala Lumpur tended to be younger than those from Gua Musang and Tasik Banding (Table 1).

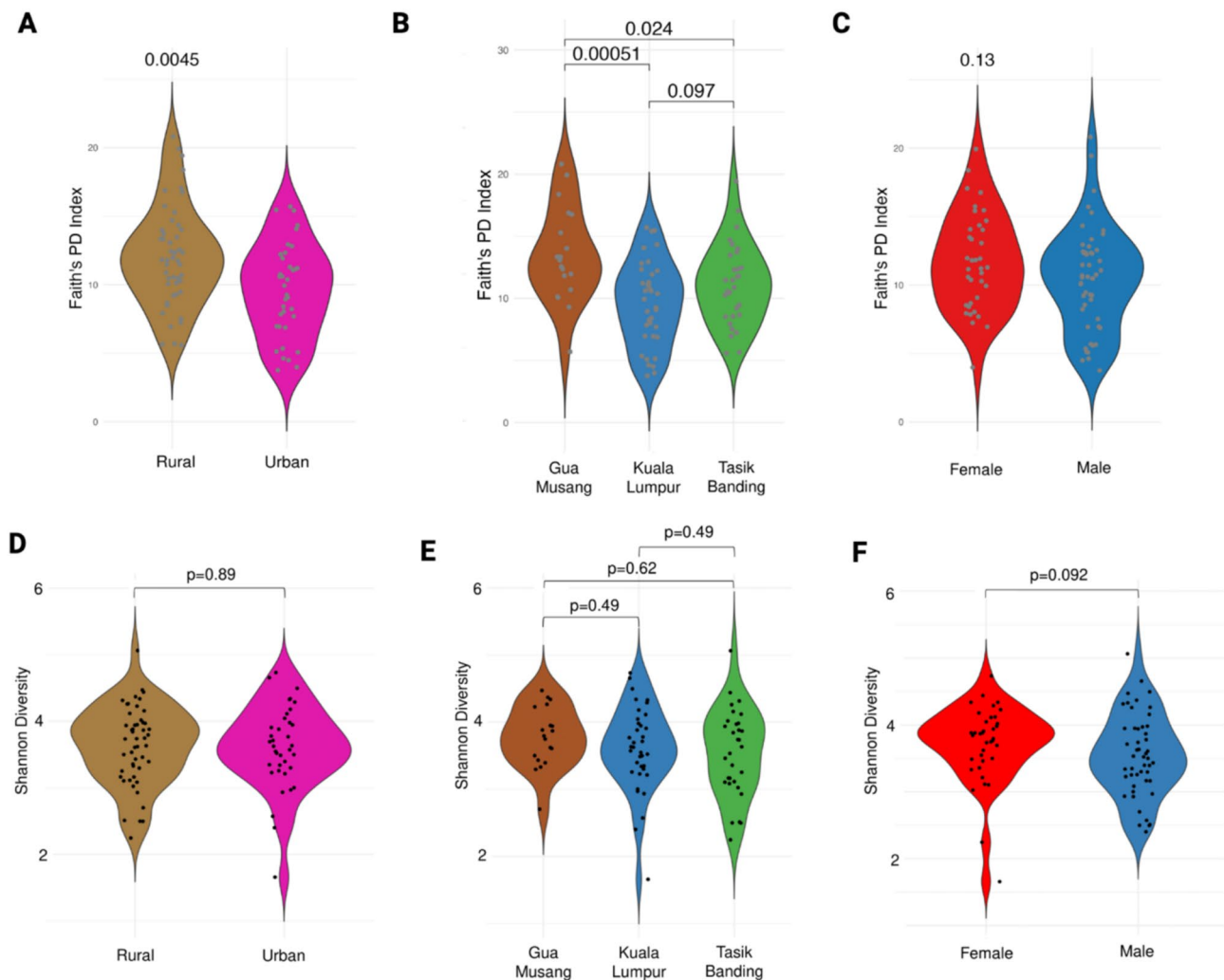


Fig. 2. Variation in alpha diversity by urbanization, locality, and sex. This plots present comparisons of gut microbiome alpha diversity across different population groups using Faith's Phylogenetic Diversity (PD) index and Shannon diversity index. Violin plots illustrate differences in Faith's PD index by urbanization (A), locality (B), and sex (C). Rural populations exhibit significantly higher phylogenetic diversity compared to urban populations (A Wilcoxon test, $p=0.0045$). Across localities, Gua Musang shows the highest diversity (B Kruskal–Wallis test, $p=0.0019$), with significant pairwise differences between Gua Musang and Kuala Lumpur ($p=5.1e-4$) and Tasik Banding and Kuala Lumpur ($p=2.4e-3$). No significant differences were observed between sexes (C, Wilcoxon test, $p=0.13$). Violin plots of Shannon diversity index by urbanization (D), locality (E), and sex (F) reveal no significant differences (Wilcoxon tests, all $p>0.05$).

Individual taxonomic profiles vary across lifestyles and localities

We next examined the relative abundance of individual gut microbial lineages across different taxonomic levels to look for associations with urbanization and locality (Figs. 5 & 6).

At the phylum level (Fig. 5A), *Firmicutes* and *Bacteroidota* were the dominant bacterial phyla across all groups. At the genus level (Fig. 5B), we observed that *Prevotella* was more abundant in rural populations, while levels of *Bacteroides* and *Phocaeicola* are elevated in urban samples. We found that the Firmicutes/Bacteroidetes (F/B) ratio does not differ significantly between urban and rural populations (Wilcoxon test, $p=0.16$) (Fig. 5C), nor among localities, suggesting that overall community balance between these two dominant phyla remains relatively stable across environments. In contrast, the Bacteroides/Prevotella ratio (Fig. 5D) was significantly lower in rural populations compared to urban populations (Wilcoxon test, $p=0.0062$), confirming a greater relative abundance of *Prevotella* in rural individuals. At the locality level, Kuala Lumpur exhibited a significantly higher Bacteroides/Prevotella ratio compared to both rural localities, Gua Musang ($p=0.0052$) and Tasik Banding ($p=0.0014$).

We next accounted for the effect of host confounders on the variation in abundance between localities. We first searched for differentially abundant taxa (Fig. 6), using MaAsLin2 and a general multivariate linear model (see Methods) to detect amplicon sequence variants (ASVs) that are differentially abundant across locality, urbanism, age, sex and BMI variables. We found that several ASVs show significant enrichment or depletion in

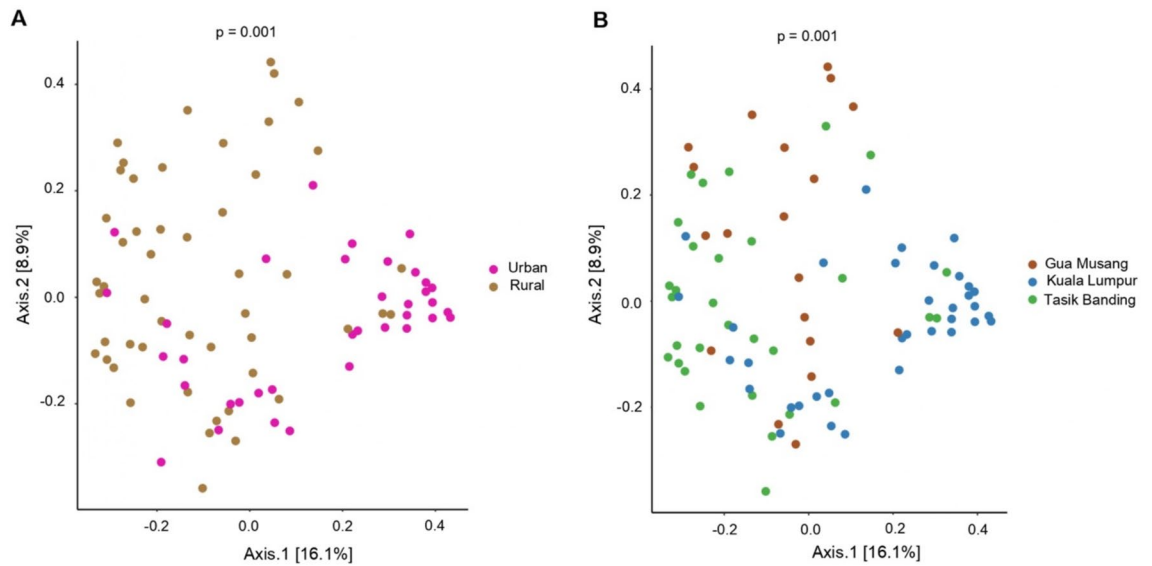


Fig. 3. PCoA Ordination of Bray–Curtis Distances showing microbial community variation by urbanization (A) and locality (B). Plots illustrate differences in gut microbiome composition based on Bray–Curtis distances. A Microbial community composition significantly differs between rural and urban populations (Adonis, $p=0.001$). B Differences in microbiome composition are also observed across localities (Adonis, $p=0.001$).

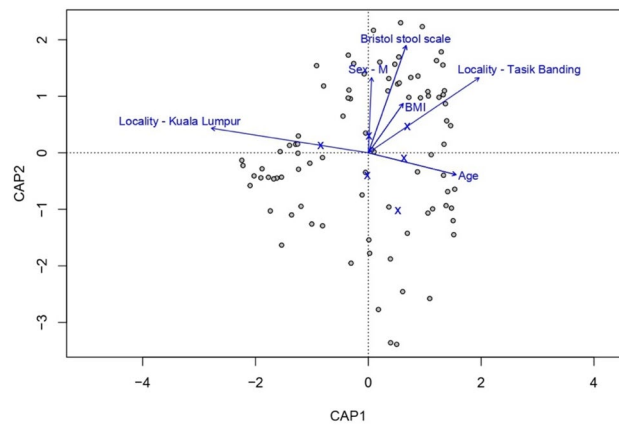


Fig. 4. Redundancy analysis of gut microbiome composition with key explanatory variables. RDA was performed to assess the influence of host and environmental factors on gut microbiome composition. The plot displays the relationship between microbial variation and key explanatory variables, with arrows indicating the direction and strength of associations.

specific localities independent of confounders, indicating that environmental and lifestyle factors strongly shape the abundance profile of individual microbial lineages (Fig. 6A–C). A taxonomic analysis shows that the most differentially abundant ASVs across locality and lifestyle predominantly belong to Bacteroidota and specific Firmicutes classes, such as Bacilli, Negativicutes, and Clostridia (Fig. 6B).

We were able to identify several ASVs down to the species level that show differences in abundance across localities (Fig. 6A, C). Notably, rural locations are characterized by a higher abundance of *Prevotella copri* (ASV_3569, adj. p -val=1.8e-02), a species known to thrive in the gut microbiome of individuals living in non-industrialized or traditional settings, often associated with high-fiber, plant-based diets²³. A *Vescimonas* ASV (ASV_7027) is also showing different abundances across localities, and is elevated in Gua Musang (adj. p -val=1.8e-02). In contrast, *Parabacteroides merdae* (ASV_4792) is more abundant in the urban Kuala Lumpur samples²⁴. The latter species is involved in the catabolism of branched-chain amino acids and has been shown to improve insulin sensitivity and glucose tolerance, reflecting a microbiome adapted to a more urban lifestyle with potentially more processed food intake. Similarly, *Phocaeicola vulgatus* (ASV_5319, adj. p -val=1.8e-02), another species prevalent in Kuala Lumpur, is a common member of the human gut microbiome with a broad capacity for carbohydrate utilization, aligning well with urban dietary patterns rich in carbohydrates, particularly

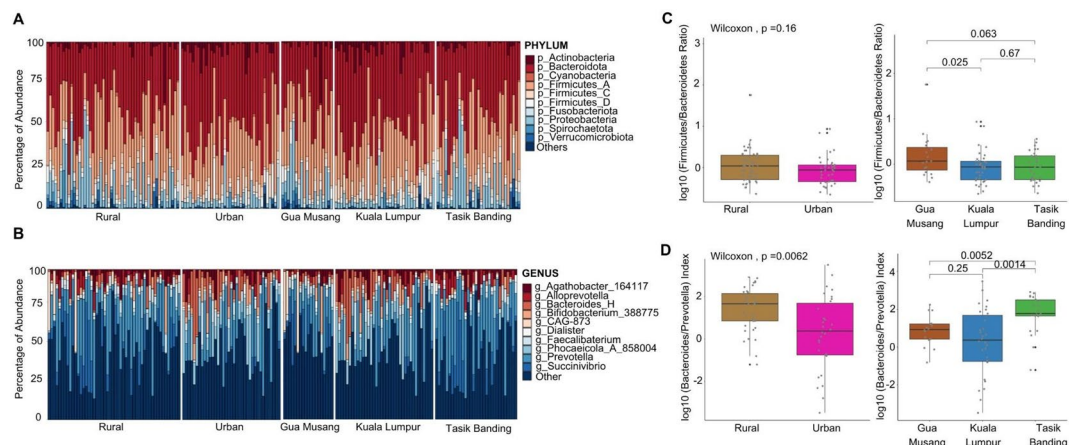


Fig. 5. Taxonomic composition (**A**, **B**) and key microbial ratios across urbanization and locality (**C**, **D**). **A** Phylum-level taxonomic composition of the gut microbiome across urbanization levels (left) and localities (right). **B** Genus-level taxonomic composition stratified by urbanization (left) and locality (right). Rural populations exhibit a higher prevalence of *Prevotella*, while urban populations show an increased abundance of *Bacteroides* and other genera associated with a more industrialized lifestyle. **C** Firmicutes/Bacteroidetes (F/B) ratio, a commonly used indicator of microbiome composition, does not differ significantly between urban and rural populations (Wilcoxon test, $p=0.16$). At the locality level, the ratio remains relatively stable across Gua Musang, Kuala Lumpur, and Tasik BANDING. **D** $\log_{10}(\text{Bacteroides/Prevotella})$ index, a proxy for dietary shifts, reveals a significant difference between urban and rural populations (Wilcoxon test, $p=0.0062$). At the locality level, Kuala Lumpur shows a significantly higher Bacteroides/Prevotella ratio compared to both rural localities, Gua Musang ($p=0.0052$) and Tasik BANDING ($p=0.0014$).

those derived from processed foods²⁵. Additionally, *Megasphaera elsdenii* (ASV_2236, adj. $p\text{-val}=3.5\text{e-}02$), found predominantly in Kuala Lumpur, is well-adapted to utilize lactate and glucose, further supporting the carbohydrate-oriented profile of an urban microbiome²⁶.

To gain further insights into the functional capabilities of these microbiomes, we conducted a pathway prediction analysis and identified significantly differentially enriched pathways (Fig. 6D, Supplementary table 3). Using hierarchical clustering, we found distinct patterns in metabolic pathways between Kuala Lumpur and rural localities. Pathways associated with several biosynthesis pathways, including the menaquinol biosynthesis (PWY-5840, PWY-5838, PWY-5897), phyloquinol biosynthesis (PWY-5863, adj. $p\text{-val}=4.7\text{e-}03$), and demethylmenaquinol-8 biosynthesis (PWY-5861, adj. $p\text{-val}=5.1\text{e-}03$), as well as (Kdo)2-lipid A biosynthesis (KDO-NAGLIPASYN-PWY, adj. $p\text{-val}=5.5\text{e-}03$) and 2-carboxy-1,4-naphthoquinol biosynthesis (PWY-5837, adj. $p\text{-val}=4.7\text{e-}03$), were significantly downregulated in the urban Kuala Lumpur group. These pathways are integral to the production of essential vitamins (e.g., vitamin K2) and bacterial cell wall components. Conversely, some pathways related to degradation and fermentation processes were upregulated in Kuala Lumpur. These included succinate fermentation to butanoate (PWY-5677, adj. $p\text{-val}=5.5\text{e-}02$), pyruvate fermentation to propanoate (P108-PWY, adj. $p\text{-val}=5.2\text{e-}03$), and pentose phosphate pathway (PENTOSE-P-PWY, adj. $p\text{-val}=4.7\text{e-}03$), as well as pathways involved in breaking down complex sugars and polysaccharides, such as chondroitin sulfate degradation (PWY-6572, adj. $p\text{-val}=2.3\text{e-}03$), β -(1,4)-mannan degradation (PWY-7456, adj. $p\text{-val}=1.2\text{e-}02$), D-fructuronate degradation (PWY-7242, adj. $p\text{-val}=4.6\text{e-}03$), and fucose and rhamnose degradation (RHAMCAT-PWY, adj. $p\text{-val}=3.7\text{e-}02$). These findings may reflect the influence of urban diets rich in processed carbohydrates and animal-derived products, which promote pathways associated with sugar metabolism and degradation of dietary glycans.

Discussion

This study provided novel insights into the influence of lifestyle and infrastructure on the gut microbiota across three distinct locations with varying levels of urbanization in Malaysia. A key limitation of this study is the use of 16S rRNA gene sequencing, which allows for taxonomic profiling but only provides inferred functional predictions. These inferences may be less accurate compared to direct functional characterization achievable through shotgun metagenomic sequencing. However, the analysis presented is adequate for capturing broad community-level differences and identifying key taxa associated with urbanization gradients.

Alpha diversity analyses revealed that rural populations exhibit higher phylogenetic diversity compared to urban populations, as evidenced by Faith's PD index. This finding aligns with previous research by Yeo et al.⁷, who observed higher alpha diversity in the semi-nomadic Jahai ethnic group compared to the urban-dwelling Temuan indigenous people. Similarly, Rosas-Plaza et al.²⁷, documented higher alpha diversity in rural populations, particularly among agricultural and hunter-gatherer communities, compared to urban populations. This supports the notion that rural lifestyles foster a more phylogenetically diverse gut microbiome.

However, our results diverge when using the Shannon diversity index, as no significant differences were observed between urban and rural populations. This discrepancy highlights how different diversity metrics

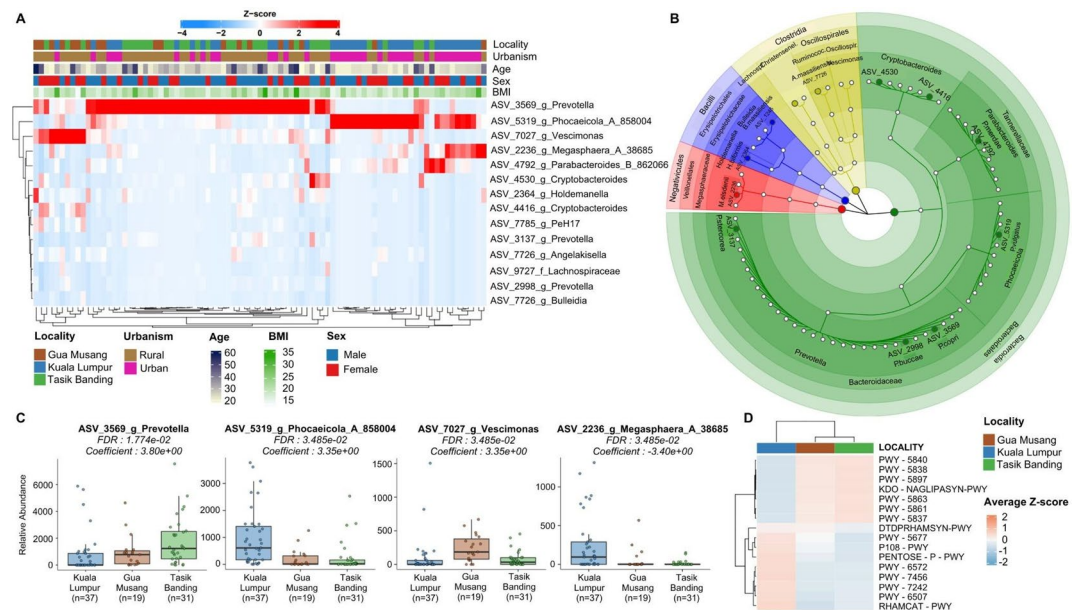


Fig. 6. Differentially abundant taxa (A), their taxonomic classification (B), and the median abundance of the top four differentially abundant taxa (C) are shown. A differential abundance analysis was also performed on predicted functional pathway profiles (D). **A** Heatmap of differentially abundant ASVs across localities. The color scale represents Z-score normalized relative abundance, with red indicating higher abundance and blue indicating lower abundance. Samples are clustered based on microbial composition, revealing distinct taxonomic signatures associated with each locality. The metadata bar on top shows locality, urbanization, age, sex, and BMI. **B** Circular phylogenetic tree generated using GraPhlAn2, displaying the taxonomic relationships of significantly differentially abundant taxa. **C** Relative abundance of the four most differentially abundant taxa across localities. Boxplots show significant differences in *Prevotella*, *Phocaeicola*, *Vescimonas*, and *Megasphaera*, with p-values and effect sizes (coefficients) indicated for each ASV. These taxa are key contributors to microbiome differentiation across populations. **D** The heatmap displays the average Z-scores of all significant differentially abundant predicted functional pathways across localities, which are represented using a hierarchical clustering based on pathway abundance profiles.

capture distinct aspects of microbial communities. While Shannon's index incorporates both species richness and evenness, Faith's PD index reflects phylogenetic breadth. Interestingly, when analyzing diversity by specific localities, only Gua Musang exhibited significantly higher phylogenetic diversity (Faith's PD) compared to Kuala Lumpur, whereas Tasik Banding's diversity profile more closely resembled that of the urban cohort. These findings suggest that microbial diversity is shaped not only by urbanization but also by nuanced, locality-specific environmental and cultural factors, even among rural populations. The reduced alpha diversity observed in Tasik Banding, despite its rural context, may reflect recent lifestyle shifts from a nomadic to a more settled existence, potentially influencing dietary and environmental exposures that shape the gut microbiome. This finding aligns with the demographic and environmental characteristics of these populations (Table 1). In Gua Musang, where individuals continue a foraging lifestyle, dietary diversity and frequent environmental microbial exposure may contribute to their higher microbiome diversity. Meanwhile, Tasik Banding represents a population that has recently transitioned from foraging to farming. Additionally, differences in living conditions could further shape microbial exposure. Gua Musang residents primarily live in homes with earth floors, which may facilitate higher microbial transfer from the environment. In contrast, the majority of individuals in Tasik Banding reside in wooden-floored homes, potentially limiting direct soil-microbe interactions.

Microbiome compositions differed significantly between rural and urban populations, with both Gua Musang and Tasik Banding showing distinct profiles from Kuala Lumpur. The RDA analysis further confirmed locality as a major factor shaping microbial variation, with more minor contributions from age, BMI, and stool consistency. An inverse association between age and microbiome composition in Kuala Lumpur suggests dietary or lifestyle differences among younger individuals in urban settings. The observed trends in microbiome composition between rural and urban localities in Malaysia presented in this study align with findings from studies conducted in other regions and countries. For example, in rural African populations, such as the Hadza hunter-gatherers in Tanzania, the gut microbiome is similarly dominated by *Prevotella* species, which are associated with high-fiber, plant-based diets typical of traditional lifestyles²⁸. These microbial profiles suggest a diet rich in unprocessed plant materials, requiring specialized fermentation capabilities, exemplified by *Prevotella* species, which are well-adapted to high-fiber diets. In contrast, urban populations in the United States and Europe often show a predominance of *Bacteroides* species, which exhibit diverse metabolic capabilities, including both fiber and sugar utilization. Similarly, the urban population in Kuala Lumpur is characterized by species like *Parabacteroides merdae* and *Phocaeicola vulgatus*, both of which are known for their carbohydrate metabolism and adaptation

to more processed diets^{26,29}. These bacteria are adept at metabolizing the more diverse and energy-rich dietary substrates common in urban environments, such as refined carbohydrates, fats, and proteins found in processed and convenience foods.

The function of these microbiome differences is also consistent with global observations. In rural populations, pathways related to amino acid synthesis and fiber fermentation, such as the ones identified in Gua Musang and Tasik Banding (e.g., PWY-5861, PWY-5897), are often enriched due to the high-fiber intake. On the other hand, urban microbiomes, as seen in Kuala Lumpur, often display enrichment in pathways related to the metabolism of simple sugars and carbohydrates, such as fucose and rhamnose breakdown pathways. A similar pattern is reported in urban Chinese populations, where diets high in sugary beverages and processed foods are linked to microbiome profiles enriched with sugar-metabolizing bacteria^{30,31}.

These findings highlight consistency of microbiome shifts driven by urbanization and lifestyle changes. The Malaysian dataset provides an additional perspective, emphasizing the role of local dietary patterns and cultural habits in shaping the microbiome. For instance, Kuala Lumpur's microbiome may also reflect the unique influences of Malaysian cuisine, which includes carbohydrate-rich dishes such as rice and noodles, alongside urban dietary staples like sugary beverages. Comparatively, rural Malaysian populations share similarities with other rural communities globally, showcasing how traditional diets foster a microbiome specialized for fiber degradation and nutrient synthesis.

Conclusion

As such, accounting for the diversity of host environments and lifestyles can provide valuable insights into the complex changes that occur in the gut microbiome in health and disease. It may also inform on the design of tailored low-threshold intervention strategies that promote healthy host-gut microbiome interactions in both rural and urban populations. Our work characterized the variability in gut microbiome profiles across three distinct communities in Malaysia: Kuala Lumpur (urban), Gua Musang (rural), and Tasik Banding (rural). Our findings underscore the importance of establishing a baseline understanding of gut microbiome diversity in healthy Malaysian populations as a foundation for future research. The next critical step is to explore how disease-associated microbiomes compare to this baseline, aiming to identify Malaysia-specific microbial signatures of disease at the population level. Such efforts will require building robust clinical cohorts within low- and middle-income countries, which will provide a crucial framework for understanding microbiome-related health outcomes.

Methods

Ethical clearance

Prior to the start of sample collection, Research & Ethics approvals were obtained from the MIT IRB (protocol #1,612,797,956) and human ethics approval was obtained from the Universiti Malaya Medical Research Ethics Committee (MREC, ID No.: 2018219–6033), which is per the Declaration of Helsinki, and complied with the international and institutional standards. All participants provided oral and written informed consent in their native language directly to the study personnel on site. Additionally, informed consent was obtained to publish the pictures shown in Fig. 1 to illustrate the lifestyle of the communities. Being in these pictures does not reflect participation in the study.

Study location and participants

The first locality involved in this study was Gua Musang. Gua Musang is located in the Northeastern part of Peninsular Malaysia, Kelantan. Indigenous communities in Kuala Koh, Gua Musang consist of Batek and Mendrik ethnicities. They live in the deeper part of forested areas and some communities still live nomadically^{21,22}. They are non-industrialized and still practicing basic traditional economic activities such as the hunter-gatherer lifestyle.

The second locality was Tasik Banding, Perak. Tasik Banding is part of Belum-Temenggor, a grouping of forest reserves in Malaysia. It has existed for over 130 million years and is the World's oldest rainforest. The ecosystem in this area is older than that of both Amazon and Congo. Indigenous people in Tasik Banding consist of the Jahai and Temiar ethnicities. They are non-industrialized. Some of the communities live in semi-nomadic living modes, shifting from hunting and gathering to an agricultural community (gatherer-hunter to farmer)³².

The third study location involved an urban-industrialized community in Kuala Lumpur. This study involved ethnicities from Chinese, Indian, Malay, and Dusun. Communities in Kuala Lumpur are industrialized and practice semi-urban lifestyles. In Kuala Lumpur, the communities are exposed to industrialization and have adopted modern lifestyles, which are completely different from the rural communities in Gua Musang and Tasik Banding.

Collection of lifestyle and dietary habit

In addition to collecting fecal samples, participants were interviewed by study personnel with questionnaire to obtain information on their health, lifestyle, and dietary.

Sample collection

The samples were collected from March 2019 to May 2019. A total of 87 stool samples were collected in this study, which included 20 stools sample from subjects in Gua Musang aged 32.5 ± 11.8 (mean \pm standard deviation) years old, 30 stools samples from subjects in Tasik Banding aged 33.6 ± 11.6 years old, and 37 stools samples from subjects in Kuala Lumpur aged 21.4 ± 3.28 years old. The subjects were given a stool container and advised to defecate directly into the stool container. All stool samples were processed within 1 h after defecation. One

gram of stool was stored in a sterile vial containing 10 mL of RNA later preservative buffer. The samples were then shipped to the United States at the Massachusetts Institute of Technology for processing and long-term preservation within the Global Microbiome Conservancy biobank at -80°C.

DNA extraction and 16S rRNA amplicon sequencing

We extracted the DNA from each sample using the PowerSoil DNA Isolation Kit (Qiagen) following the manufacturer protocol. We characterized the stool bacterial compositions of each sample by sequencing the V4 region of the 16S ribosomal RNA (rRNA) gene on the Illumina MiSeq machine at The Broad Institute of MIT and Harvard, using previously-published methods³³.

Data and statistical analysis

We processed the 16S rRNA amplicon data with Qiime2 and Dada2³⁴. Statistical significance was determined using p-values adjusted for multiple testing with the Benjamini–Hochberg correction (q-values). Analyses were performed in R environment (R version 4.4.1 (2024-06-14)).

To ensure reliable and standardized comparisons of microbial diversity across samples, we performed rarefaction to a depth of 8,000 reads using the phyloseq package (version 1.48.0). Alpha diversity metrics were computed on the rarefied dataset using the vegan package (version 2.7–0). A linear regression model was employed to compare Shannon diversity indices across localities or levels of urbanism while controlling for the covariates sex, age, and BMI. For general comparisons between multiple groups, we applied the Kruskal–Wallis test, followed by pairwise comparisons using the Wilcoxon rank-sum test where applicable. Beta diversity distance metric was calculated using the Bray–Curtis dissimilarity metric, followed by ordination with Principal Coordinates Analysis (PCoA). Significant differences in community composition across groups, while controlling for the covariates sex, age, and BMI, were tested using the PERMANOVA tests (adonis function from the vegan package). We conducted a constrained analysis using the redundancy analysis (RDA) model, specifically using the capscale function from the vegan package in R. We constrained the ordination of beta diversity Bray–Curtis distances based on locality and the covariates mentioned above.

The Firmicutes/Bacteroidetes ratio was calculated as the total abundance of taxa classified under the phylum Firmicutes divided by the total abundance of taxa classified under the phylum Bacteroidota and applying a log10 transformation. The Bacteroides/Prevotella index was derived by computing the ratio of the total relative abundance of Bacteroides spp. to Prevotella spp. and applying a log10 transformation.

For differential abundance analysis of 16S rRNA gene sequencing data, we used the Maaslin2 package (version 1.18.0) based on a multivariate linear model with the following covariates: locality, urbanism (urban vs. rural), sex, age, and BMI. The analysis was performed using total sum scaling for normalization and log transformation for feature scaling. Benjamini–Hochberg (BH) correction applied to adjust for multiple testing. Standardization of features was enabled to ensure comparability across variables.

To visualize the phylogenetic relationships between differentially abundant taxa, we employed the GraPhlAn2 tool (version 1.1.3).

To predict functional pathways based on our 16S rRNA data, we used PICRUSt2 tool (version 2.5.3). The predicted MetaCyc (<https://metacyc.org/>) pathway abundances were subsequently analyzed using DESeq2 package (version 1.44.0) to assess differential pathway abundance. Similar to the taxonomic analysis, DESeq2 incorporated locality, sex, age, and BMI as covariates.

Data availability

The author confirms that the data supporting the findings of this study are available within the article and its Supplementary Material. The data supporting the findings of this study are not freely and publicly available due to controlled-access restrictions. The dataset has been deposited in the Database of Genotypes and Phenotypes (dbGaP, Study Accession: phs002235.v1), a controlled-access database. Applicants must comply with the data usage policies and provide justification for access in line with the dbGaP requirements and with the terms of the participant consenting mechanism.

Received: 21 March 2025; Accepted: 12 June 2025

Published online: 05 July 2025

References

1. Chua, E. G. et al. The influence of modernization and disease on the gastric microbiome of Orang Asli, Myanmar and modern Malaysians. *Microorganisms* <https://doi.org/10.3390/microorganisms7060174> (2019).
2. Das, B. et al. Analysis of the gut microbiome of rural and urban healthy Indians living in sea level and high altitude areas. *Sci. Rep.* **8**(1), 1–15. <https://doi.org/10.1038/s41598-018-28550-3> (2018).
3. De Filippo, C. et al. Diet, environments, and gut microbiota: a preliminary investigation in children living in rural and Urban Burkina Faso and Italy. *Front. Microbiol.* **7**, 8–11. <https://doi.org/10.3389/fmicb.2017.01979> (2017).
4. Kabwe, M. H., Vikram, S., Mulaudzi, K., Jansson, J. K. & Makhalanyane, T. P. The gut microbiota of rural and urban individuals is shaped by geography. *BMC Microbiol.* <https://doi.org/10.1186/s12866-020-01907-3> (2020).
5. Oduaran, O. H. et al. Gut microbiome profiling of a rural and urban South African cohort reveals biomarkers of a population in lifestyle transition. *BMC Microbiol.* <https://doi.org/10.1186/s12866-020-02017-w> (2020).
6. Stražar, M. et al. Gut microbiome-mediated metabolism effects on immunity in rural and urban African populations. *Nat. Commun.* <https://doi.org/10.1038/s41467-021-25213-2> (2021).
7. Yeo, L.-F. et al. The oral, gut microbiota and cardiometabolic health of indigenous Orang Asli communities. *Front. Cell. Inf. Microbiol.* **12**, 36. <https://doi.org/10.3389/fcimb.2022.812345> (2022).
8. Obregon-Tito, A. J. et al. Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* <https://doi.org/10.1038/ncomms7505> (2015).

9. Lokmer, A. et al. Response of the human gut and saliva microbiome to urbanization in Cameroon. *Sci. Rep.* <https://doi.org/10.1038/s41598-020-59849-9> (2020).
10. Harith Wafi Mohd Rosman, M., Lay Yong, C., Umais Azman, M., & Izzuan Mohd Ishar, M. (2020). The health issue in Orang Asli community: outbreak of measles. *Malaysian J. Soc. Sci. Hum. (MJSSH)* 5(2). www.msocialsciences.com
11. Mahmud, M. H., Baharudin, U. M. & Md Isa, Z. Diseases among Orang Asli community in Malaysia: a systematic review. *BMC Public Health* <https://doi.org/10.1186/s12889-022-14449-2> (2022).
12. Nicholas, C. *The Orang Asli and The Contest for Resources - Indigenous Politics, Development and Identity in Peninsular Malaysia*. International Work Group for Indigenous Affairs (2000).
13. Masron, T., Masami, F., & Ismail, N. (2013). Orang Asli in Peninsular Malaysia: population, spatial distribution and socio-economic condition. *J. Ritsumekan Soc. Sci. Hum.* 6. <https://www.researchgate.net/publication/286193594>
14. Deepshika, R. et al. Helminth infection promotes colonization resistance via type 2 immunity. *Science* 352(6285), 608–612 (2016).
15. Lee, S. C. et al. Linking the effects of helminth infection, diet and the gut microbiota with human wholeblood signatures. *PLoS Pathog.* <https://doi.org/10.1371/journal.ppat.1008066> (2019).
16. Lee, S. C. et al. Helminth colonization is associated with increased diversity of the gut microbiota. *PLoS Negl. Trop. Dis.* <https://doi.org/10.1371/journal.pntd.0002880> (2014).
17. Tee, M. Z. et al. Gut microbiome of helminth-infected indigenous Malaysians is context dependent. *Microbiome* <https://doi.org/10.1186/s40168-022-01385-x> (2022).
18. Chong, C. W. et al. Effect of ethnicity and socioeconomic variation to the gut microbiota composition among pre-adolescent in Malaysia. *Sci. Rep.* <https://doi.org/10.1038/srep13338> (2015).
19. Dwiyanto, J. et al. Ethnicity influences the gut microbiota of individuals sharing a geographical location: a cross-sectional study from a middle-income country. *Sci. Rep.* <https://doi.org/10.1038/s41598-021-82311-3> (2021).
20. Strachan, D. P. Hay fever, hygiene, and household size. *BMJ* 299(6710), 1259–1260. <https://doi.org/10.1136/bmj.299.6710.1259> (1989).
21. Alias, A., et al (2018). Traditional knowledge management and usage of medicinal plants as daily medication in healing rituals among the Batek of Kuala Koh, Gua Musang, Kelantan: an exploratory evidence. *Int. J. Eng. Technol.* 7 (2). www.sciencepubco.com/index.php/IJET
22. Koh, K., Musang, G., Amran Alias, M., & Salleh, H. The challenges of managing traditional knowledge related to medicinal plants among the Batek Community in. *J. Soc. Dev. Sci.* 5(4) (2014).
23. Tett, A. et al. The *Prevotella copri* complex comprises four distinct clades underrepresented in westernized populations. *Cell Host Microbe* 26(5), 666–679.e7. <https://doi.org/10.1016/j.chom.2019.08.018> (2019).
24. Qiao, S. et al. Gut Parabacteroides merdae protects against cardiovascular damage by increasing commensal bacteria-driven branched-chain amino acid catabolism. <https://doi.org/10.21203/rs.3.rs-1127540/v1> (2021).
25. Da Silva Morais, E., Grimaud, G. M., Warda, A., Stanton, C. & Ross, P. Genome plasticity shapes the ecology and evolution of *Phocaeicola dorei* and *Phocaeicola vulgatus*. *Sci. Rep.* <https://doi.org/10.1038/s41598-024-59148-7> (2024).
26. Cabral, L. S. & Weimer, P. J. *Megasphaera elsdenii*: its role in ruminant nutrition and its potential industrial application for organic acid biosynthesis. *Microorganisms* <https://doi.org/10.3390/microorganisms12010219> (2024).
27. Rosas-Plaza, S. et al. Human gut microbiome across different lifestyles: from hunter-gatherers to urban populations. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2022.843170> (2022).
28. Schnorr, S. L. et al. Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* <https://doi.org/10.1038/ncomms4654> (2014).
29. Syromyatnikov, M. et al. Characteristics of the gut bacterial composition in people of different nationalities and religions. *Microorganisms* <https://doi.org/10.3390/microorganisms10091866> (2022).
30. Rodionova, I. A. et al. Comparative genomics and functional analysis of rhamnose catabolic pathways and regulons in bacteria. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2013.00407> (2013).
31. Yan, T. et al. Habitual intakes of sugar-sweetened beverages associated with gut microbiota-related metabolites and metabolic health outcomes in young Chinese adults. *Nutr. Metab. Cardiovasc. Dis.* 33(2), 359–368. <https://doi.org/10.1016/j.numecd.2022.10.016> (2023).
32. Tacey, I. et al. *Proceedings of the 2nd Temenggor Scientific Expedition*, 73–87 (2012).
33. Poyet, M. et al. A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. *Nat. Med.* 25(9), 1442–1452. <https://doi.org/10.1038/s41591-019-0559-3> (2019).
34. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37(8), 852–857. <https://doi.org/10.1038/s41587-019-0209-9> (2019).

Acknowledgements

This research was supported by grants from the Center for Microbiome Informatics and Therapeutics at MIT under the Global Microbiome Conservancy Project (IF030-2018) and the Universiti Malaya Partnership Grant (RK019-2018). Further support was received from the German Science Foundation through the Collaborative Research Center (CRC) 1182 on the Origin and Function of Metaorganisms (Project-ID 261376515 – SFB 1182), with project C5.1 awarded to MG and project C5.2 to MP. This work was also funded by the German Research Foundation (DFG) under the NFDI4Microbiota project (Flex Fund grant, DFG project no. 28/1 AOB): 700895 Bio4ALL, NFDI4Microbiota) as part of the German National Research Data Infrastructure (NFDI).

Author contributions

M.P., and M.G., designed this study. M.P., M.G., and E.J.A. founded the Global Microbiome Conservancy Project under which field collections occurred. F.I., Y.A.L.L., T.M.P., L.S.C., C.D., C.H.C. E.J.A., M.G. and M.P. managed field administrative work and performed the collection of data and samples. M.P. performed DNA extraction from stool samples and library preparation for 16S sequencing. M.G., M.P., A.R.A.M., N.N.R., B.O. and N.F. performed computational work and data analyses. N.F., O.B., N.N.R., M.P. and M.G. wrote the manuscript, with the supervision of F.I. All authors reviewed the final version of the manuscript. C.H.C. provided the pictures to document the Global Microbiome Conservancy’s work in Malaysia, which is featured in this article.

Funding

This research was supported by grants from the Center for Microbiome Informatics and Therapeutics at MIT under the Global Microbiome Conservancy Project (IF030-2018) and the Universiti Malaya Partnership Grant (RK019-2018). Additional support was provided by the German Research Foundation (DFG) through the Collaborative Research Center (CRC) 1182 “Origin and Function of Metaorganisms” (Project-ID 261376515 – SFB 1182), with project C5.1 awarded to MG and project C5.2 to MP. This work was also funded by the DFG under

the NFDI4Microbiota project (Flex Fund grant, DFG Project No. 28/1, AOBJ: 700895, Bio4ALL, NFDI4Microbiota) as part of the German National Research Data Infrastructure (NFDI).

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval and consent to participate

Ethical approvals were obtained from the Massachusetts Institute of Technology Institutional Review Board (MIT IRB; Protocol No. 1612797956) and the Universiti Malaya Medical Research Ethics Committee (MREC; ID No. 2018219–6033). All procedures involving human participants were conducted in accordance with the ethical standards of the respective institutional and national research committees, as well as the Declaration of Helsinki and its later amendments or comparable ethical guidelines. Informed consent was obtained from all individual participants included in the study.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-07117-z>.

Correspondence and requests for materials should be addressed to M.P. or F.I.

Reprints and permissions information is available at www.nature.com/reprints.

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

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

© The Author(s) 2025