

Comprehensive Analysis of Gut Microbiota of a Healthy Population and Covariates Affecting Microbial Variation in Two Large Japanese Cohorts

Jonguk Park

National Institute of Biomedical Innovation, Health and Nutrition <https://orcid.org/0000-0003-4872-385X>

Kumiko Kato

Morinaga Milk Industry Co., Ltd.

Haruka Murakami

National Institutes of Biomedical Innovation, Health and Nutrition

Koji Hosomi

National Institutes of Biomedical Innovation, Health and Nutrition

Kumpei Tanisawa

National Institutes of Biomedical Innovation, Health and Nutrition

Takashi Nakagata

National Institutes of Biomedical Innovation, Health and Nutrition

Harumi Ohno

National Institutes of Biomedical Innovation, Health and Nutrition

Kana Konishi

National Institutes of Biomedical Innovation, Health and Nutrition

Hitoshi Kawashima

National Institutes of Biomedical Innovation, Health and Nutrition

Yi-An Chen

National Institutes of Biomedical Innovation, Health and Nutrition

Attayeb Mohsen

National Institutes Biomedical Innovation, Health and Nutrition

Jin-zhong Xiao

Morinaga Milk Industry CO., Ltd.

Toshitaka Odamaki

Morinaga Milk Industry Co., Ltd.

Jun Kunisawa

National Institutes of Biomedical Innovation, and Health and Nutrition

Kenji Mizuguchi

National Institutes of Biomedical Innovation, Health and Nutrition

Motohiko Miyachi (✉ miyachi@nibiohn.go.jp)

National Institutes of Biomedical Innovation, Health and Nutrition, Tokyo <https://orcid.org/0000-0002-1146-0905>

Research

Keywords: 16S rRNA, gut microbiota, covariates, large cohort, Japanese population

Posted Date: October 14th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-89427/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Microbiology on May 20th, 2021. See the published version at <https://doi.org/10.1186/s12866-021-02215-0>.

Abstract

Background: Inter-individual variations in gut microbiota composition are observed even among healthy populations. As gut microbiota in Japanese populations exhibit a unique composition compared to others, it is informative to conduct large-scale Japanese cohort studies for a comprehensive understanding of the association between healthy gut microbiota and host status. The aim of the present study was to elucidate the integrated and non-redundant factors associated with gut microbiota composition within the Japanese population by 16S rRNA sequencing of fecal samples and questionnaire-based covariate analysis.

Results: A total of 1,596 healthy Japanese individuals participated in this study via two independent cohorts, NIBIOHN cohort (n = 954) and MORINAGA cohort (n = 642). Gut microbiota composition was described and the interaction of these microorganisms with metadata parameters such as anthropometric measurements, bowel habits, medical history, and lifestyle were obtained. Thirteen genera, including *Alistipes*, *Anaerostipes*, *Bacteroides*, *Bifidobacterium*, *Blautia*, *Eubacterium halli* group, *Faecalibacterium*, *Fusicatenibacter*, *Lachnoclostridium*, *Parabacteroides*, *Prevotella_9*, *Roseburia* and *Subdoligranurum* were predominant among the two cohorts. On the basis of univariate analysis for overall microbiome variation, 18 matching variables exhibited significant association in both cohorts. A stepwise redundancy analysis revealed that there were four common covariates, Bristol Stool Scale (BSS) scores, gender, age, and defecation frequency, displaying non-redundant association with gut microbial variance.

Conclusions: We conducted a comprehensive analysis of gut microbiota in healthy Japanese individuals, based on two independent cohorts, and obtained reliable evidence that questionnaire-based covariates such as frequency of bowel movement and dietary intake affects the microbial composition of the gut. To our knowledge, this was the first study to investigate integrated and non-redundant factors associated with gut microbiota among Japanese populations.

Background

Trillions of microorganisms constitute the gut microbiota in the human gut; it is a complex ecosystem and is directly or indirectly related to intestinal [1] and metabolic [2, 3] diseases, and neurological disorders [4, 5]. Owing to inter-individual variations in the gut microbiota composition among healthy populations [6], it is important to comprehend the characteristics of “healthy” gut microbiota and its determinants for the maintenance of human health and well-being [7]. Major host factors that have been reported to affect the gut microbial community structure include: age [8–10], ethnicity [11, 12], geography [13], host genetics [14], gender [15], and life style factors such as dietary habits [16], smoking [17], exercise [18], and medications of non-antibiotics [19] and antibiotics [20]. Certain patterns and transitions in the gut microbiota composition are observed with age [8] and altered dietary intake [21–23]. Furthermore, intestinal transit time, which is measured directly or assessed using Bristol Stool Scale (BSS) scores [24], is also a factor influencing the composition and diversity of gut microbiota [25, 26].

The association between intestinal transit time and gut microbiota is complicated and interactive [27, 28], and therefore, studies involving the analysis of transit time as one of the investigating factors will help us to improve our understanding of the gut microbiome.

Several reports have investigated the relationship between gut microbiota and variable factors, however, most of those reports identified only a few factors without integration and non-redundant estimation of host and environmental impact. In 2016, Falony *et al.* [29] and Zhernakova *et al.* [30] reported the estimation of variable factors in Belgian and Dutch cohorts. However, those studies only compared results separately obtained from two independent cohorts and ignored the disagreement of some variable factors among different cohorts. Significant effects of determinants such as age, BSS score, and body mass index (BMI) on gut microbiota were analyzed in two independent cohorts in the study by Falony *et al.* [29]. These determinants are reported to considerably affect not only European populations, but also cause gut microbiota variations in Chinese populations [31].

As the gut microbiome of Japanese populations exhibit a unique composition [32], it is informative to conduct large-scale analyses of Japanese cohorts for comprehensive understanding of the association between healthy gut microbiota and host status. Three previous reports [8, 33, 34] described determinants of healthy gut microbiota in Japanese cohorts and analyzed the relationship of gut microbiota with these determinants in more than one hundred subjects, however, the scale and metagenomic data analysis were comparatively limited [29, 30]. Thus, an integrated view of the non-redundant factors affecting gut microbiota composition among the Japanese population is yet to be established.

In this study, we analyzed gut microbiota composition, including the structure and interaction of gut microorganisms with metadata parameters such as anthropometric measurements, bowel habits, medical history, and lifestyles of Japanese populations. We obtained data from two independent cohorts, NIBIOHN cohort (v20.1) and MORINAGA cohort (v20.1) including 954 and 642 Japanese individuals, respectively. Moreover, we have not only compared the statistical results from our two cohorts but also unified the sets of phenotypic parameters used in the two studies as much as possible. The comparative data analysis between the two large distinct cohorts improved the robustness and reliability of our results. This is the first report to identify the non-redundant association of host factors with gut microbiota among Japanese populations.

Materials And Methods

Study population

We evaluated fecal samples and metadata of community-dwelling Japanese volunteers from two independent cohorts: health and nutrition-based cohort study conducted by National Institutes of Biomedical Innovation, Health and Nutrition (NIBIOHN), hereafter referred to as NIBIOHN cohort, and FAECES-02 study conducted by Morinaga Milk Industry, hereafter referred to as MORINAGA cohort.

In the NIBIOHN cohort, a total of 954 healthy Japanese adult volunteers were analyzed (age range of 20 – 80, male: 391, female: 563). In the MORINAGA cohort, a total of 642 healthy Japanese adult volunteers were analyzed (age range of 21 – 83, male: 129, female: 513). There were some differences in the characteristics of the two cohorts, in particular the distribution of residential area of the participants was extremely different; NIBIOHN cohort comprised of individuals from a limited area with a high concentration of subjects in each area, and MORINAGA cohort comprised of individuals from a dispersed and continuous area (Fig. S1). Most of the participants in the MORINAGA cohort were customers who purchased products of the Morinaga Milk Industry Co., Ltd. including dairy products or probiotics supplement. All participants of both cohorts had no history of cancer, cardiovascular, liver, or gastrointestinal diseases, and candidates who consumed antibiotics, laxatives, and antifatulents within two weeks prior to the study and whose data quality did not satisfy the criteria described below were excluded. Informed consent was obtained from all participants. These cohort studies were approved by the Ethics Committee of National Institutes of Biomedical Innovation, Health and Nutrition (Osaka, Japan), and by the Ethics Committee of Japan Conference of Clinical Research (Tokyo, Japan) and adhered to all guidelines.

Questionnaire about covariates

All participants answered a questionnaire on anthropometric data, medical history, fecal characteristics, dietary habits, physical activity, and sleep; a total of 76 items common to two cohorts were extracted as variables (Table S1). Fecal volume, form and color were evaluated by using our assessment card tool for NIBIOHN cohort as described previously [35] with slight modifications related to volume and color for MORINAGA cohort. To assess daily dietary intake, we asked subjects to answer a different questionnaire, the Brief Self-Administered Diet History Questionnaire (BDHQ), to estimate the intake volume of 58 food and beverage items consumed in the preceding month [36] (Table S1). The subjects whose estimated total energy intake was under 600 kcal or over 4,000 kcal were excluded due to lack of data reliability.

Fecal sampling

Stool samples of subjects from both cohorts were obtained, and an aliquot was mixed with 3 ml of guanidine thiocyanate solution (TechnoSuruga Laboratory, Shizuoka, Japan) prior to its transport to the laboratory.

DNA extraction and 16S rRNA gene amplicon sequencing

Fecal sample mixtures were mechanically disrupted by bead beating method using 0.1 mm glass beads, DNA was extracted using an automated extraction machine, and the V3-V4 region of the bacterial 16S rRNA gene was amplified by PCR. All samples from both cohorts were processed similarly, with some modifications as mentioned below.

For NIBIOHN cohort, the fecal sample mixtures were mechanically disrupted using Cell Destroyer PS1000 (Bio Medical Science, Tokyo, Japan). DNA was extracted by using Gene Prep Star PI-80X device (Kurabo Industries, Osaka, Japan). The V3-V4 region of the 16S rRNA gene was amplified using KOD-Plus-v2 (Toyobo, Osaka, Japan) and sequenced by paired-end method using Illumina MiSeq instrument and the MiSeq v3 Reagent Kit (Illumina, San Diego, CA, United States). All the steps from fecal sampling to 16S rRNA gene amplicon sequencing was performed according to a previously described protocol [37].

For MORINAGA cohort, the fecal sample mixture was mechanically disrupted using FastPrep-24 5G (MP Biomedicals, Santa Ana, CA, United States). DNA was extracted by using Gene Prep Star PI-480 device (Kurabo Industries). The V3-V4 region of the 16S rRNA gene was amplified using TaKaRa Ex Taq HS Kit (Takara Bio, Kusatsu, Japan) and sequenced by paired-end method using Illumina MiSeq instrument and the MiSeq v3 Reagent Kit (Illumina). 16S rRNA gene amplicon sequencing was performed according to a previously described protocol [38]. 10,000 reads per sample for NIBIOHN cohort and 5,000 reads per sample for MORINAGA cohort were randomly selected for further analysis. Samples with insufficient read numbers were resequenced, and samples with repeated insufficient read numbers were thereafter excluded.

Bioinformatics analysis

The obtained paired end FASTQ data were trimmed and merged before selection of the operational taxonomic units (OTUs). The OTU classification and diversity analysis were performed using QIIME pipeline (v1.9.1) [39]. All the steps from FASTQ trimming to gut microbiota diversity analysis were automatically performed according to a previously described method [40]. The OTUs were clustered against SILVA 128 reference database [41] at 97% similarity using USEARCH algorithm [42]. Taxonomic classification was performed using SILVA 128 reference database to the genus level taxa (hereafter referred as genera). Taxonomy name is expressed with a specific taxonomy name based on the SILVA database phylogenetic classification standard (<https://www.arb-silva.de/browser/ssu/>).

Statistical analysis

The output of QIIME pipeline in Biom table format was imported and analyzed in R (version 3.5.1). The alpha-diversity indices were calculated by the *estimate_richness* function in the “phyloseq” R-package. For alpha-diversity comparison analysis between NIBIOHN and MORINAGA cohorts, 5,000 reads were randomly reselected from NIBIOHN cohort. The beta-diversity index, calculated by Bray-Curtis distance using genus level data, was generated using the *vegdist* function in the “vegan” R-package. For enterotype analysis, we used Jensen-Shannon divergence (JSD) using a previously described method [6]. Principal coordinate analysis (PCoA) was performed using the *dudi.pco* function in the “ade4” R-package. Covariates of gut microbiome β -diversity were identified by calculating the association between continuous or categorical phenotypes and genus-level community coordinate with *envfit* function in the “vegan” R-package. This function performs MANOVA and linear correlations for categorical and

continuous variables, respectively. Over all metadata categories, 68 covariates in the NIBIOHN cohort and 32 covariates in the MORINAGA cohort were identified. To identify non-redundant determinants of microbiota variation, the covariates selected by *envfit* function were sub-selected by forward stepwise redundancy analysis on genus-level community ordination calculated by Bray-Curtis distance with the *ordiR2step* function in the “vegan” R-package. The dominant bacteria from phylum to genus level were defined as the mean of the distribution of bacterial composition with at least 1% correlation analysis. We used the Wilcoxon rank sum test (*wilcox.test* function in “stats” R-package) and Spearman correlation analysis (*cor* function in “stats” R-package) for comparison and correlation analysis, respectively. The comparison analysis of metadata was based on summary statistics of data derived from each cohort. Statistical analysis in this study was performed using R (version 3.5.1). Heatmaps were created using “corrplot” and “superheat” R-package, and PCoA figures and Boxplots were created using R package “ggplot2”. All statistical tests were two-sided, with a p-value < 0.05 considered significant.

Data availability

The DNA sequences corresponding to the 16S rRNA gene for NIBIOHN cohort have been deposited in DNA Databank of Japan (DDBJ) under accession numbers DRA010837 – DRA010841, and the DNA sequences for MORINAGA cohort have been deposited in DDBJ under accession numbers DRA009764 – DRA009767.

Results

Distribution of gut bacterial community in healthy individuals of two independent cohorts

First, we described the gut bacterial community structure of two large-scale healthy Japanese cohorts based on genus-level PCoA and enterotype analysis by partitioning around medoids (PAM) clustering using JSD (Fig. 1a and Fig. S2a, S2b). Values of the Calinski-Harabasz index (CH index) suggested that each cohort was divided into three clusters (Fig. 1b), which were characterized by the predominance of *Bacteroides*, *Prevotella_9*, and *Faecalibacterium* (in the NIBIOHN cohort) or *Bifidobacterium* (in the MORINAGA cohort). The ratio of each enterotype were 37:16:47 (*Bacteroides*-enterotype:*Prevotella*-enterotype:*Faecalibacterium*-enterotype) in the NIBIOHN cohort and 50:8:42 (*Bacteroides*-enterotype:*Prevotella*-enterotype:*Bifidobacterium*-enterotype) in the MORINAGA cohort. These results indicated that *Faecalibacterium* in the NIBIOHN cohort and *Bifidobacterium* in the MORINAGA cohort largely contributed to the individual differences in characteristics and distribution of the gut bacterial community. PCoA based on the integrated data from the two cohorts showed that no obvious difference in overall distribution between the two cohorts was observed (Fig. 1c). We also integrated the two cohorts and analyzed the enterotype. The integrated data of Japanese gut microbiota could be divided into two enterotypes based on the CH-index (Fig. S1d), *Bacteroides*-enterotype and *Prevotella*-enterotype (Fig. S2c). The integrated data also showed that the other enterotype-related genera of each cohort,

Faecalibacterium and *Bifidobacterium*, exhibited similar directionalities on gut microbial variation (Fig. S2c). The comparison of the alpha diversity indices showed differences between the two cohorts; NIBIOHN cohort contained a significantly higher number of observed OTUs and species richness Chao1 related to alpha diversity index (p-value < 0.01), whereas MORINAGA cohort was significantly higher in the Shannon and Simpson diversity index related to alpha diversity evenness (p-value < 0.05 and p-value < 0.01, respectively) (Fig. 1d). Next, we compared the predominant genus of gut microbiota in the two cohorts. Thirteen genera including *Alistipes*, *Anaerostipes*, *Bacteroides*, *Bifidobacterium*, *Blautia*, *Eubacterium halli* group, *Faecalibacterium*, *Fusicatenibacter*, *Lachnoclostridium*, *Parabacteroides*, *Prevotella_9*, *Roseburia* and *Subdoligranulum* were the predominant bacterial genera and were common among the two cohorts (Fig. 1e). The distribution of these genera was slightly different between the two cohorts. NIBIOHN cohort showed a significant abundance of *Bacteroides*, *Faecalibacterium*, and *Roseburia*, and MORINAGA cohort showed an abundance of *Bifidobacterium*, *Blautia*, *Lachnoclostridium*, *Fusicatenibacter*.

Identification of metadata covarying with gut microbiota in Japanese populations

Our results highlighted the common determinants that lead to variations in gut microbiota in the Japanese population. We tested a total of 134 variables to identify covariates associated with gut bacteria (Table S1). The comparison of metadata among the two cohorts indicated there were certain cohort-specific features (Table S2), for example, intake of dairy products and area of residence (Fig. S2). Univariate analysis, using the *envfit* function, for overall variations in the microbiome revealed 18 matching variables to be significantly associated with microbiome composition in both cohorts (p-value < 0.05, Fig. 2a, Table S3). A majority of the matching variables indicated a comparatively high effect size in each cohort. In particular, bowel habit-related variables showed strong associations with gut microbiota. Five anthropometric variables such as height, weight, age, gender, and BMI, were also common between the two cohorts. Although several diet-related variables were particularly significant, alcoholic drinks (sake, beer, shochu, whiskey, and wine), sweets (cake and Japanese cake), and fatty fish were common in the two cohorts. In addition, a high consumption of beer, among all alcoholic drinks, (Table S2) was also common. In the lifestyle category, the frequency of social drinking and total weekly working hours were common in the two cohorts. There were no significant common variables in the medical history or physical activity category. Some variables also showed similar directionalities for the ordination of gut microbiota composition in the two cohorts, for example, stool odor, BSS, and a feeling of exhilaration during bowel movement, and frequency of bowel movement per week (Fig. 2b).

We observed a redundancy among covariates as there was an overlap in the directionalities of variables (Fig. 2b and 2c). To identify non-redundant covariates of microbiome variations, we performed a forward stepwise redundancy analysis (*vegan* : *ordiR2 step* function). The number of non-redundant covariates were 12 out of 68 variables in the NIBIOHN cohort, and 10 out of 32 variables in the MORINAGA cohort, and their cumulative effect size on microbial community variation were 10.7% in the NIBIOHN cohort and

7.1% in the MORINAGA cohort, respectively. Furthermore, we demonstrated that four covariates, BSS, gender, age, and frequency of bowel movement per week, matched in the two cohorts (Fig. 2c). In addition, although the proposed question in the questionnaire was different in the two cohorts, factors related to beer consumption were identified as non-redundant factors. The other non-redundant covariates showed diverse characteristic features in each cohort. In particular, the residential area of participants had no significant influence on the gut microbial variance in the MORINAGA cohort, but showed the highest effect size on gut microbial variance in the NIBIOHN cohort (p-value < 0.001). Furthermore, the Spearman correlation analysis based on the dominant-genus data indicated that there were corresponding associations between these four factors (BSS, gender, age, and frequency of bowel movement) and gut bacteria (Table S4), as well as overall microbiome community variations. BSS showed inverse correlation with *Bifidobacterium* and *Alistipes*. The differences between genders were illustrated by high prevalence of *Prevotella_9* in men, and *Alistipes*, *Bifidobacterium*, *Faecalibacterium* and *Subdoligranulum* in women. The relative abundance of *Roseburia* increased with age, whereas that of *Blautia* and *Parabacteroides* decreased with age. Bowel movement frequencies per week were positively associated with the abundance of *Subdoligranulum* and *Alistipes*, and negatively associated with the presence of *Blautia*.

Discussion

In the present study, we report that not only the predominant bacterial genera (Fig. 1e), but also their interactions between metadata covariates were common (Fig. 2c) in two independent cohorts and seemed to show the general characteristics of healthy gut microbiota in the Japanese. First, we investigated the overall structural characteristics by using enterotype analysis. Although one of the dominant genera, *Faecalibacterium* in the NIBIOHN cohort and *Bifidobacterium* in the MORINAGA cohort, seemed to be different between the two cohorts (Fig. 1a, 1b), integrated data indicated that these genera had a similar third directionality on gut microbial variations (Fig. S2c). Because the directionalities of the two genera did not exactly correspond, the third directionality was not chosen as an enterotype. Taken together with a previous report analyzing the relationship between enterotypes and assemblages [43], the gut microbiota could be divided into three clusters.

To clarify the features of gut microbiota in Japanese populations in detail, we focused on the dominant genus in each cohort, and showed 13 genera were common in the two cohorts (Fig. 1e). Compared with the dominant genera in a dataset containing a total of 2,186 North American and European individuals [44], most of our dominant genera were common, but the average relative abundance varied. The characteristic differences in abundance were that of *Bifidobacterium*, which was 4.4% in the NIBIOHN cohort, 8.4% in the MORINAGA cohort, while 1.4% in the previous report [44], and that of *Blautia*, which was 4.7% in the NIBIOHN cohort, 6.5% in the MORINAGA cohort (Fig. 1e), while 2.9% in the previous report [44]. The higher abundance of *Bifidobacterium* and *Blautia* is a characteristic feature of gut microbiota in Japanese individuals [32]. Altogether, the dominant genera seemed to be common regardless of the population, however, the abundance of genera showed population-specific features.

There were some cohort-specific features in gut bacteria (Fig. 1c) and alpha diversity indices (Fig. 1d) in the two cohorts, even though the same Japanese population. It was suggested that these differences between the two cohorts were due to subjects having different backgrounds as indicated in Table S2. Together with a previous study [38], which reported that consumption of dairy products increases *Bifidobacterium* abundance in Japanese population, the higher intake of normal fat dairy products in the MORINAGA cohort compared to the NIBIOHN cohort (Table S2) was thought to be one of the reasons for a significant difference in *Bifidobacterium* abundance. Most of the subjects in the MORINAGA cohorts were customers who purchased products from the Morinaga Milk Industry Co., Ltd.

The comparison of results by *envfit* analysis between the two Japanese cohorts revealed 18 common covariates, showing significant association with structural variations in microbiota (Fig. 2a). Seven common covariates including BSS, height, weight, BMI, alcohol consumption (beer, rice wine, and low alcohol liquor), age, and gender, were in agreement with previous reports of three different cohorts in Belgium [29], Netherland [30] and China [31]. It appears that regardless of the population, these seven covariates commonly associate with structural differences in gut microbiota. Significant common covariates involved in microbial variation in the three cohorts [29–31] of previous studies, such as smoking, disease state, medication, fruit and meat consumption, were not observed in the NIBIOHN and MORINAGA cohorts, while the significant association of cake and fatty fish consumption was not found in previous cohorts [29–31]. These differences may be attributed to distinctive dietary habits between populations. For instance, Japanese population show the highest intake of seafood in the world [45], which may be a contributing factor to gut microbial variations. Although Falony *et al.* [29] revealed that medication resulted in the largest total variance and interacted with other covariate-microbiota associations, the target of our study was different and the effects of medication showed no significant association with healthy gut microbiota.

A forward stepwise redundancy analysis showed a remarkable cumulative effect size on community variation, specifically 10.7% in the NIBIOHN cohort and 7.1% in the MORINAGA cohort (Fig. 2c); effect size in both Japanese cohorts were comparable to 7.7% reported in Belgian population [29]. These results indicated that the proportion of gut microbial variation explained by questionnaire-based covariates seemed to be approximately 10% regardless of the population or number of covariates, suggesting there were several additional intrinsic or extrinsic contributors such as immunity, host genetics, bacterial-bacterial interaction, as well as unknown factors [46]. Non-redundant determinants such as BSS, gender, and age were common in our two cohorts (Fig. 3c), in accordance with the previous report involving a Belgian cohort [29], indicating that these covariates are common among healthy populations independent of factors like country of origin. Falony *et al.* [29] reported that stool consistency showed the largest effect size on bacterial variations; BSS score for gut_stool_shape showed a higher non-redundant effect on total composition variation in Japanese cohorts (Fig. 2c). Gender [15] and age [9] intricately associated with life style factors such as dietary habits, therefore, the selection of gender and age as non-redundant covariates was justified. Interestingly, non-redundant analysis also showed population- or cohort-specific results. Beer consumption and related covariates were identified as non-redundant covariates in each cohort (Fig. 2c), presumably because beer was the most popular alcoholic drink in the

Japanese cohorts (Table S2). This result is in line with a previous study [47] stating that alcohol affects the composition of gut bacteria.

Another cohort-specific feature with a high effect size was observed to be residential area in the NIBIOHN cohort, whereas no association was observed in the MORINAGA cohort (Table S3). The distribution of residential area was extremely different between the two cohorts; NIBIOHN cohort comprised of individuals from a limited area with a high concentration of subjects in each area (Fig S2), and MORINAGA cohort comprised of individuals from a dispersed and continuous area. Notably, a previous report, involving a 16S rRNA gene analysis of fecal samples collected from 516 healthy Japanese adults residing in various regions of Japan [33], was very similar to the MORINAGA cohort, and demonstrated no association between residential area and gut microbiota variation. However, the residential area was found to have the highest influence on gut bacterial variance in the NIBIOHN cohort. Further large-scale nationwide cohort studies are required to understand the effect of residential area on total gut microbial variance among Japanese populations.

Besides the common interactions of covariates and compositional variation, some associations between gut bacteria and covariates were also common between the two cohorts (Table S4). The negative association of BSS score with *Alistipes* was in accordance with a previous report [29], whereas the negative association with *Bifidobacterium* was not previously reported. Furthermore, the reported association of Christenellaceae, Mogibacteriaceae, and Rikenellaceae with bowel movement frequency in the Japanese population [33] was not observed in our study. Interestingly, regardless of the influence of BSS score and frequency of bowel movement on colon transit time, both covariates showed non-redundant associations with inter-individual gut microbial variation. This difference was represented by the fact that *Subdoligranulum* and *Blautia* only associated with defecation frequency. This difference highlighted the complex association between colon transit time, gut microbiota, and diet [27]. In relation to the common association of gender and gut bacteria, the higher abundance of *Bifidobacterium* in women and *Prevotella_9* in men (Table S4) was in agreement with a previous report of Japanese gut microbiota [34], indicating the possibility of population-specific results. In contrast, the higher abundance of *Faecalibacterium* and *Alistipes* in women (Table S4) is first reported in the present study. The common association of age with the abundance of *Blautia*, *Parabacteroides*, and *Roseburia* was not shown in the previous report [29] and seemed to be population-specific. A decrease of *Bifidobacterium* abundance with age in the NIBIOHN cohort was in accordance with previous reports [8, 29], identifying *Bifidobacterium* as an adult-enriched bacteria. However, in the MORINAGA cohort, *Bifidobacterium* showed no significant association with age. One of the reasons for this discrepancy was presumed to be the unique dietary habits, which was high consumption of dairy products, and the high relative abundance of *Bifidobacterium* in the MORINAGA cohort.

In this study, we reported the characteristics of gut microbiota and comprehensively examined the major microbiome-associated variables in Japanese gut microbiota using two independent large-scale cohort data. The comparison between these large, distinct cohorts provided reliability and robustness to our study. One of the limitations of our study was that it was a cross-sectional study and did not show the

causal relationships between gut microbiota and metadata. Prospective observational or interventional studies would be required to delineate these relationships. Another limitation was that all participants were applicants, which may be a source of potential bias in this study. Random sampling of subjects from all over the country or a larger sample size would be essential for overcoming this bias.

Conclusions

Our results described the features of healthy Japanese gut microbiota, including the 13 predominant genera and high abundance of *Bifidobacterium* and *Blautia* in two independent cohorts. The comparison of two diverse and independent cohorts increased the reliability of our results. Furthermore, 18 covariates, including anthropometric measurements, bowel habits, lifestyle, and dietary habits, commonly associated with gut bacterial variations were identified. The BSS score, gender, age, and frequency of bowel movement independently affected gut microbiota composition and were inferred to be essential factors influencing microbial communities in healthy Japanese populations. To our knowledge, this was the first study to report integrated and non-redundant associations of factors affecting structural characteristics of gut microbiota among Japanese populations.

List Of Abbreviations

National Institutes of Biomedical Innovation, Health and Nutrition (NIBIOHN)

Bristol Stool Scale (BSS)

partitioning around medoids (PAM)

body mass index (BMI)

Brief Self-Administered Diet History Questionnaire (BDHQ)

operational taxonomic unit (OTU)

Calinski-Harabasz index (CH index)

Jensen-Shannon divergence (JSD)

principal coordinate analysis (PCoA)

Declarations

Ethics approval and consent to participate

NIBIOHN cohort study was approved by the Ethics Committee of National Institutes of Biomedical Innovation, Health and Nutrition (Approval No. KENEI3-06, KENEI78-05, KENEI116, 154-09, 177-07, Osaka,

Japan) and MORINAGA cohort study was approved by the Ethics Committee of Japan Conference of Clinical Research (Approval No. 246, Tokyo, Japan), in accordance to their respective guidelines.

Consent for publication

Not applicable.

Availability of data and material

The DNA sequences corresponding to the 16S rRNA gene for NIBIOHN cohort have been deposited in DDBJ under accession numbers DRA010837 – DRA010841, and the sequences for MORINAGA cohort have been deposited in DDBJ under accession numbers DRA009764 – DRA009767. The datasets supporting the results of this study are included within the article and Additional files.

Competing interests

The authors of this manuscript have the following competing interests: Kumiko Kato, Jin-zhong Xiao and Toshitaka Odamaki are employees of Morinaga Milk Industry Co., Ltd.

Funding

This study was supported by Grant-in-Aid for Scientific Research (C) 6K00944 and by Health and Labour Sciences Research Grant 201709002B (to MM); the Japan Agency for Medical Research and Development (AMED; JP20gm1010006h004 to JK); the Ministry of Health and Welfare of Japan and Public/Private R&D Investment Strategic Expansion Program: PRISM (to JK, KM, and MM); the Japan Society for the Promotion of Science KAKENHI (Grant nos. JP18H05280 to KH), and the Canon Foundation (to JK).

Authors' contributions

MM, TO, KM and JK contributed to the study conception; JP, KK (Kato) and HM contributed to the design study; HM, KK (Kato), HO, KK (Konishi), KT, TN and KH contributed to data collection; KK (Kato) and KH contributed to the experiments; JP, YAC, AM and HK contributed to the bioinformatics analysis; JP, KK (Kato), HM, KH contributed to the data analysis; JP, KK (Kato), MM, TO, KM, JK, HK, HM, KH and JX contributed to the interpretation data; JP and KK (Kato) contributed to the drafting manuscript; All authors reviewed the manuscript and the final version of the manuscript was approved by all authors.

Acknowledgements

The authors are grateful to all of the participants who provided fecal samples for use in this research. We are grateful to the support members of each laboratory at NIBIOHN and Morinaga Milk Industry, especially Aya Mizuno and Yuki Yamauchi. We would like to thank Editage (www.editage.com) for English language editing.

References

1. Mizutani S, Yamada T, Yachida S. Significance of the gut microbiome in multistep colorectal carcinogenesis. *Cancer Sci.* 2020;111:766–73. doi: 10.1111/cas.14298.
2. Gurung M, Li Z, You H, Rodrigues R, Jump DB, Morgun A, et al. Role of gut microbiota in type 2 diabetes pathophysiology. *EBioMedicine.* 2020;51:102590. doi:10.1016/j.ebiom.2019.11.051.
3. Tang WHW, Kitai T, Hazen SL. Gut microbiota in cardiovascular health and disease. *Circ Res.* 2017;120:1183–96. doi: 10.1161/CIRCRESAHA.117.309715.
4. Kho ZY, Lal SK. The human gut microbiome - a potential controller of wellness and disease. *Front Microbiol.* 2018;9:1835. doi:10.3389/fmicb.2018.01835.
5. Baizabal-Carvalho JF, Alonso-Juarez M. The link between gut dysbiosis and neuroinflammation in Parkinson's disease. *Neuroscience.* 2020;432:160–73. doi:10.1016/j.neuroscience.2020.02.030.
6. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. *Nature.* 2011;473:174–80. doi: 10.1038/nature09944.
7. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. *Cell.* 2012;148:1258–70. doi:10.1016/j.cell.2012.01.035.
8. 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;16:90. doi:10.1186/s12866-016-0708-5.
9. An R, Wilms E, Masclee AAM, Smidt H, Zoetendal EG, Jonkers D. Age-dependent changes in GI physiology and microbiota: time to reconsider? *Gut.* 2018;67:2213–22. doi: 10.1136/gutjnl-2017-315542.
10. O'Toole PW, Jeffery IB. Gut microbiota and aging. *Science.* 2015;350:1214–5. doi: 10.1126/science.aac8469.
11. Gupta VK, Paul S, Dutta C. Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. *Front Microbiol.* 2017;8:1162. doi:10.3389/fmicb.2017.01162.
12. Deschasaux M, Bouter KE, Prodan A, Levin E, Groen AK, Herrema H, et al. Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nat Med.* 2018;24:1526–31. doi:10.1038/s41591-018-0160-1.

13. Gaulke CA, Sharpton TJ. The influence of ethnicity and geography on human gut microbiome composition. *Nat Med.* 2018;24:1495–6. doi:10.1038/s41591-018-0210-8.
14. Kurilshikov A, Wijmenga C, Fu J, Zhernakova A. Host genetics and gut microbiome: challenges and perspectives. *Trends Immunol.* 2017;38:633–47. doi:10.1016/j.it.2017.06.003.
15. Kim YS, Unno T, Kim B-Y, Park M-S. Sex differences in gut microbiota. *World J Mens Health.* 2020;38:48–60. doi: 10.5534/wjmh.190009.
16. Zmora N, Suez J, Elinav E. You are what you eat: diet, health and the gut microbiota. *Nat Rev Gastroenterol Hepatol.* 2019;16:35–56. doi:10.1038/s41575-018-0061-2.
17. Lee SH, Yun Y, Kim SJ, Lee E-J, Chang Y, Ryu S, et al. Association between cigarette smoking status and composition of gut microbiota: population-based cross-sectional study. *J Clin Med.* 2018;7:282. doi:10.3390/jcm7090282.
18. Barton W, Penney NC, Cronin O, Garcia-Perez I, Molloy MG, Holmes E, et al. The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. *Gut.* 2018;67:625–33. doi:10.1136/gutjnl-2016-313627.
19. Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature.* 2018;555:623–8. doi:10.1038/nature25979.
20. Zhao J, Murray S, Lipuma JJ. Modeling the impact of antibiotic exposure on human microbiota. *Sci Rep.* 2014;4:4345. doi:10.1038/srep04345.
21. Kushida M, Sugawara S, Asano M, Yamamoto K, Fukuda S, Tsuduki T. Effects of the 1975 Japanese diet on the gut microbiota in younger adults. *J Nutr Biochem.* 2019;64:121–7. doi:10.1016/j.jnutbio.2018.10.011.
22. Meslier V, Laiola M, Roager HM, De Filippis F, Roume H, Quinquis B, et al. Mediterranean diet intervention in overweight and obese subjects lowers plasma cholesterol and causes changes in the gut microbiome and metabolome independently of energy intake. *Gut.* 2020;69:1258–68. doi:10.1136/gutjnl-2019-320438.
23. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature.* 2014;505:559–63. doi:10.1038/nature12820.
24. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. *Scand J Gastroenterol.* 1997;32:920–4.
25. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut.* 2016;65:57–62. doi: 10.1136/gutjnl-2015-309618.
26. Roager HM, Hansen LBS, Bahl MI, Frandsen HL, Carvalho V, Gøbel RJ, et al. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat Microbiol.* 2016;1:16093. doi:10.1038/nmicrobiol.2016.93.
27. Kashyap PC, Marcobal A, Ursell LK, Larauche M, Duboc H, Earle KA, et al. Complex interactions among diet, gastrointestinal transit, and gut microbiota in humanized mice. *Gastroenterology.* 2013;144:967–77. doi:10.1053/j.gastro.2013.01.047.

28. Ge X, Zhao W, Ding C, Tian H, Xu L, Wang H, et al. Potential role of fecal microbiota from patients with slow transit constipation in the regulation of gastrointestinal motility. *Sci Rep.* 2017;7:441. doi:10.1038/s41598-017-00612-y.
29. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. *Science.* 2016;352:560–4. doi:10.1126/science.aad3503.
30. Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science.* 2016;352:565–9. doi: 10.1126/science.aad3369.
31. He Y, Wu W, Zheng H-M, Li P, McDonald D, Sheng H-F, et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat Med.* 2018;24:1532–5. doi:10.1038/s41591-018-0164-x.
32. Nishijima S, Suda W, Oshima K, Kim S-W, Hirose Y, Morita H, et al. The gut microbiome of healthy Japanese and its microbial and functional uniqueness. *DNA Res.* 2016;23:125–33. doi:10.1093/dnares/dsw002.
33. Oki K, Toyama M, Banno T, Chonan O, Benno Y, Watanabe K. Comprehensive analysis of the fecal microbiota of healthy Japanese adults reveals a new bacterial lineage associated with a phenotype characterized by a high frequency of bowel movements and a lean body type. *BMC Microbiol.* 2016;16:284. doi:10.1186/s12866-016-0898-x.
34. 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. doi:10.1007/s00535-018-1488-5.
35. Ohno H, Murakami H, Tanisawa K, Konishi K, Miyachi M. Validity of an observational assessment tool for multifaceted evaluation of faecal condition. *Sci Rep.* 2019;9:3760. doi:10.1038/s41598-019-40178-5.
36. Kobayashi S, Murakami K, Sasaki S, Okubo H, Hirota N, Notsu A, et al. Comparison of relative validity of food group intakes estimated by comprehensive and brief-type self-administered diet history questionnaires against 16 d dietary records in Japanese adults. *Public Health Nutr.* 2011;14:1200–11. doi: 10.1017/S1368980011000504.
37. Hosomi K, Ohno H, Murakami H, Natsume-Kitatani Y, Tanisawa K, Hirata S, et al. Method for preparing DNA from feces in guanidine thiocyanate solution affects 16S rRNA-based profiling of human microbiota diversity. *Sci Rep.* 2017;7:4339. doi:10.1038/s41598-017-04511-0.
38. Kato K, Ishida S, Tanaka M, Mitsuyama E, Xiao J-Z, Odamaki T. Association between functional lactase variants and a high abundance of *Bifidobacterium* in the gut of healthy Japanese people. *PLoS One.* 2018;13:e0206189. doi:10.1371/journal.pone.0206189.
39. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods.* 2010;7:335–6. doi: 10.1038/nmeth.f.303.

40. Mohsen A, Park J, Chen YA, Kawashima H, Mizuguchi K. Impact of quality trimming on the efficiency of reads joining and diversity analysis of Illumina paired-end reads in the context of QIIME1 and QIIME2 microbiome analysis frameworks. *BMC Bioinformatics*. 2019;20:581. doi: 10.1186/s12859-019-3187-5.
41. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41:D590–6. doi: 10.1093/nar/gks1219.
42. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. 2011;27:2194–200. doi:10.1093/bioinformatics/btr381.
43. Hosoda S, Nishijima S, Fukunaga T, Hattori M, Hamada M. Revealing the microbial assemblage structure in the human gut microbiome using latent Dirichlet allocation. *Microbiome*. 2020;8:95. doi:10.1186/s40168-020-00864-3.
44. Feng Y, Duan Y, Xu Z, Lyu N, Liu F, Liang S, et al. An examination of data from the American Gut Project reveals that the dominance of the genus *Bifidobacterium* is associated with the diversity and robustness of the gut microbiota. *Microbiologyopen*. 2019;8:e939. doi: 10.1002/mbo3.939.
45. Micha R, Khatibzadeh S, Shi P, Andrews KG, Engell RE, Mozaffarian D, et al. Global, regional and national consumption of major food groups in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys worldwide. *BMJ Open*. 2015;5:e008705. doi:10.1136/bmjopen-2015-008705.
46. Schmidt TSB, Raes J, Bork P. The human gut microbiome: from association to modulation. *Cell*. 2018;172:1198–215. doi:10.1016/j.cell.2018.02.044.
47. Bjørkhaug ST, Aanes H, Neupane SP, Bramness JG, Malvik S, Henriksen C, et al. Characterization of gut microbiota composition and functions in patients with chronic alcohol overconsumption. *Gut Microbes*. 2019;10:663–75. doi:10.1080/19490976.2019.1580097.

Figures

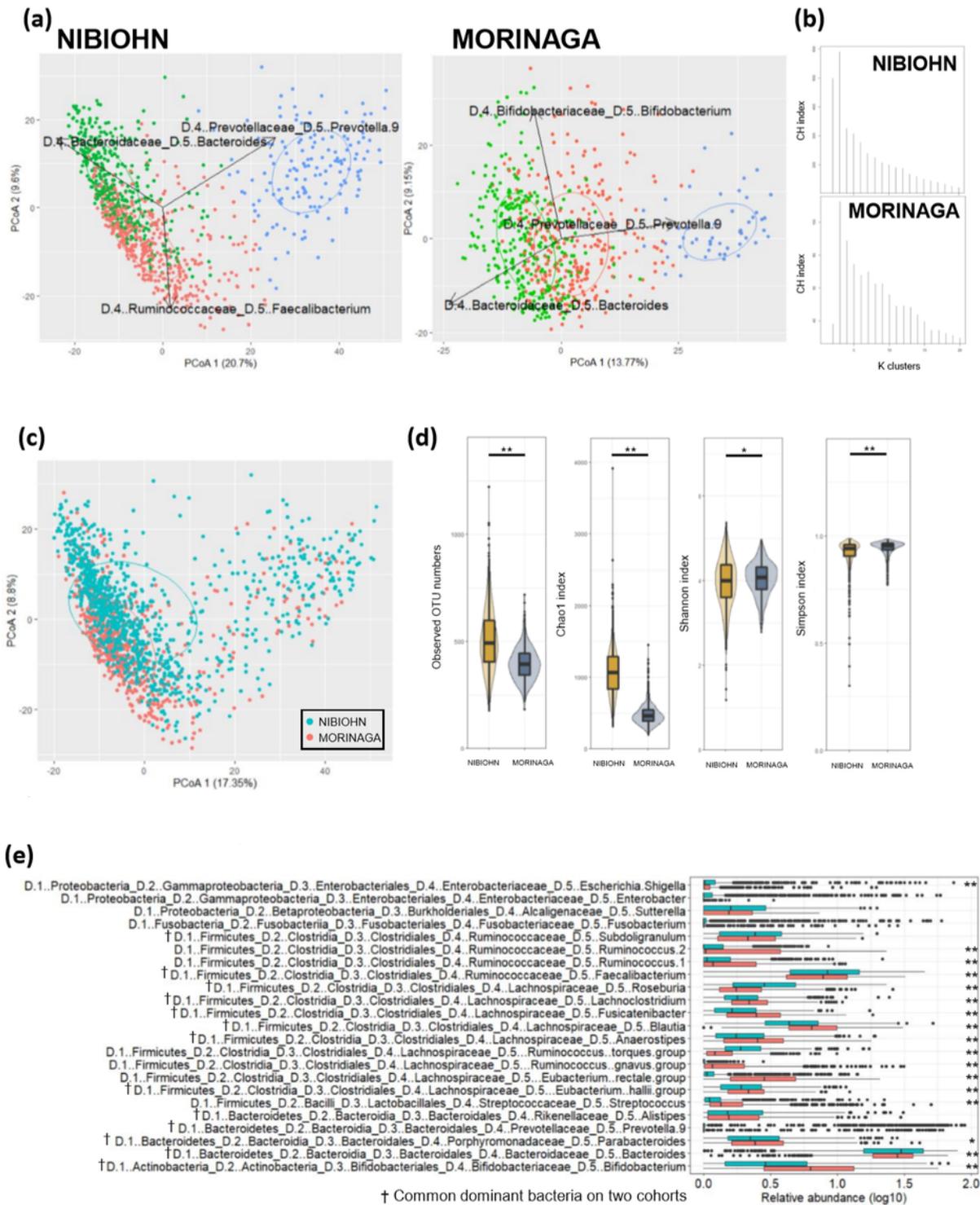


Figure 1

Gut microbiota distribution of the NIBIOHN cohort and MORINAGA cohort (a) Microbial community variation in each cohort represented by principal coordinates analysis (PCoA, genus-level JSD) and PAM clustering. Arrows indicated enterotype drivers; Bacteroides-enterotype (green), Prevotella-enterotype (blue), and Faecalibacterium-enterotype or Bifidobacterium-enterotype (red). (b) The estimated result of suitability of cluster number Calinski-Harabasz index (CH index) (c) The distribution of integrated data of

two cohorts (d) The distribution of alpha diversity indices of NIBIOHN cohort (yellow) and MORINAGA cohort (blue) (e) The dominant genus among the two cohorts and their composition. **p < 0.01, *p < 0.05 (Wilcoxon rank sum test).

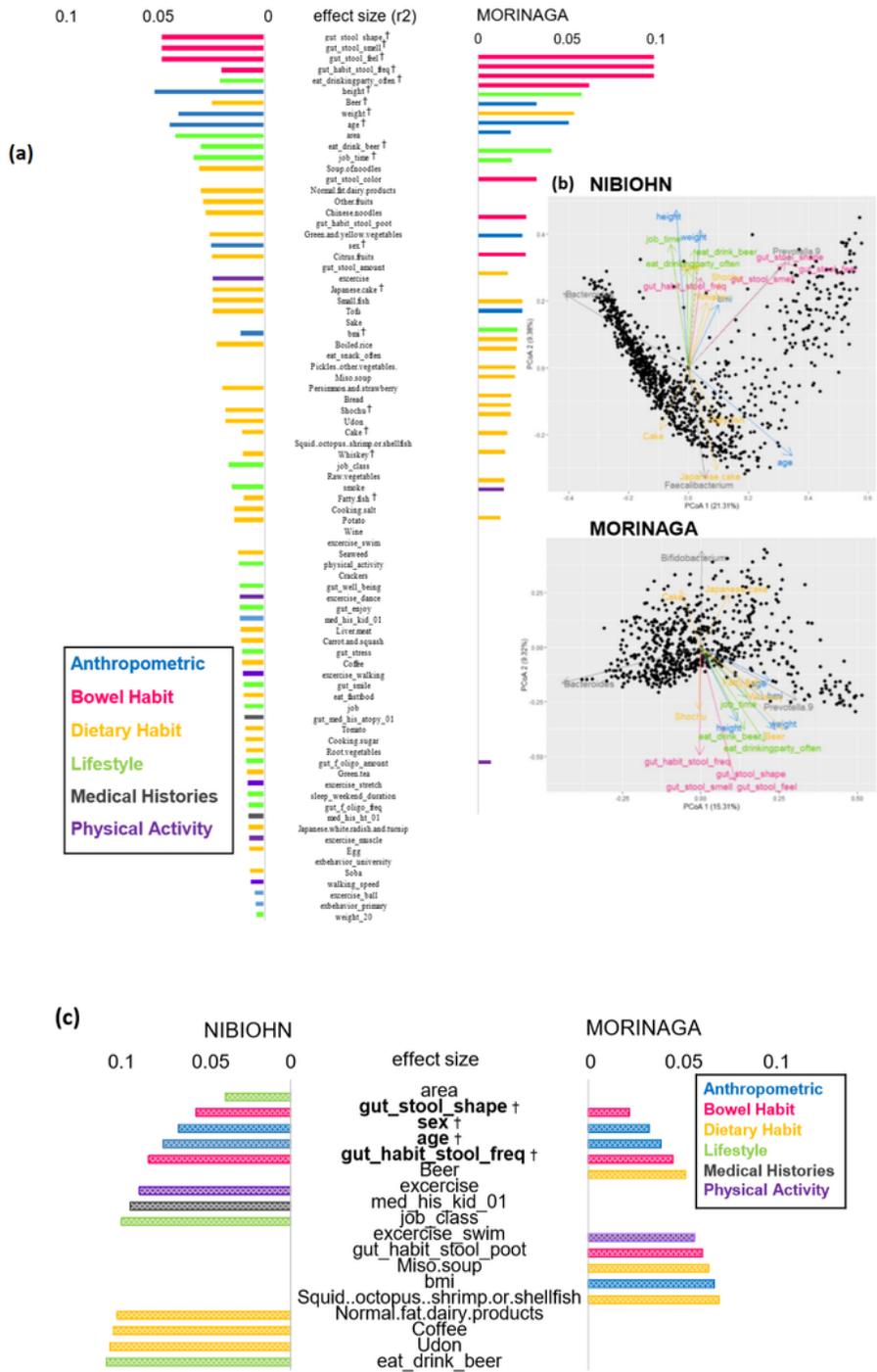


Figure 2

Effect size of microbiome covariates (a) Effect size identified in the NIBIOHN cohort (left) and the MORINAGA cohort (right). Factors are sorted according to their effect size and colored based on

metadata category (Table. S1) (b) PCoA-based on Bray-Curtis distance. Arrows show the ordination of 18 common covariates for overall microbiome community variation in the NIBIOHN cohort (top) and the MORINAGA cohort (bottom) (c) Cumulative effect size of non-redundant covariates. Microbial covariates selected by stepwise redundancy analysis in the NIBIOHN cohort (left) and the MORINAGA cohort (right). †Common covariates in the two cohorts.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile6.pptx](#)
- [Additionalfile5.pptx](#)
- [Additionalfile4.xlsx](#)
- [Additionalfile3.xlsx](#)
- [Additionalfile2.xlsx](#)
- [Additionalfile1.xlsx](#)