Identification of age-associated microbial changes via long-read 16S sequencing

Kai Yee Toh[1*], Tzi Shin Toh[2], Khi Pin Chua[3], Priscilla Rajakumar[1], Jonathan Wei Jie Lee[1,4,7,8,9] and Chun Wie Chong[5,6]

Abstract

Background Age-related gut microbial changes have been widely investigated over the past decade. Most of the previous age-related microbiome studies were conducted on the Western population, and the short-read sequencing (e.g., 16S V4 or V3-V4 region) was the most common microbiota profiling method. We evaluated the gut compositional differences using the long-read sequencing approach (i.e., PacBio sequencing targeting the full-length V1-V9 regions) to enable a deeper taxonomic resolution and better characterize the gut microbiome of Singaporeans from different age groups.

Results A total of 83 research participants were included in this study. Although no significant differences were detected in alpha and beta diversity, our study demonstrated several bacterial taxa with abundances that were significantly different across age groups. With young individuals as the reference group, *Eggerthella lenta* and *Bacteroides uniformis* were found to be significantly altered in the middle-aged group, while *Catenibacterium mitsuokai* and *Bacteroides plebeius* were significantly altered in the elderly group. These age-related differences in the gut microbiome were associated with aberrations in several predicted functional pathways, including dysregulations of pathways related to lipopolysaccharide and tricarboxylic acid cycle in older adults.

Conclusions The utilization of long-read sequencing facilitated the identification of species- and strain-level differences across age groups, which was challenging with the partial 16S rRNA sequencing approach. Nevertheless, replication studies are warranted to confirm our findings, and if confirmed, further in vitro and in vivo studies are crucial to better understand the impact of the altered levels of age-related bacterial taxa. Additionally, the modest performance of strain-level taxonomic classification using 16S-ITS-23S gene sequences, likely due to the limited depth of currently available alignment databases, highlights the need for optimization and refinement in curating these databases for the long-read sequencing approach.

Keywords Gut microbiome, Aging, Long-read sequencing

*Correspondence:
Kai Yee Toh
kaiyee.toh@amili.asia
[1]AMILI Pte Ltd, 89 Science Park Drive #03-09, The Rutherford, Lobby C, Singapore Science Park 1, Singapore 118261, Singapore
[2]Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[3]Pacific Biosciences of California, Menlo Park, CA, USA
[4]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore
[5]School of Pharmacy, Monash University Malaysia, Selangor, Malaysia
[6]MUM Microbiome Research Centre, Monash University Malaysia, Selangor, Malaysia
[7]Division of Gastroenterology and Hepatology, Department of Medicine , National University Health System, Singapore 119228, Singapore
[8] iHealthtech, National University of Singapore, Singapore 117599, Singapore
[9] SynCTI, National University of Singapore, Singapore 117456, Singapore

Background

According to the 2023 United Nations Report, the global aging population over the age of 65 is projected to reach 1.6 billion by 2050 [1]. While this significant increase in life expectancy is a remarkable achievement, it is accompanied by a rise in disability rates, potentially leading to higher healthcare costs. Understanding how environmental and biological factors contribute to unhealthy aging is therefore essential. Generally, the composition of the human gut microbiota remains relatively stable during early and mid-adulthood. However, as individuals age, the stability of the gut microbiome may be disrupted due to various factors such as biological aging processes, changes in health status, and diet [2, 3].

A growing body of evidence suggests alterations in gut microbial composition in aged individuals. However, current knowledge on aging and the gut microbiome is predominantly derived from studies conducted in the Western world [4, 5] and, to some extent, China [6, 7]. There is a lack of similar studies from ASEAN countries, such as Singapore, which has an ethnically diverse population (i.e., Chinese, Malay, and Indian) practicing contrasting lifestyles. Furthermore, most previous studies have utilized the 16S ribosomal ribonucleic acid (rRNA) sequencing method (e.g., V3-V4 region [5, 8, 9]), which has several limitations.

Short-read sequencing of the bacterial 16S rRNA gene is often used to profile the gut microbiome due to its cost-effectiveness. The 16S rRNA gene, which is 1.5 kb in length, covers nine hypervariable regions. Amplicon-based 16S rRNA gene sequencing allows the amplification of selected variable regions only. The limited read length of the short-read sequencing approach, combined with amplification bias, often results in low taxonomic resolution [10, 11]. To overcome these limitations, some studies have utilized shotgun metagenomics sequencing, which provides better sensitivity and resolution by fragmenting the whole genome into smaller pieces of deoxyribonucleic acid (DNA) and then reassembling the fragments using overlapping regions. However, shotgun metagenomics sequencing often requires significant computational power and is costly.

Long-read sequencing is a cost-effective alternative to shotgun metagenomics and offers better taxonomic resolution than short-read 16S rRNA sequencing. This method allows the entire 16S rRNA gene (i.e., all nine variable regions), extending to the Internal Transcribed Spacer (ITS) and 23S rRNA gene regions, to be sequenced. Although some may argue that long-read sequencing has lower accuracy ($\sim 90\%$), the recent development of circular consensus sequencing (CCS) from PacBio and improved methodologies to remove polymerase chain reaction (PCR)-related errors have enabled researchers to discriminate sequences differing by a single nucleotide across an entire gene with an average length of up to 13.5 kb with 99.8% accuracy [11, 12]. This allows for the identification of intragenomic 16S gene copy variants within a taxon, enabling easier distinction of species and strains (sub-species) and preventing the over-estimation of bacterial diversity [10, 13].

Therefore, in this study, we utilized the long-read sequencing approach to investigate the gut microbiome profiles of individuals from different age groups, with the aim to better characterize and understand the gut microbial differences at species-level.

Methods

Data and sample collection

Singapore residents above age of 21 were recruited from AMILI's Poop Saves Life campaign which was a public outreach program, where 83 participants were selected for this study. Participants were selected if they fulfilled the following criteria: (1) ethnicity of Chinese, Malay, or Indian, (2) no use of antibiotics three months prior to recruitment, (3) no regular use of proton pump inhibitors, and (4) absence of gastrointestinal diseases such as inflammatory bowel disease and colon disease. Written informed consent was obtained from all participants and the study was approved by AMILI Institutional Review Board (reference number 2020/0501) in accordance with ethical guidelines.

Basic socio-demographic information and medical history, including age and body mass index (BMI), were collected from each participant. Information on daily dietary intake were recorded using a locally-validated Food Frequency Questionnaire (FFQ) [14]. Stool samples were collected in DNA/RNA shield stool collection kit (Zymo Research, Irvine, California, USA). The stool collection kit and a detailed instruction sheet on stool collection (Additional File 1) were provided to the participants. After sample collection, stool samples were sent to the laboratory via post. All samples reached the laboratory within a week from sample collection. Upon receiving, the fecal samples were aliquoted and frozen at -80 °C until further sample processing.

Sample preparation

Total DNA was extracted from the fecal samples using the QIAmp PowerFecal Pro DNA extraction kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The full-length 16S-ITS-23S rRNA gene was amplified by PCR using consensus sequences 27F 5'-AGRRTTYGATYHTDGYTYAG-3' and 23SR 5'-AGTACYRHRARGGAANGR-3'. Amplicon libraries were created using the Shoreline Complete StrainID kit (Shoreline Biome, Farmington, Connecticut, US) and sequenced on PacBio Sequel II system.

Sequence and predictive functional analyses

The raw 16S-ITS-23S rRNA gene sequences were first quality-filtered and denoised using DADA2 version 1.22.0 [15] and subsequently imported into SBanalyzer software for strain-level microbial identification using the Athena database. For the analysis of 16S V1-V9 gene sequences, the raw 16S-ITS-23S rRNA gene sequences were first trimmed to the 16S V1-V9 region using cut-adapt version 3.7 [16] based on primer sequences F27 (AGRGTTYGATYMTGGCTCAG) and R1492 (AAGTC-GTAACAAGGTARCY). The trimmed sequences were then quality-filtered and denoised using DADA2 version 1.22.0 [15]. After that, filtered reads were grouped into amplicon sequence variants (ASVs). To maximize the assignment rate, each ASV was first taxonomically classified based on GTDB r207 [17], followed by Silva v138 [18], and lastly RefSeq+RDP [19, 20]. The microbial gene content was then inferred from the taxa abundance using PICRUSt2 [21].

ASV sequences corresponding to the differentially abundant bacterial species detected using the 16S V1-V9 dataset were extracted and searched against the ASV taxonomy table of the 16S-ITS-23S dataset. The 16S-ITS-23S ASVs that have similar gene sequences and that were assigned to the same species were analyzed to obtain strain-level information.

Statistical analysis

All statistical analyses were carried out using R version 4.1.2. All tests were considered significant at p-value < 0.05. All p-value adjustments were based on the Benjamini-Hochberg procedure. Alpha diversity measurements (Shannon, Chao1, and Pielou's index) were computed using the microbiome R package version 1.16.0 [22]. Beta diversity measures were inferred by first transforming the data based on centered log-ratio to account for compositionality [22, 23]. The Aichitson distance was used for principal coordinate analysis (PCoA) and permutational multivariate analysis of variance (PERMANOVA) (1000 permutations). PERMANOVA was carried out using the adonis2 function in vegan version 2.6.2 [24]. Differentially abundant taxa/pathways across different age groups (adjusted p-value < 0.05) were identified based on the analysis of compositions of microbiomes with bias correction 2 (ANCOM-BC2) method [25].

Results

Subjects' characteristics

A total of 83 participants were included in this study. Following the age cut-offs used by several prior studies [9, 26, 27], participants were grouped into different age groups, as follows: (1) young (age < 36; n = 25), (2) middle-aged (36 ≤ age < 56; n = 30), and (3) old (age ≥ 56; n = 28).

The demographics and basic characteristics of the cohort are summarized in Table 1. There were no significant differences in sex, BMI, race, smoking status and daily dietary intake among the three groups included in this study.

Comparisons of taxonomy classification of 16S-ITS-23S vs. 16S V1-V9 datasets

The original 16S-ITS-23S sequencing data yielded a total of 2,327,728 gene sequences (mean = 28,045; standard deviation = 8,494; min = 11,147; max = 49,313). After the alignment against Athena database, nine phyla, 14 classes, 26 orders, 45 families, 83 genera, and 144 species were detected (Additional File 2: Supplementary Table 1). Of the 2,740 features, 1,843 (67.3%) and 1,111 (40.5%) reached species- and strain-level resolution, respectively. Notably, there were only 59 different strains found in the output of this sequencing dataset. Given that bacterial strains can differ by only a few bases, the limited strain-level information obtained using the 16S-ITS-23S sequences is likely due to the limited depth of the Athena database.

To maximize the utility of our data, we trimmed the 16S-ITS-23S sequences to the 16S V1-V9 region, aiming to identify more distinct species. This 16S V1-V9 sequencing dataset yielded 1,937,509 gene sequences, with an average of 23,343 sequences per sample (ranging from 9,127 to 44,382, with a standard deviation of 8,783). After quality filtering using the DADA2 R-package and assigning taxonomy based on three different reference databases, we identified eight phyla, 19 classes, 40 orders, 76 families, 182 genera, and 299 species. Of the 2,307 ASVs, 2,031 (88.0%) reached species-level resolution (Additional File 2: Supplementary Table 1). Since the taxonomic classification of 16S V1-V9 data provided more information, we proceeded the subsequent analyses using this dataset, unless otherwise specified.

Comparisons of bacterial richness (alpha diversity) and composition (beta diversity) across age groups

No significant differences were detected in all diversity indices across the three age groups (Wilcoxon rank-sum test, adjusted p-value > 0.05; Fig. 1A). However, participants of older age generally exhibited higher alpha diversity than younger participants.

PERMANOVA revealed that age groups explained about 2.8% of the overall variation in gut microbiome composition, although this was not statistically significant ($R^2 = 0.028$, p-value = 0.098). Additionally, the complementary PCoA analysis showed modest separation among age groups (Fig. 1B). Age (on a continuous scale) ($R^2 = 0.015$, p-value = 0.089), as well as the rest of the potential confounding factors, including sex ($R^2 = 0.012$, p-value = 0.399), BMI ($R^2 = 0.016$, p-value = 0.155), race

Toh *et al. Gut Pathogens*   (2024) 16:56

Page 4 of 10

Table 1 Demographics and characteristics of subjects included in the study

	Young (< 36 years)	Middle-aged (36–55 years)	Old (≥ 56 years)	*p*-value
Sample size (N)	25	30	28	-
Age (years)	28.0 [6.0]	43.5 [8.8]	66.0 [10.3]	**0.001***
Sex				
Male	12 (48.0)	15 (50.0)	13 (46.4)	0.963
Female	13 (52.0)	15 (50.0)	15 (53.6)	
BMI[a]	24.1 [4.2]	22.7 [5.3]	22.9 [5.7]	0.496
BMI category[a, b]				
Underweight	1 (4.0)	0 (0.0)	4 (14.3)	0.169
Normal	6 (24.0)	14 (46.7)	10 (35.7)	
Overweight	11 (44.0)	7 (23.3)	9 (32.1)	
Obese	3 (12.0)	6 (20.0)	4 (14.3)	
Race				
Chinese	24 (84.0)	27 (90.0)	27 (96.4)	0.252
Malay	1 (4.0)	0 (0.0)	0 (0.0)	
Indian	0 (0.0)	3 (10.0)	1 (3.6)	
Smoking status				
Non-smoker	24 (84.0)	26 (86.7)	28 (100.0)	0.276
Past smoker	1 (4.0)	3 (10.0)	0 (0.0)	
Current smoker	0 (0.0)	1 (3.3)	0 (0.0)	
Dietary intake[c]				
Total energy (kcal/day)	2417.6 [977.1]	2052.9 [773.4]	1986.4 [1221.5]	0.396
Total protein (g/day)	91.5 [42.6]	83.8 [43.1]	83.3 [73.4]	0.550
Total fat (g/day)	101.3 [23.8]	87.6 [26.8]	77.5 [58.5]	0.230
Saturated fatty acids (g/day)	38.3 [14.4]	31.7 [11.6]	25.5 [17.8]	0.099
Monounsaturated fatty acids (g/day)	38.8 [11.8]	34.1 [15.7]	29.2 [26.8]	0.219
Polyunsaturated fatty acids (g/day)	17.0 [6.5]	17.0 [6.6]	17.3 [11.7]	0.919
Total carbohydrate (g/day)	278.6 [192.6]	245.0 [112.1]	221.2 [189.0]	0.449
Starch (g/day)	172.8 [158.6]	145.8 [103.0]	148.8 [159.2]	0.191
Sugar (g/day)	62.6 [43.7]	67.6 [41.8]	73.9 [34.8]	0.647
Total fiber (g/day)	18.1 [7.8]	18.0 [7.9]	18.0 [11.3]	0.922
Vitamin A (mg/day)	871.8 [489.1]	758.5 [249.3]	771.1 [523.4]	0.863
Vitamin C (mg/day)	82.5 [48.1]	82.6 [49.1]	90.9 [63.8]	0.578
Calcium (mg/day)	597.7 [410.8]	619.9 [293.9]	648.9 [419.7]	0.786
Iron (mg/day)	16.8 [6.3]	15.7 [8.2]	15.5 [7.8]	0.735
β-carotene (mg/day)	2920.6 [2586.1]	2396.3 [1710.9]	2709.5 [2312.2]	0.299
Thiamin (mg/day)	1.4 [1.1]	1.5 [0.7]	1.4 [0.9]	0.899
Riboflavin (mg/day)	1.7 [0.7]	1.6 [0.8]	1.6 [1.0]	0.971
Potassium (mg/day)	2436.0 [1012.6]	2183.9 [749.2]	2398.8 [1510.2]	0.296
Zinc (mg/day)	11.1 [4.9]	10.2 [5.7]	10.1 [6.7]	0.371

Non-normally distributed continuous data are presented as median [interquartile range], while categorical data are reported as *n* (%). *p*-values are obtained using Kruskal-Wallis test (for non-normally distributed data) or chi-square test (for categorical data). [a]Information available for 75/83 participants. [b]Based on the Asian BMI cut-off (i.e., underweight: BMI < 18.5 kg/m^2; normal: 18.5 kg/m^2 ≤ BMI < 23.0 kg/m^2; overweight: 23.0 kg/m^2 ≤ BMI < 27.5 kg/m^2; obese: BMI ≥ 27.5 kg/m^2). [c]Information available for 79/83 participants

($R^2 = 0.025$, *p*-value = 0.435), smoking status ($R^2 = 0.029$, *p*-value = 0.117), and various dietary parameters, did not significantly contribute to differences in gut microbial composition (Table 2).

Differentially abundant taxa across age groups

Using young participants as the reference group, we identified twelve and ten bacterial species that were differentially abundant in the middle-aged and old groups, respectively (adjusted *p*-value < 0.05, Fig. 2A). However, only four differentially abundant taxa (i.e., *Eggerthella lenta* [adjusted *p*-value = 0.049], *Bacteroides uniformis* [adjusted *p*-value = 0.033], *Catenibacterium mitsuokai* [adjusted *p*-value = 0.004] and *Bacteroides plebeius*

[adjusted *p*-value = 0.030]) passed sensitivity tests for the pseudo-count addition. This step helps to mitigate the risk of inflated false-positive rates due to the implementation of pseudo-count in ANCOM-BC2 [25]. *E. lenta* was significantly increased in the middle-aged group, while *B. uniformis* was significantly reduced in the same group. On the other hand, the abundances of *B. plebeius* and *C. mitsuokai* were significantly increased and reduced, respectively, in the old group. Interestingly, when the ASV gene sequences of the four differentially abundant bacterial species were compared against the 16S-ITS-23S output, we identified *E. lenta DSM 2243* to be exclusively present in the middle-aged subgroup. However, it is important to note that the bacterial strain was detected
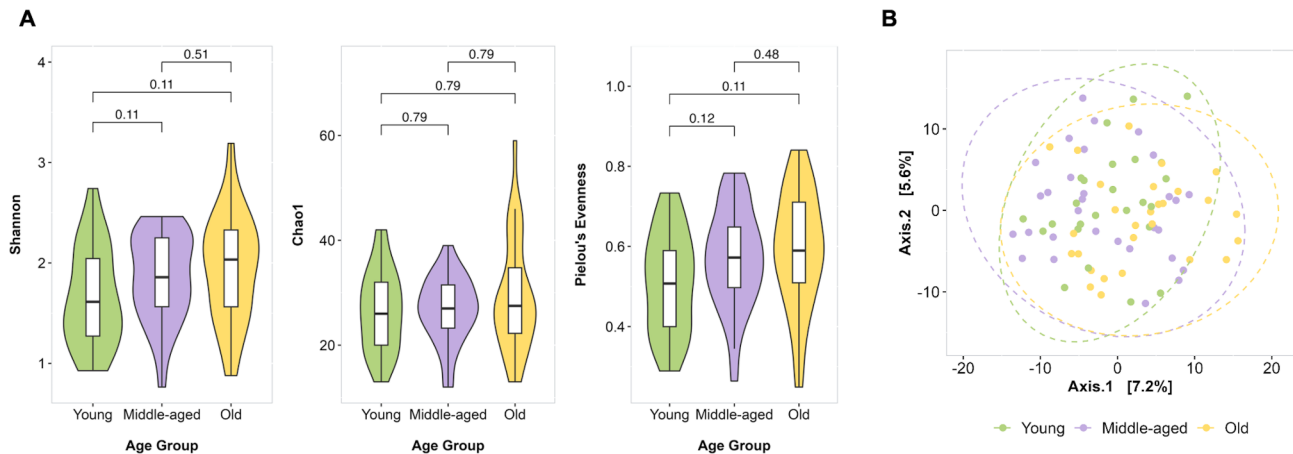
Fig. 1 Comparisons of alpha diversity and beta diversity across age groups. (**A**) Alpha diversity is measured using the Shannon diversity index, Chao1 diversity index, and Pielou's evenness index. No significant differences were detected in any of the diversity indices across the three age groups (Wilcoxon rank-sum test, adjusted $p > 0.05$). (**B**) PCoA demonstrating insignificant separation in the gut microbial composition of individuals from different age groups (PERMANOVA; $R^2 = 0.028$, $p = 0.098$). Each data point represents a single sample. The distance between data points reflects the differences between samples

Table 2 Effects of age groups and various potential confounders on gut microbiome compositional differences

Factor	R^2 value	p-value
Age groups	0.028	0.098
Age	0.015	0.089
Sex	0.012	0.399
BMI[a]	0.016	0.155
Race	0.025	0.435
Smoking status	0.029	0.117
Dietary intake[b]		
Total energy	0.015	0.211
Total protein	0.015	0.218
Total fat	0.014	0.222
Saturated fatty acids	0.016	0.131
Monounsaturated fatty acids	0.014	0.279
Polyunsaturated fatty acids	0.011	0.718
Total carbohydrate	0.015	0.214
Starch	0.014	0.306
Sugar	0.017	0.079
Total fiber	0.014	0.299
Vitamin A	0.013	0.436
Vitamin C	0.013	0.510
Calcium	0.015	0.251
Iron	0.014	0.301
β-carotene	0.011	0.727
Thiamin	0.014	0.290
Riboflavin	0.015	0.210
Potassium	0.014	0.270
Zinc	0.013	0.464

All datasets were first transformed using centered log-ratio transformation to account for compositionality. The Euclidean distance was then used for permutational multivariate analysis of variance (PERMANOVA). The R^2 value indicates the effect size, while p-value < 0.05 is considered significant. [a]Calculated based on 75 samples. [b]Calculated based on 79 samples

in only 8/30 (26.7%) of them (Additional File 2: Supplementary Table 2).

At the genus level, *Olsenella* was found to be significantly increased in middle-aged participants (adjusted p-value < 0.001), while *Holdemanella* (adjusted p-value < 0.001) was significantly increased and *Catenibacterium* (adjusted p-value = 0.004) was significantly reduced in the old subgroup. Additionally, bacteria from the *Atopobiaceae* family were found to be increased in the gut microbiome of middle-aged individuals (adjusted p-value < 0.001). All four differentially abundant bacterial taxa had an adjusted p-value < 0.05 and passed the sensitivity tests. The complete list of bacterial classes, orders, families, and genera detected to be differentially abundant in the middle-aged and old groups, with young individuals as the reference, is provided in Additional File 2: Supplementary Fig. 1.

Differentially abundant bacterial taxa were also identified in pairwise comparisons between all possible combinations of the age groups. Bacterial species with abundances that were significantly different between any two groups (adjusted p-value < 0.05), as well as passed the sensitivity tests, are listed in Supplementary Table 3 (Additional File 2), along with their respective log fold changes and adjusted p-values.

Predictive functional analysis

With young individuals as the reference, predictive functional analysis revealed a total of eight metabolic pathways that differed significantly across age groups (adjusted p-value < 0.05; Fig. 2B). Among these, three pathways were significantly downregulated in the middle-aged subgroup, including (1) superpathway of UDP-*N*-acetylglucosamine-derived O-antigen building blocks
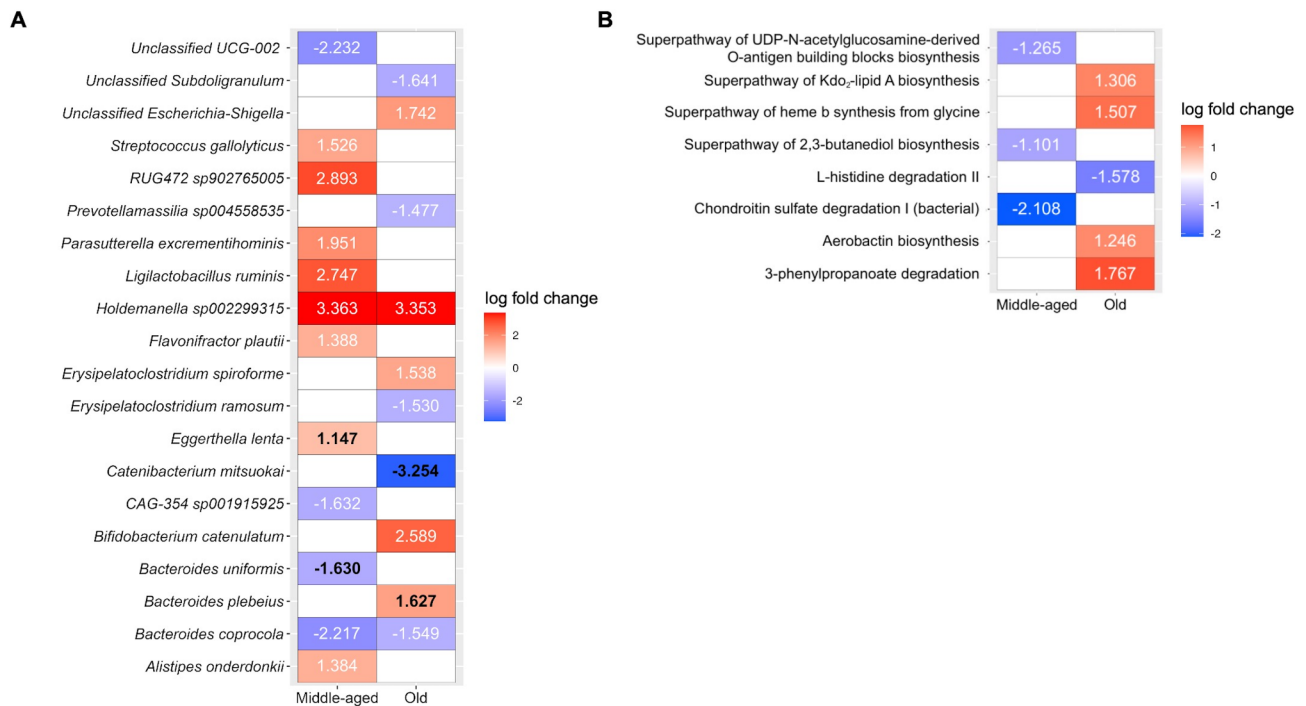
A

B

Fig. 2 Differentially abundant bacterial species and predicted functional pathways. The analysis was conducted using the ANCOM-BC2, with young individuals as the reference group. The columns represent age groups, while the rows show differentially abundant (**A**) bacterial taxa or (**B**) predicted functional pathways. Cell colors indicate log fold changes, with blue representing a significant reduction and red representing a significant increase (darker colors correspond to greater log fold changes). The log fold changes relative to the reference (i.e., young) group are noted in each cell. Values in white font indicate log fold changes that are significant at an adjusted p-value < 0.05, while values in bold black font indicate log fold changes that are significant at an adjusted p-value < 0.05 and passed sensitivity tests for pseudo-count addition

biosynthesis, (2) superpathway of 2,3-butanediol biosynthesis and (3) chondroitin sulfate degradation I. The remaining five functional pathways were differentially abundant in the elderly group, with the superpathway of $(Kdo)_2$-lipid A biosynthesis, superpathway of heme b synthesis from glycine, aerobactin biosynthesis and 3-phenylpropanoate degradation being significantly increased, while L-histidine degradation II being significantly downregulated. However, it is important to note that these pathways did not pass the ANCOM-BC2 sensitivity tests.

Discussion

In this study, we investigated the association between age and gut microbial composition in individuals residing in Singapore. To the best of our knowledge, this is the first full-length 16S rRNA gene assessment of the gut microbiota in a multi-ethnic country. Previous studies evaluating the effect of age on the gut microbiome primarily used a short-read 16S rRNA gene sequencing approach (e.g., V3-V4 region: ~460 base pairs [5, 8, 9, 28]). The limited resolution of short-read sequencing often fails to detect bacterial changes at the species/strain level, which can affect data interpretation. In our work, we utilized the long-read sequencing approach (i.e., PacBio sequencing) to explore age-related gut microbiome alterations for

the first time. In addition to replicating previous findings, our study unveiled several novel differentially abundant taxa and predicted functional pathways associated with age.

Overall, no significant differences were detected in alpha diversity across all age groups, although there was an increasing trend with age. de la Cuesta-Zuluaga et al. (2019) reported a minimal association between gut biodiversity and age in a Chinese cohort, contrasting with the significant positive correlation found in cohorts from the United States, United Kingdom, and Colombia [29]. This finding supports our results, as nearly all participants in our study are Chinese (n=78; 94.0%). However, it is well established that individuals of the same ethnicity may exhibit different gut microbial profiles when residing in different geographical regions [30]. Therefore, factors such as ethnicity, geographical location, and dietary habits must be controlled when investigating the gut microbiome.

To ensure a more robust interpretation of age-related gut microbiome changes, we investigated the effects of various potential confounders, including sex, BMI, ethnicity, smoking status, and daily dietary intake, on the overall gut microbiome composition of our cohort. None of these potential confounders significantly contributed

to the overall gut microbiome differences in the cohort. We also did not observe significant effects of age groups or age (on a continuous scale) on the overall gut microbial composition of the participants, which was further supported by the modest separation observed in the PCoA analysis. While these insignificant findings could be attributed to the study's limited sample size, the lack of significant differences in gut microbial composition across age groups is consistent with results from two independent Japanese studies [9, 28]. Although the exact reasons for discrepancies with other age-related gut microbiome studies are unclear, this could be partly due to their Asian origin or the health of participants (i.e., participants were relatively healthy). Further investigations are warranted to better understand these observations.

Despite the insignificant differences in alpha and beta diversity, several differentially abundant bacterial taxa were detected among the age groups. For instance, a notable decrease in *Bacteroides uniformis* was observed in the gut microbiome of middle-aged individuals, while *Bacteroides plebeius* was significantly elevated in the old group. These findings align with multiple earlier studies, which reported contrasting findings on the abundances of *Bacteroides* in the gut microbiome of individuals from different age groups. Some studies documented increased levels in younger adults [4, 8], while others reported higher abundances in the elderly group [28, 31]. Interestingly, recent investigations have also linked the differential abundance of *Bacteroides* to the overall health status of the study group [5, 32]. Collectively, these findings underscore the importance of employing sequencing techniques that provide precise taxonomic assignments down to the species/strain level, thereby facilitating a more comprehensive delineation of the gut microbiome and permitting a more accurate understanding of the microbial ecology associated with specific conditions.

In the pairwise comparison between middle-aged and old groups, we found that elderly individuals exhibited a significantly higher abundance of *Klebsiella pneumoniae* (as well as the *Klebsiella* genus) in their gut microbiome. The elevated level of *K. pneumoniae* in older individuals is believed to be associated with factors such as increased use of medication [33] or inflammation linked to interleukin-6 [34], both of which are common in older individuals. This bacterium, which is known to be a pathogen, may contribute to health issues frequently observed in this age group. Given the potential health implications, future investigation should aim to validate this finding through larger, longitudinal studies. If confirmed, further research should explore the specific implications of changes in *K. pneumoniae* level in older adults.

Besides replicating previous research findings, our study revealed several novel differentially abundant

bacterial species that have not been previously reported. These include increased abundances of *Eggerthella lenta* in middle-aged participants and reduced levels of *Catenibacterium mitsuokai* in elderly individuals. *E. lenta*, a bacterium belonging to the *Coriobacteriaceae* family, is known as an opportunistic pathogen implicated in various conditions and infections [35]. This gut bacterium is also involved in the inactivation of the cardiac drug digoxin [36]. While *E. lenta* may negatively impact human health, it is crucial to obtain strain-level information to elucidate its exact role in the aging-related gut microbiome. *C. mitsuokai*, on the other hand, is generally considered part of the normal human gut microbiome [37]. Previous studies have linked *C. mitsuokai* with dyslipidemia and insulin resistance [38], and a higher abundance of the *Catenibacterium* genus has been associated with a potentially lower risk of frailty [39]. Altogether, these findings suggest potential health implications related to changes in the levels of *C. mitsuokai* in the gut microbiome. The observed reduction of *C. mitsuokai* in elderly individuals of our cohort could either reflect age-related alterations in gut microbiome composition or represent a compensatory response to the health changes commonly seen in old individuals. With the increasing emphasis on independent replication of scientific findings, especially in the field of gut microbiome where numerous confounders exist [40], further investigations are warranted to confirm these findings and explore their roles in relation to human age.

Given the importance of understanding the functional implications of changes in the gut microbiome across different age groups, we conducted a predictive functional analysis using PICRUSt2. Through differential abundance analysis, we identified several metabolic pathways that may be impacted by alterations in the gut microbiome. Notably, we observed dysregulated pathways related to lipopolysaccharide (LPS) biosynthesis, including the superpathway of UDP-*N*-acetylglucosamine-derived O-antigen building blocks biosynthesis and the superpathway of $(Kdo)_2$-lipid A biosynthesis, in older individuals. LPS are outer membrane components of gram-negative bacteria, composed of three main domains: lipid A, the core oligosaccharide, and the O antigen [41]. The dysregulations of LPS-related pathways observed in this study suggest the presence of gut inflammation in older individuals, potentially activated by LPS through toll-like receptor (TLR)-4 signaling [42]. Interestingly, we also observed a significant reduction in chondroitin sulfate degradation in middle-aged individuals. The degradation and fermentation of chondroitin sulfate by the human gut microbiota produce significant amounts of short-chain fatty acids, which are also known to regulate gut permeability and inflammation [43]. Additionally, elderly individuals demonstrated downregulated

degradation of L-histidine and increased degradation of 3-phenylpropanoate. Aberrations of these two metabolic pathways could lead to the production of compounds like fumarate, pyruvate and acetoacetate, which are involved in the tricarboxylic acid (TCA) cycle, potentially disrupting central metabolism [44, 45]. However, the functional analysis conducted in this study is predicted based on microbial abundances, and the differentially abundant predicted functional pathways did not pass the ANCOM-BC2 sensitivity tests, and thus requires cautious interpretation. Further in vitro and in vivo studies are warranted to better elucidate the functional implications of age-related microbial changes in the gut.

Our study is not free from limitations. Although we managed to identify several age-related differentially abundant bacterial species, most of our results were derived using the 16S V1-V9 sequences. This approach was necessitated by the modest taxonomic classification of 16S-ITS-23S sequences down to species/strain level, which is likely attributed to the limited depth of currently available alignment databases. For instance, using Athena database (16S-ITS-23S), only 67.3% and 40.5% of the features reached species- and strain-level resolution, respectively.

Furthermore, the small sample size of our study may impact the robustness of detecting microbiome differences across age groups, thereby requiring cautious interpretation and further validation in future studies. The use of the FFQ provides only a snapshot of the participants' nutritional intake and may not fully capture the complexity of diet-microbiota interactions in this study. Additionally, the age cut-offs used in this study were primarily based on the quartiles and median age of the overall cohort, as well as the cutoffs employed by a few previous studies. The lack of consistent definitions of age cut-offs for young, middle-aged, and old adults in gut microbiome research was similarly highlighted by a previous study [46]. Lastly, the effect of age (as a continuous variable) on the gut microbiome was not explored in this study due to the lack of significance in beta diversity and differential abundance analyses.

Conclusions

Although no significant differences were observed in alpha and beta diversity, our study revealed several age-related differentially abundant bacterial taxa using the long-read sequencing method. Specifically, we observed altered abundances of *Eggerthella lenta* (enriched) and *Bacteroides uniformis* (reduced) in middle-aged individuals, while the old individuals exhibited significant alterations in the levels of *Bacteroides plebeius* (enriched) and *Catenibacterium mitsuokai* (reduced). These changes in the gut microbiome across age groups had some functional implications, such as dysregulations of LPS-related

pathways and metabolic pathways involved in the TCA cycle. Further studies are warranted to confirm our findings and to better elucidate the impact of age-related changes in the gut microbiome. Importantly, the higher taxonomic resolution provided by the long-read sequencing technology used in this study allowed for more specific identification of bacterial taxa that may have been overlooked in previous studies employing the more commonly used short-read sequencing method. However, the limited depth of currently available alignment databases for 16S-ITS-23S sequences hindered the detection of a broader range of bacterial species/strains. Further optimization and refinement in the curation of alignment databases for 16S-ITS-23S sequences are required to resolve the issues faced in this study, as well as to enhance the accuracy and comprehensiveness of microbial identification in future research.

Abbreviations

rRNA	Ribosomal ribonucleic acid
DNA	Deoxyribonucleic acid
ITS	Internal Transcribed Spacer
CCS	Circular consensus sequencing
PCR	Polymerase chain reaction
BMI	Body mass index
FFQ	Food Frequency Questionnaire
ASV	Amplicon sequence variant
PCoA	Principal coordinate analysis
PERMANOVA	Permutational multivariate analysis of variance
LPS	Lipopolysaccharides
TLR	Toll-like receptor
TCA	Tricarboxylic acid

Declarations

Ethics approval
The study was approved by AMILI Institutional Review Board (reference number 2020/0501) in accordance with ethical guidelines.

Consent for publication
Not applicable.

Competing interests
KYT is an employee of AMILI Pte Ltd. JWJL is a co-founder of AMILI and serves as a member of the scientific advisory board. CWC and RPR are data consultants to AMILI. KPC is an employee of Pacific Biosciences. All other authors declare no financial or non-financial competing interests.

References
1. Wilmoth JR, Bas D, Mukherjee S, Hanif N. World social report 2023: leaving no one behind in an ageing world. UN; 2023.
2. Boehme M, Guzzetta KE, Bastiaanssen TF, Van De Wouw M, Moloney GM, Gual-Grau A, et al. Microbiota from young mice counteracts selective age-associated behavioral deficits. Nat Aging. 2021;1(8):666–76.
3. Ghosh TS, Shanahan F, O'Toole PW. The gut microbiome as a modulator of healthy ageing. Nat Reviews Gastroenterol Hepatol. 2022;19(9):565–84.
4. Herzog EL, Wäfler M, Keller I, Wolf S, Zinkernagel MS, Zysset-Burri DC. The importance of age in compositional and functional profiling of the human intestinal microbiome. PLoS ONE. 2021;16(10):e0258505.
5. Wilmanski T, Diener C, Rappaport N, Patwardhan S, Wiedrick J, Lapidus J, et al. Gut microbiome pattern reflects healthy ageing and predicts survival in humans. Nat Metabolism. 2021;3(2):274–86.
6. Kong F, Hua Y, Zeng B, Ning R, Li Y, Zhao J. Gut microbiota signatures of longevity. Curr Biol. 2016;26(18):R832–3.
7. Wang F, Yu T, Huang G, Cai D, Liang X, Su H, et al. Gut microbiota community and its assembly associated with age and diet in Chinese centenarians. J Microbiol Biotechnol. 2015;25(8):1195–204.
8. Chaudhari DS, Dhotre DP, Agarwal DM, Gaike AH, Bhalerao D, Jadhav P, et al. Gut, oral and skin microbiome of Indian patrilineal families reveal perceptible association with age. Sci Rep. 2020;10(1):5685.
9. Takagi T, Naito Y, Inoue R, Kashiwagi S, Uchiyama K, Mizushima K, et al. Differences in gut microbiota associated with age, sex, and stool consistency in healthy Japanese subjects. J Gastroenterol. 2019;54:53–63.
10. Gehrig JL, Portik DM, Driscoll MD, Jackson E, Chakraborty S, Gratalo D, et al. Finding the right fit: evaluation of short-read and long-read sequencing approaches to maximize the utility of clinical microbiome data. Microb Genomics. 2022;8(3):000794.
11. Johnson JS, Spakowicz DJ, Hong B-Y, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun. 2019;10(1):5029.
12. Wenger AM, Peluso P, Rowell WJ, Chang P-C, Hall RJ, Concepcion GT, et al. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nat Biotechnol. 2019;37(10):1155–62.
13. Sun D-L, Jiang X, Wu QL, Zhou N-Y. Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. Appl Environ Microbiol. 2013;79(19):5962–9.
14. Whitton C, Ho JCY, Tay Z, Rebello SA, Lu Y, Ong CN, Van Dam RM. Relative validity and reproducibility of a food frequency questionnaire for assessing dietary intakes in a multi-ethnic Asian population using 24-h dietary recalls and biomarkers. Nutrients. 2017;9(10):1059.
15. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581–3.
16. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17(1):10–2.
17. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for Bacteria and Archaea. Nat Biotechnol. 2020;38(9):1079–86.
18. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and all-species living tree project (LTP) taxonomic frameworks. Nucleic Acids Res. 2014;42(D1):D643–8.
19. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal database project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2014;42(D1):D633–42.
20. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44(D1):D733–45.
21. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol. 2020;38(6):685–8.
22. Lahti L, Shetty S. Introduction to the microbiome R package. Preprint at https://microbiome github io/tutorials. 2018.
23. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome datasets are compositional: and this is not optional. Front Microbiol. 2017;8:294209.
24. Oksanen J, Blanchet F, Friendly M, Kindt R, Legendre P, Mcglinn D et al. Vegan: community ecology package. Ordination methods, diversity analysis and other functions for community and vegetation ecologists, 05–26. Version 2.5-1. 2019.
25. Lin H, Peddada SD. Multigroup analysis of compositions of microbiomes with covariate adjustments and repeated measures. Nat Methods. 2024;21(1):83–91.
26. Ghosh TS, Das M, Jeffery IB, O'Toole PW. Adjusting for age improves identification of gut microbiome alterations in multiple diseases. eLife. 2020;9:e50240.
27. Woodmansey EJ, McMurdo ME, Macfarlane GT, Macfarlane S. Comparison of compositions and metabolic activities of fecal microbiotas in young adults and in antibiotic-treated and non-antibiotic-treated elderly subjects. Appl Environ Microbiol. 2004;70(10):6113–22.
28. Odamaki T, Kato K, Sugahara H, Hashikura N, Takahashi S, Xiao J-z, et al. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. BMC Microbiol. 2016;16:1–12.
29. de la Cuesta-Zuluaga J, Kelley ST, Chen Y, Escobar JS, Mueller NT, Ley RE, et al. Age-and sex-dependent patterns of gut microbial diversity in human adults. mSystems. 2019;4(4). https://doi.org/10.1128/msystems. 00261 – 19.
30. Dwiyanto J, Ayub Q, Lee SM, Foo SC, Chong CW, Rahman S. Geographical separation and ethnic origin influence the human gut microbial composition: a meta-analysis from a Malaysian perspective. Microb Genomics. 2021;7(8):000619.
31. La-Ongkham O, Nakphaichit M, Nakayama J, Keawsompong S, Nitisinprasert S. Age-related changes in the gut microbiota and the core gut microbiome of healthy Thai humans. 3 Biotech. 2020;10:1–14.
32. Singh H, Torralba MG, Moncera KJ, DiLello L, Petrini J, Nelson KE, Pieper R. Gastro-intestinal and oral microbiome signatures associated with healthy aging. Geroscience. 2019;41(6):907–21.
33. Leite G, Pimentel M, Barlow GM, Chang C, Hosseini A, Wang J et al. Age and the aging process significantly alter the small bowel microbiome. Cell Rep. 2021;36(13).
34. Biagi E, Nylund L, Candela M, Ostan R, Bucci L, Pini E, et al. Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. PLoS ONE. 2010;5(5):e10667.
35. Dong X, Guthrie BG, Alexander M, Noecker C, Ramirez L, Glasser NR, et al. Genetic manipulation of the human gut bacterium Eggerthella lenta reveals a widespread family of transcriptional regulators. Nat Commun. 2022;13(1):7624.
36. Haiser HJ, Gootenberg DB, Chatman K, Sirasani G, Balskus EP, Turnbaugh PJ. Predicting and manipulating cardiac drug inactivation by the human gut bacterium Eggerthella lenta. Science. 2013;341(6143):295–8.
37. Kageyama A, Benno Y. Catenibacterium mitsuokai gen. nov., sp. nov., a gram-positive anaerobic bacterium isolated from human faeces. Int J Syst Evol MicroBiol. 2000;50(4):1595–9.
38. Brahe LK, Le Chatelier E, Prifti E, Pons N, Kennedy S, Hansen T, et al. Specific gut microbiota features and metabolic markers in postmenopausal women with obesity. Nutr Diabetes. 2015;5(6):e159–e.
39. Cui G, Li S, Ye H, Yang Y, Jia X, Lin M, et al. Gut microbiome and frailty: insight from genetic correlation and mendelian randomization. Gut Microbes. 2023;15(2):2282795.
40. Schmidt TS, Raes J, Bork P. The human gut microbiome: from association to modulation. Cell. 2018;172(6):1198–215.
41. Whitfield C, Williams DM, Kelly SD. Lipopolysaccharide O-antigens—bacterial glycans made to measure. J Biol Chem. 2020;295(31):10593–609.

Toh *et al. Gut Pathogens*   (2024) 16:56

Page 10 of 10

42. Park BS, Lee J-O. Recognition of lipopolysaccharide pattern by TLR4 complexes. Exp Mol Med. 2013;45(12):e66–e.
43. Wang Y, Ma M, Dai W, Shang Q, Yu G. Bacteroides salyersiae is a potent chondroitin sulfate-degrading species in the human gut microbiota. Microbiome. 2024;12(1):41.
44. Zmuda AJ, Kang X, Wissbroecker KB, Freund Saxhaug K, Costa KC, Hegeman AD, Niehaus TD. A universal metabolite repair enzyme removes a strong inhibitor of the TCA cycle. Nat Commun. 2024;15(1):846.
45. Martínez-Reyes I, Chandel NS. Mitochondrial TCA cycle metabolites control physiology and disease. Nat Commun. 2020;11(1):102.
46. An R, Wilms E, Masclee AA, Smidt H, Zoetendal EG, Jonkers D. Age-dependent changes in GI physiology and microbiota: time to reconsider? Gut. 2018;67(12):2213–22.