Genome Medicine

RESEARCH

Open Access

# Early-life and concurrent predictors of the healthy adolescent microbiome in a cohort study

Hannah E. Laue[1*], Amy D. Willis[2], Fang Wang[3], Melinda C. MacDougall[4], Yingying Xu[4,5], Margaret R. Karagas[6], Juliette C. Madan[6,7], Abby F. Fleisch[8,9], Bruce P. Lanphear[10], Kim M. Cecil[4,5,11,12], Kimberly Yolton[4,5,12], Aimin Chen[13], Jessie P. Buckley[14] and Joseph M. Braun[15]

## Abstract

**Background** The microbiome of adolescents is poorly understood, as are factors influencing its composition. We aimed to describe the healthy adolescent microbiome and identify early-life and concurrent predictors of its composition.

**Methods** We performed metagenomic sequencing of 247 fecal specimens from 167 adolescents aged 11–14 years participating in the Health Outcomes and Measures of the Environment (HOME) Study, a longitudinal pregnancy and birth cohort (Cincinnati, OH). We described common features of the adolescent gut microbiome and applied self-organizing maps (SOMs)—a machine-learning approach—to identify distinct microbial profiles ($n=4$). Using prospectively collected data on sociodemographic characteristics, lifestyle, diet, and sexual maturation, we identified early-life and concurrent factors associated with microbial diversity and phylum relative abundance with linear regression models and composition with Kruskal–Wallis and Fisher's exact tests.

**Results** We found that household income and other sociodemographic factors were consistent predictors of the microbiome, with higher income associated with lower diversity and differential relative abundances of Firmicutes (increased) and Actinobacteria (decreased). Sexual maturation, distinct from chronological age, was related to higher diversity in females and differences in phylum relative abundances and compositional profiles in both males and females.

**Conclusions** Our study suggests that adolescence is a unique window for gut microbial composition and that it may be shaped by both early-life and concurrent exposures, highlighting its potential in future epidemiologic research.

**Keywords** Gut microbiome, Adolescence, Metagenomics, Self-organizing maps

*Correspondence:
Hannah E. Laue
hlaue@umass.edu
Full list of author information is available at the end of the article

## Background

The gut microbiome—the microorganisms that reside in the gastrointestinal tract—is a critical component of human health [1]. Most prior microbiome research has been conducted in infants, adults, or clinical populations, leaving a critical gap in understanding the adolescent microbiome [2, 3]. The prevailing theory is that an infant's gut microbiome composition fluctuates over the first few years of life, with diversity steadily increasing until a diet of solid food is established [4, 5], and then reaches a mature, adult state [6]. However, the gut microbiome, which responds to hormonal changes during pregnancy [7–9] and menopause [10], may be sensitive to the changes of estrogen and testosterone during puberty. In fact, studies have found that subgingival [11] and skin [12] microbiomes change across stages of sexual maturation. The few studies of the adolescent gut microbiome also highlight differences in prevalent and abundant bacteria relative to adults and younger children [13–16]. Thus, adolescence may be a critical period for development and alteration of the gut microbiome.

In addition to clarifying distinctive features of the adolescent gut microbiome, it is crucial to understand factors that shape microbial communities in this life stage. Few prospective studies have examined the microbiome in relation to gestational and early-life exposures. Despite this, some studies have reported differences in the gut microbiome related to self-report of early-life factors like delivery mode and the initiation or duration of consumption of human milk that persist into adulthood [17, 18]. Elucidating the influence of early-life and concurrent factors on the adolescent microbiome using prospective data has potential implications for interventions to improve lifelong well-being.

We investigated the fecal microbiome of adolescents (ages 11–14 years) in the Health Outcomes and Measures of the Environment (HOME) Study. This prospective pregnancy and birth cohort has extensive longitudinal data collected through structured questionnaires, clinical assessments, and dietary recalls, allowing for a comprehensive analysis of early-life and concurrent factors associated with the adolescent microbiome.

## Methods

### Cohort and study participants

The Health Outcomes and Measures of the Environment (HOME) Study is a prospective pregnancy and birth cohort based in Cincinnati, OH, USA. Prior publications described the cohort and schedule of study visits from pregnancy through age 12 (range: 11–14) years [19, 20]. Briefly, 420 pregnant people were recruited during prenatal visits at Cincinnati metropolitan-area obstetric practices between 2003 and 2006. Participants were excluded if they were age < 18 years, > 19 weeks gestation, or planning to move before delivery. Additionally, because the cohort was designed to study prenatal lead exposure, participants had to live in homes build before 1978 when lead paint was banned for residential use [21]. Participants attended annual in-person study visits through age 5 years and again at age 8 and 12 years. Each visit included the collection of biospecimens, anthropometry, and questionnaires to capture sociodemographic characteristics. Two hundred fifty-six completed the 12-year study visit and 167 provided at least one stool sample [20]. Five sets of twins were included in analyses. Caregivers provided written, informed consent and adolescents provided written, informed assent. The institutional review boards (IRBs) of Cincinnati Children's Hospital Medical Center (CCHMC) and participating delivery hospitals approved this study. Dartmouth College, Brown University, and the University of Massachusetts Amherst deferred to the CCHMC IRB as the IRB of record.

### Stool collection

At the 12-year study visit, staff asked participants to collect up to three stools within the month following the study visit using Hemoccult II SENSA fecal occult blood test (FOBT) cards (Beckman Coulter, Brea, CA, USA). Stool samples were not collected at prior study visits. Females were instructed not to collect a stool while they were menstruating. For each stool collection ("sample"), participants were provided three FOBT cards, two labeled "end" and one labeled "middle," and were instructed to swab from each end and the middle of the stool for the appropriate cards (each a "specimen"; Additional file 1: Fig. S1). This resulted in up to nine distinct specimens from three stool samples from each participant. Participants were provided a toilet hat, all necessary collection materials, and detailed written instructions that were reviewed at the study visit. Participants mailed the FOBT cards to CCHMC, where they were stored at − 80 °C until shipment to the laboratory for analysis. For a subset of participants ($n = 20$), we extracted all specimens from the first stool sample and the "middle" specimens from subsequent two stool samples. To select the subset from the 95 (38 male and 57 female) participants who provided nine specimens, we restricted to those who had data on the timing of each stool, complete data for other research aims, and had variability in stool consistency, reported using the Bristol scale ($n = 45$; 17 male, 28 female; Additional file 1: Fig. S1) [22]. These were then narrowed to ten male and ten female participants based on medication use, symptoms (e.g., stomachache), duration between samples, and whether the participant urinated in the stool collection hat (Additional file 2:

Table S1). For the remaining participants, we extracted and sequenced DNA from the "middle" card of the first stool sample.

#### DNA extraction and sequencing

DNA was extracted from the FOBT cards at the Laboratory of Precision Environmental Health at Columbia Mailman School of Public Health following adapted protocols for DNA extraction from solid stool [23]. DNA was extracted from the eluent using the QIAamp PowerFecal Pro kit following the manufacturer's protocol (Qiagen, Hilden, Germany). Briefly, the stool specimen was cut from the FOBT card with a sterile scissor and homogenized in a bead-beating tube. Total genomic DNA was then isolated in a spin column and quantified with a Nano spectrophotometer. For downstream sequencing, total DNA (mean ± SD yield: 12,694 ± 11,212 ng) was normalized to 35 ng/µl in 40 µl and sent to the New York University Langone Genome Technology Center (RRID: SCR_017929; New York, USA) for shotgun metagenomic sequencing. Technicians pooled libraries with Illumina Flex (San Diego, CA, USA) library preparation and sequenced the specimens with an Illumina NovaSeq 6000 using an S2 flowcell to obtain 2× 150 bp paired-end reads.

#### Sequence processing

Raw sequences were uploaded to the National Center for Biotechnology Information Sequence Read Archive (Accession: PRJNA1139690) after the removal of human sequences [24]. Prior to processing, there were (mean ± SD) 34,265,115 ± 7,762,788 reads per specimen. Sequences were processed using the Biobakery pipeline in Python 3.7 [25]. Human and ribosomal RNA reads were removed, and sequences were trimmed based on quality scores (minimum: 25) with kneaddata (v0.12.0 [26] with Bowtie2 and trimmomatic v0.39 [27] using databases hg37 dec v0.1 for human reads and SILVA 128 LSU and SSU for ribosomal RNA reads). After this preprocessing step, specimens had a mean ± SD read depth of 30,322,565 ± 6,719,222. We aligned sequences to known microbial genomes and metagenome-assembled genomes (MAGs) using MetaPhlAn 4 with the CHOCOPhlAnSGB vJun23 database [28].

#### Metadata

We identified early-life and concurrent factors that may predict adolescent fecal microbial diversity and composition based on the extant literature of infant and adult gut microbiome studies. Delivery mode (vaginal, cesarean) and sex assigned at birth (male, female) were abstracted from medical records. Delivery mode was missing for 3 of the 167 participants who provided a stool sample. Adolescent's race/ethnicity was assessed via standardized questionnaire as a proxy for systemic racism and cultural factors related to the microbiome. We dichotomized race/ethnicity as non-Hispanic white and minoritized racial/ethnic groups (primarily non-Hispanic Black, with some American Indian, Asian/Pacific Islander, Hispanic) due to small sample size of non-Hispanic Black groups. Caregivers reported the duration of human milk consumption (exclusive and any) using standardized questionnaires at multiple times during the first 3 years of life [29]. We categorized these non-normally distributed variables as follows: for duration of any human milk consumption (exclusive or supplemented with formula or complementary foods)—fed for 6 months or less, or fed longer than 6 months; for duration of exclusive human milk consumption—2 days or fewer (normal duration of hospital stay after delivery [30]), or more than 2 days. Human milk consumption data were missing for 4 of the 167 participants who provided a stool sample. We assessed tobacco use and/or exposure of the pregnant person with serum cotinine concentrations at 16 weeks gestation. Serum cotinine concentrations were missing for 3 of the 167 participants who provided a stool sample. Based on prior findings, we categorized cotinine concentrations as no smoke exposure (< 0.015 ng/mL), exposure to second-hand smoke (0.015–3 ng/mL), and active smoking (≥ 3 ng/mL) for descriptive statistics [31], and used $\log_{10}$-transformed continuous cotinine concentrations in analyses.

We calculated participant's age at the age 12 study visit as the time (days) between the date of the study visit and the recorded date of birth. At the visit, caregivers reported their household income. Trained research staff measured participants' weights and standing heights in triplicate, from which we calculated age- and sex-standardized body mass index (BMI) *Z*-scores [32]. As previously described, participants were asked to evaluate their sexual maturation using pictograph Tanner scales with written descriptions (breast development and pubic hair growth for females and pubic hair growth only for males) [33, 34]. Three females and no males who contributed stool samples declined to report sexual maturation. Due to few participants reporting Tanner stages 1 and 5, we grouped stages 1 and 2 as "early" development and 4 and 5 as "advanced" development, as previously described [35]. Female participants and their caregivers were asked if the participant had reached menarche, and if so, at what age. One participant who provided a stool sample declined to report on their menarchal status. As previously described [34, 35], we collected a morning fasting blood sample at the study visit and stored serum samples at −80 °C until analysis. Serum was unavailable for 19 females and 8 males. We measured estradiol in females

with quantitative chemiluminescent immunoassay at Associated Regional and University Pathologist laboratories in Salt Lake City, UT. Testosterone was quantified in males by liquid chromatography tandem mass spectrometry at CCHMC. We $\log_{10}$-transformed hormone concentrations to reduce the influence of outliers.

At the time of each stool sample collection, participants reported their current antibiotic use and the stool's consistency using the Bristol scale. Briefly, participants were provided with cartoon images and descriptions of stool consistency ranging from 1 (most solid) to 7 (most liquid) and were asked to endorse the category that most closely resembled their stool. Based on prior literature, we grouped categories 1–3 as "constipated" and categories 6 and 7 as "diarrhea," with categories 4 and 5 representing "healthy" stool [36]. Two participants endorsed multiple Bristol categories (Additional file 2: Table S2).

At the age 12 study visit and on two subsequent days, study staff administered 24-h dietary recalls (one weekend day and two weekdays) using the Nutrition Data Systems for Research software. One participant who provided a stool sample did not complete the dietary recalls. Using these data, we computed 2010 Healthy Eating Index (HEI) scores, including average daily intake of soluble, insoluble, and total fiber [37]. For analyses, we considered the three fiber measures, total HEI scores, and the 12 HEI components as potential predictors of microbiome diversity and composition (Additional file 2: Table S3). Sparse components (greens and beans—58.7% zeroes; seafood and plant protein—43.1% zeroes) were treated dichotomously (any vs. none).

#### Statistical analysis

In the subset of participants ($n = 20$) with multiple specimens sequenced ($n = 5$ per participant), we sought to understand factors that contributed to variability in species-level diversity and composition. To assess the contribution of factors [subject, sample order, location on stool (end vs. middle), time between stool samples, stool consistency] to variability in Shannon and inverse Simpson indices and phylum relative abundances, we fit linear mixed effects models with nested random effects using the "nlme" R package sequentially adding nested variables [38, 39]. We determined the best fit model via the minimum Akaike's information criterion (AIC) and determined the proportion of variance explained by each variable with interclass correlation coefficients (ICCs). We calculated Bray–Curtis distances between observations from the species matrix to assess the separation between specimens and performed principal coordinates analysis [40, 41]. The first three components were then plotted against each other to visualize separation. We quantified the contribution of factors to variability in species-level composition using marginal adonis2 models with 9999 permutations from the "vegan" R package [40]. Briefly, we conducted sequential models with each predictor in the model individually and then adjusting for prior predictors and reported $R^2$ and $p$ values. Our final model included all variables that were significantly associated in the last sequential model.

We then conducted a descriptive analysis of the adolescent gut microbiome. To reduce the influence of autocorrelated specimens, we restricted to the first middle specimen for each participant ($n = 167$). We calculated specimen species diversity and richness with the Shannon and inverse Simpson indices using the "vegan" R package [40, 42, 43]. The inverse Simpson index gives more weight to the evenness of taxon relative abundance than the Shannon index, and, unlike the Simpson index, is not bound between zero and one. To describe microbiome composition, we explored prevalent (≥ 80% detection in specimens) and abundant (≥ 1% median relative abundance) species and genera.

We also sought to identify microbial composition profiles using self-organizing maps (SOMs) [44–46]. This unsupervised, machine-learning data-reduction method is related to multidimensional scaling and $k$-means clustering. Briefly, specimens with similar microbiome profiles cluster together. These clusters, or nodes, are arranged spatially based on similarity to each other to aid visualization and interpretation. For this analysis, we evaluated hexagonal SOMs ranging from 4 to 20 nodes in size, fitting robust centered log ratio (rclr) transformed relative abundances of prevalent genera (> 80% detection) as the correlated data variables (Additional file 1: Figs. S2, S3) [47]. We selected the optimal map size based on established cluster statistics including the ratio of within-cluster to between-cluster sum of squared errors, and the percent of variability in phylum relative abundance explained by the SOM (adjusted $R^2$) [44, 48]. Consideration was also given to maintaining sufficient observations in each node for downstream analyses. We then grouped nodes into clusters, prioritizing similarities in genera that previously were found to distinguish adolescent gut microbiomes from those of adults or younger children (Additional file 2: Table S4).

To estimate associations between predictors of microbiome diversity and composition, we restricted to the first middle specimen for each participant ($n = 167$) to eliminate correlation between observations from the same participant. We transformed phylum, genus, and species relative abundances using the rclr method to account for outliers and the compositional nature of the data [47]; thus, effect estimates can be interpreted as log-fold change in phylum geometric mean abundance compared to the average phylum/genus/species. We

estimated differences in Shannon and inverse Simpson indices and log-fold change in phylum, genus, and species relative abundances associated with early-life and concurrent predictors using linear regression. For continuous variables, we compared median and interquartile ranges (IQRs) across nodes and tested for significant differences using Kruskal–Wallis tests. For categorical predictors, we calculated Fisher's exact *p* values. When there were more than two ordered categories, we treated the predictor as an ordered factor in linear models to calculate a *p*-for-trend. For most analyses, we considered $p < 0.05$ as significant, but for genus and species associations we considered a more conservative $p < 0.01$ significant. For models in which sexual maturation was the predictor, we ran analyses separately in males and females. We hypothesized that the association of early-life and concurrent diet with the microbiome may be confounded by household income as a correlate of baseline health and that the association between sexual maturation and the microbiome may be confounded by age. Thus, we adjusted for these variables in sensitivity analyses. Also as a sensitivity analysis, we excluded the four participants who were taking antibiotics at the time of stool collection and repeated the analysis of predictors of microbial diversity and SOM node membership.

## Results

### Intra-individual gut microbiome variability

In a subset of participants ($n = 20$), we sequenced 5 specimens from 3 stool samples collected over 2 weeks, including three specimens from different locations on the first sample (Additional file 1: Fig. S1, Additional file 2: Table S1). Shannon diversity was predicted by participant, sample, and location within the stool sample, while inverse Simpson diversity was predicted by participant and sample (Additional file 2: Table S5). Sixty percent of Shannon diversity and 37% of inverse Simpson diversity was explained by participant, with sample explaining ~25% of diversity, and location within stool explaining <1% (Additional file 2: Table S6). Variability in relative abundance of Firmicutes (76%), Bacteroidota (75%), Actinobacteria (83%), and Proteobacteria (78%) was also attributable to participant, with 7–15% attributable to sample (Additional file 2: Tables S5, S6). The species-level composition of specimens from the same participant were more similar to each other than they were to other specimens (Fig. 1, Additional file 2: Table S7). Approximately 85% of the variability was explained by participant in univariate models. Bristol score (14.9%) and Bristol category (4.0%) were also significant univariate predictors of Bray–Curtis distances, although each explained only a small proportion of the variance. Our final model included participant (71.2% of variance), days between sample collection (0.3%), and Bristol score (1.8%).

### Adolescent gut microbiome composition

After restricting to one specimen per participant ($n = 167$), a median (interquartile range; IQR) of 120 (101, 135) species were detected per specimen, with a median (IQR) Shannon index of 3.5 (3.3, 3.7), reflecting moderate diversity. Firmicutes was the most abundant phylum on average, although there were differences across specimens (Fig. 2). Other prevalent phyla (sequenced in ≥80% of individuals) included Bacteroidetes, Actinobacteria, and Proteobacteria. We found 37 prevalent species (sequenced in ≥80% of individuals), with *Faecalibacterium prausnitzii* detected in all but one individual (Additional file 2: Table S8). *Eubacterium rectale* was the most abundant species on average [median relative abundance ($Med_{RA}$) =6.9%], followed by *Faecalibacterium prausnitzii* ($Med_{RA} = 6.0\%$). Twelve other species had an average median relative abundance greater than 1% (Additional file 2: Table S8). There were 33 prevalent genera (sequenced in ≥80% of individuals), with *Bacteroides* detected in all individuals and *Mediterraneibacter*, *Faecalibacterium*, and *Blautia* detected in all but one specimen (Additional file 2: Table S9). On average, the most abundant genera were an unclassified *Lachnospiraceae* genus ($Med_{RA} = 7.7\%$), *Bifidobacterium* ($Med_{RA} = 6.4\%$), and *Faecalibacterium* ($Med_{RA} = 6.2\%$; Additional file 2: Table S9).

Differences between specimens and moderate correlations between prevalent genus relative abundances (Additional file 1: Fig. S3) enabled us to identify genus-level microbial profiles using self-organizing maps (SOMs) in 167 specimens (one per participant). The best fit SOM had 15 nodes or profiles (Additional file 1: Figs. S2, S4). We grouped these nodes into four clusters (Fig. 2), preferentially grouping nodes with similar median relative abundances of genera previously used to discriminate fecal microbiome samples of adolescents from adults or younger children (Additional file 2: Table S4). Clusters A and C were characterized by high *Bifidobacterium* and *Bacteroides* (A) and high *Ruminococcus* (C). Cluster B was characterized by low relative abundance of a candidate genus (candidatus *Cibionibacter*) and high *Ruthenibacterium*. Cluster D was characterized by high relative abundance of *Flavonifractor*. Species-level Shannon diversity, but not inverse Simpson diversity, was associated with SOM cluster (Additional file 2: Table S10).

### Factors associated with the adolescent gut microbiome

To gain a better understanding of factors associated with the adolescent microbiome, we estimated differences in

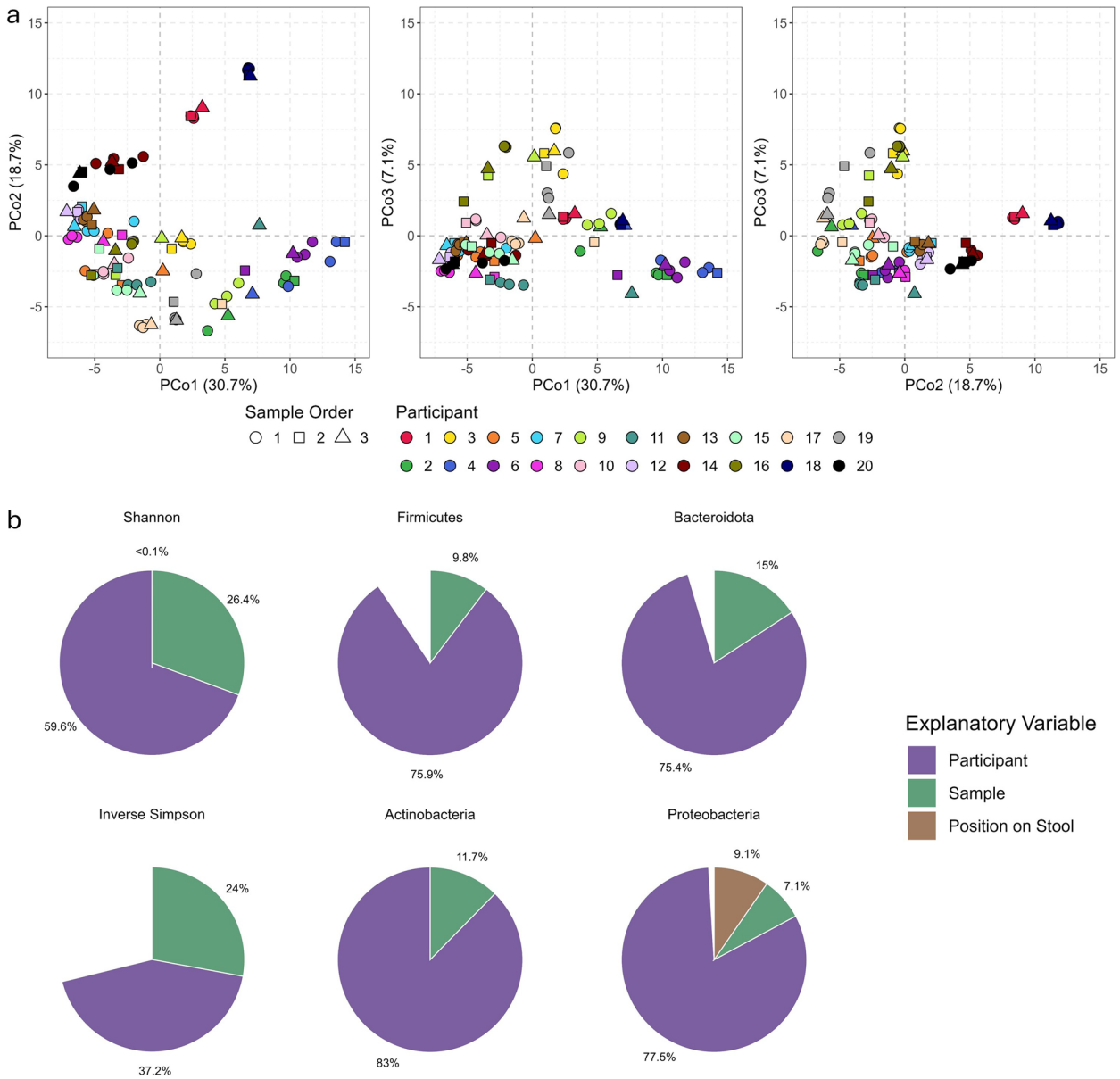


**Fig. 1** Variability of gut microbiome within and between stool samples. **a** Principal coordinate (PCo) plots for the first three PCos of Bray–Curtis distances. Percent variability explained by each PCo is in parentheses. Each point represents a specimen, with color reflecting participant (characteristics in Supplementary Table 1) and shape representing stool sample order. **b** Variance in diversity (Shannon, inverse Simpson) and phylum relative abundance (Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria) attributable to participant, sample, and position on stool (Shannon, inverse Simpson, Bacteroidetes). Remaining variability (white space) is unexplained. Exact values can be found in Supplementary Table 30

species-level diversity and relative abundances of prevalent phyla, genera, and species across factors collected prospectively or concurrent with the stool collection and examined how these factors differed across the four SOM clusters. These factors fell into four broad categories: early-life factors (sex assigned at birth, child's race/ethnicity, delivery mode, infant diet, gestational exposure to tobacco smoke), measures of sexual maturation (self-reported, hormone concentrations), non-dietary concurrent factors (age at follow-up, body mass index (BMI) *Z*-score, household income, stool consistency, current antibiotic use), and concurrent dietary components. The characteristics of participants who provided at least one stool sample were similar to those of all participants who completed the follow-up visit at 12 years of age (Table 1), although females who contributed a stool

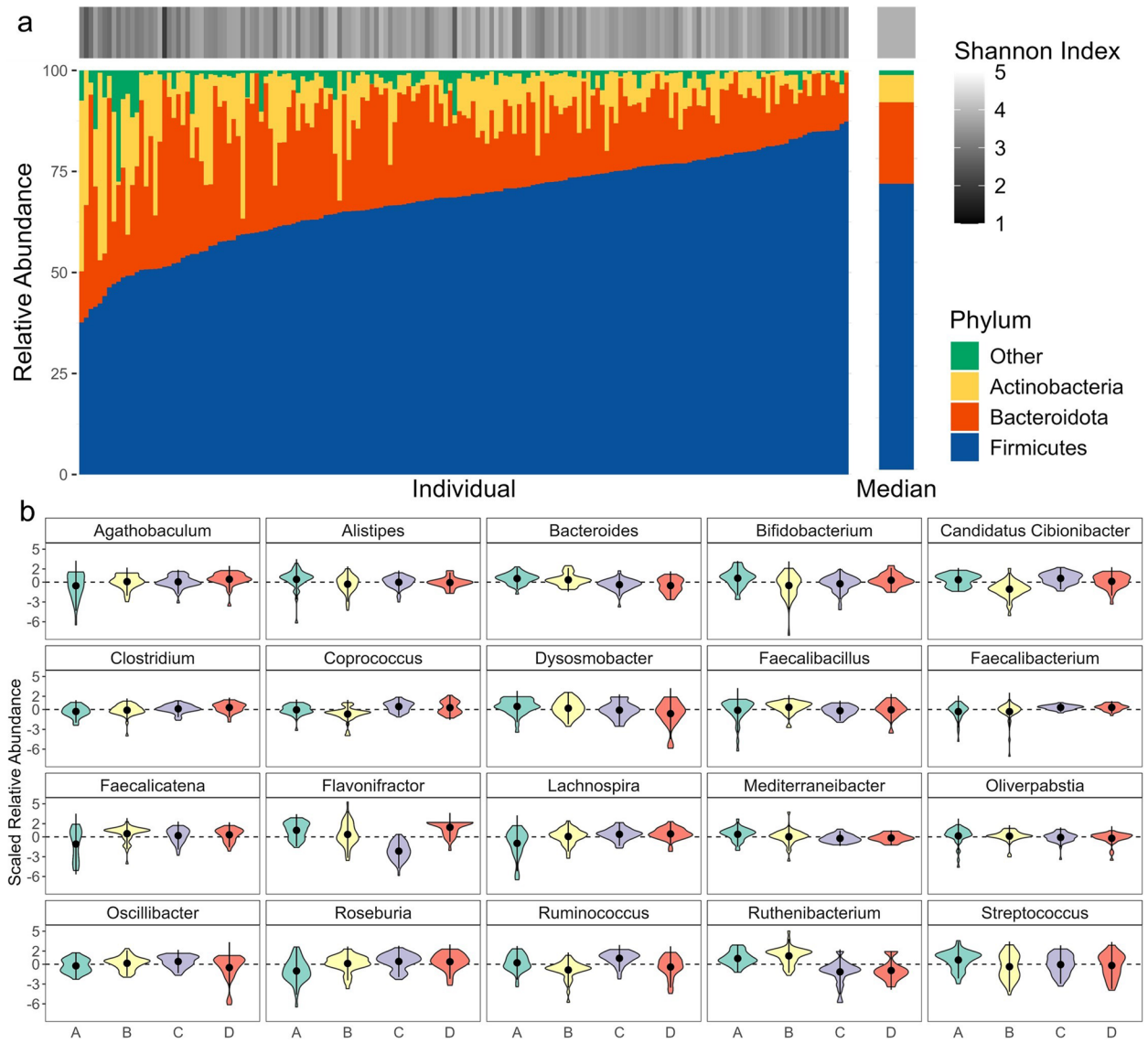


**Fig. 2** Distributions of phylum relative abundance in adolescent fecal microbiome specimens ($n = 247$) and common genus-based profiles ($n = 167$). **a** The relative abundance (y-axis) and Shannon index of each specimen (x-axis) compared to the median sample. Samples are arranged by increasing Firmicutes relative abundance. **b** Distribution of robust centered log ratio transformed and centered genus relative abundances (y-axis) in self-organizing map (SOM) clusters (x-axis). Only prevalent genera (≥ 80%) that have different median relative abundances in at least two clusters are shown

sample were more likely to be menarchal than those who did not contribute a stool sample.

### *Early-life factors*

Of the early-life factors investigated, race/ethnicity was the most consistently associated with the adolescent gut microbiome (Fig. 3). Non-Hispanic white participants had lower Actinobacteria [$\widehat{\beta} = -0.5$ log-fold difference (95%CI: $-0.8$, $-0.2$), $p = 0.002$] and modestly higher Proteobacteria [$\widehat{\beta} = 0.3$ log-fold difference (95%CI: $-0.1$, $0.6$), $p = 0.105$] than participants from marginalized populations, although the latter estimate did not reach statistical significance (Additional file 2: Table S11). Non-Hispanic white participants also had lower abundances of the genera *Bifidobacterium*, *Dorea*, *Coprococcus*, and *Oscillibacter*, and higher abundance of *Intestinibacter* and *Lachnospira* than participants from marginalized populations (Additional file 2: Table S12). However, these differences were not reflected in genus-level SOM cluster distribution of race (Additional file 2:

**Table 1** Characteristics of participants in the Health Outcomes and Measures of the Environment (HOME) Study followed through adolescence compared to those contributing at least one fecal sample [*n* (%), mean ± standard deviation, median (25 th %ile, 75 th %ile)]

| | **Participants followed through 12 years of age (*n* = 256)** | | **Participants contributing a fecal sample (*n* = 167)** | |
|---|---|---|---|---|
| Maternal serum cotinine at 16 weeks gestation | | | | |
| Not exposed (< 0.015 ng/mL) | 73 (28.5) | | 52 (31.1) | |
| Second-hand smoke (0.015–3 ng/mL) | 149 (58.2) | | 97 (58.1) | |
| Active smoker (≥ 3 ng/mL) | 29 (11.3) | | 15 (9.0) | |
| Missing | 5 (2.0) | | 3 (1.8) | |
| Child's race/ethnicity | | | | |
| Non-Hispanic white | 147 (57.4) | | 103 (61.7) | |
| Minoritized race/ethnicity | 109 (42.6) | | 64 (38.3) | |
| Delivery mode | | | | |
| Vaginal | 176 (68.8) | | 118 (70.7) | |
| Cesarean | 69 (27.0) | | 46 (27.5) | |
| Missing | 11 (4.3) | | 3 (1.8) | |
| Sex assigned at birth | | | | |
| Male | 113 (44.1) | | 72 (43.1) | |
| Female | 143 (55.9) | | 95 (56.9) | |
| Fed exclusively human milk for at least 2 days | | | | |
| Yes | 99 (38.7) | | 72 (43.1) | |
| No | 140 (54.7) | | 91 (54.5) | |
| Missing | 17 (6.6) | | 4 (2.4) | |
| Duration of human milk consumption | | | | |
| ≤ 6 months | 141 (55.1) | | 89 (53.3) | |
| > 6 months | 98 (38.3) | | 74 (44.3) | |
| Missing | 17 (6.6) | | 4 (2.4) | |
| Family annual income (USD) | 75,000 (35,000, 145,000) | | 85,000 (45,000, 145,000) | |
| Child's age at 12-year assessment (y) | 12.4 ± 0.7 | | 12.3 ± 0.7 | |
| Child's BMI *Z*-score at 12-year assessment | 0.4 ± 1.2 | | 0.4 ± 1.2 | |
| Current antibiotic use[a] | | | | |
| No | – | | 163 (97.6) | |
| Yes | – | | 4 (2.4) | |
| Tanner pubic hair stage | Male | Female | Male | Female |
| Early (1, 2) | 54 (47.8) | 36 (25.2) | 40 (55.6) | 21 (22.1) |
| Mid (3) | 30 (26.5) | 42 (29.4) | 21 (29.2) | 30 (31.6) |
| Advanced (4, 5) | 28 (24.8) | 62 (43.4) | 11 (15.3) | 41 (43.2) |
| Missing | 1 (0.9) | 3 (2.1) | – | 3 (3.2) |
| Tanner breast stage | | | | |
| Early (1, 2) | – | 31 (21.7) | – | 21 (22.1) |
| Mid (3) | – | 68 (47.6) | – | 42 (44.2) |
| Advanced (4, 5) | – | 41 (28.7) | – | 29 (30.5) |
| Missing | – | 3 (2.1) | | 3 (3.2) |
| Menarchal status | | | | |
| Non-menarchal | – | 73 (51) | – | 39 (41.1) |
| Menarchal | – | 64 (44.8) | – | 55 (57.9) |
| Missing | – | 6 (4.2) | – | 1 (1.1) |
| Estradiol (pg/dL) | – | 43 (23, 64) | – | 44 (19, 65) |
| Testosterone (ng/dL) | 138 (27, 383) | – | 120 (27, 323) | – |

[a] Data on current antibiotic use were collected as part of a survey administered at the time of stool sample collection, thus are missing for participants who did not provide a stool sample

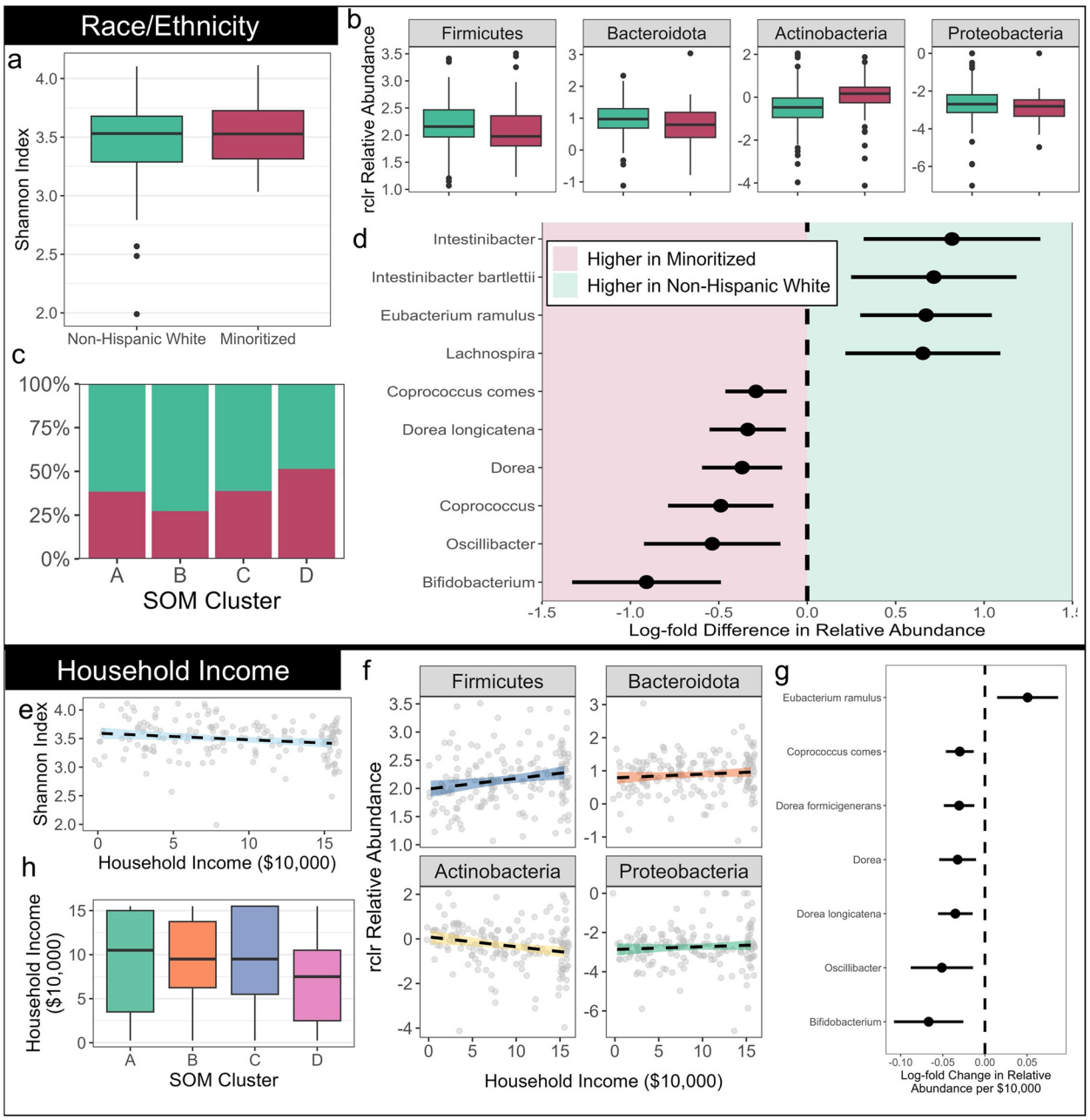


**Fig. 3** Selected associations of early-life (race/ethnicity; **a**–**d**) and concurrent (household income, **d**–**g**) predictors with adolescent gut microbiome diversity (Shannon index; **a**, **e**), robust centered log ratio (rclr) transformed phylum relative abundances (**b**, **f**), and genus-level self-organizing map (SOM) node membership (**c**, **h**) and rclr transformed genus and species relative abundances (**d**, **g**). Genus and species estimates with $p < 0.01$ are displayed

Table S10). Associations with species were consistent with genus-level associations; specifically, non-Hispanic white participants had lower *Coprococcus comes* and *Dorea longicatena* and higher *Eubacterium ramulus* than participants from marginalized populations (Additional file 2: Table S13).

Maternal gestational serum cotinine, an indicator of prenatal tobacco smoke exposure, was also consistently associated with differences in the adolescent gut microbiome. Higher concentrations were associated with increased relative abundance of Actinobacteria [$\widehat{\beta} = 0.2$ log-fold difference per tenfold cotinine increase (95%CI:

0.1, 0.4), $p = 0.001$] and decreased relative abundance of Proteobacteria [$\widehat{\beta} = -0.1$ log-fold difference per tenfold cotinine increase (95%CI: −0.3, 0.0), $p = 0.064$; Additional file 2: Table S11]. Higher gestational cotinine concentrations were negatively associated with abundance of *Bacteroides* and *Phocaeicola*, and positively associated with *Bifidobacterium* (Additional file 2: Table S12). Specifically, higher gestational cotinine was associated with abundance of *Bacteroides uniformis* [$\widehat{\beta} = -0.3$ log-fold difference per tenfold cotinine increase (− 0.45, −0.14), $p = 3.33E-04$] and *Dorea formicigenerans* [$\widehat{\beta} = 0.11$ log-fold difference per tenfold cotinine increase (0.03, 0.19), $p = 0.007$] (Additional file 2: Table S13). While early-life diet was not associated with microbiome composition at the phylum level, both exclusive consumption of human milk for at least 2 days and consumption of human milk for more than 6 months were each associated with unique differences in the abundances genera and species. The distribution of early-life diet trended toward being different across SOM clusters, with those in cluster D least likely to be fed exclusively human milk for at least 2 days (26.5%) and most likely to be fed human milk for more than 6 months (55.9%). Exclusive consumption of human milk for at least the first 2 days of life was associated with lower abundance of *Blautia wexlerae*, *Dorea formicigenerans*, and *Eubacterium rectale*, and consumption of human milk for more than 6 months was associated with higher abundances of *Oliverpabstia intestinalis* (Additional file 2: Table S13) as well as the genera *Roseburia* and *Eubacterium* (Additional file 2: Table S12). Duration of human milk consumption is strongly associated with socioeconomic factors; thus, we conducted a sensitivity analysis adjusting our early-diet models for household income, which did not meaningfully alter our findings, although the association with *Dorea formicigenerans* was somewhat attenuated (Additional file 2: Tables S12–S16). Sex assigned at birth and birth mode were not associated with adolescent gut microbiome composition (Additional file 2: Tables S10–S13), and no early-life factors were associated with Shannon or inverse Simpson diversity (Additional file 2: Tables S16, S17).

#### *Sexual maturation factors*

Self-reported sexual maturation was associated with gut microbial diversity and composition in males and females. We measured fasting morning serum gonadal hormones estradiol and testosterone in females and males, respectively, as objective markers of sexual maturation. Among females, tenfold higher morning serum estradiol was associated with 0.4 log-fold higher Proteobacteria abundance [(95%CI: 0.0, 0.9), $p = 0.046$; Fig. 4, Additional file 2: Table S11]. Estradiol was also positively associated with abundance of *Faecalibacterium*, especially *Faecalibacterium prausnitzii* [$\widehat{\beta} = 0.82$ log-fold higher per tenfold estradiol increase (0.24, 1.41), $p = 0.007$; Additional file 2: Tables S12, S13]. Self-reported female sexual maturation (menarche, breast development, pubic hair growth) was also associated with differences in the gut microbiome. Menarchal females [$\widehat{\beta} = 0.48$ log-fold higher compared to premenarchal (0.18, 0.78), $p = 0.002$] and those reporting advanced breast development [$\widehat{\beta} = 0.49$ log-fold higher compared to those reporting early development (0.06, 0.91), $p$-for-trend $= 0.027$] had higher *Dorea* abundance (Additional file 2: Table S12), including both *Dorea longicatena* and *Dorea formicigenerans* (Additional file 2: Table S13). Those reporting advanced breast development and pubic hair growth had higher abundances of *Coprococcus*, specifically *Coprococcus comes*, and lower abundances of *Romboutsia*, specifically *Romboutsia timonensis*, than those reporting earlier development. Breast development was associated with higher species-level Shannon ($p$-for-trend $= 0.005$) and inverse Simpson diversity ($p$-for-trend $= 0.025$; Additional file 2: Table S17). Sample size limits inference about associations between sexual maturation and SOM nodes (Additional file 2: Table S10). However, we found that the proportion of females reporting advanced pubic hair growth varied across SOM clusters, with cluster D having the highest proportion reporting advanced growth (58.8) and those in cluster B reporting the lowest proportion (25%, $p = 0.062$).

Among males, self-reported pubic hair growth was associated with higher relative abundance of Actinobacteria ($p$-for-trend $= 0.024$) and somewhat lower relative abundance of Bacteroidota ($p$-for-trend $= 0.067$; Fig. 4, Additional file 2: Table S11). Males reporting advanced pubic hair growth had higher relative abundances of *Bifidobacterium* ($p$-for-trend $= 0.026$) and *Dorea* ($p$-for-trend $= 0.016$), and lower relative abundance of *Lachnospira* compared to those with less advanced growth ($p$-for-trend $= 0.016$; Additional file 2: Table S12). In addition to positive associations with *Dorea* species, we found that males reporting advanced pubic hair growth had lower relative abundance of *Bacteroides uniformis* ($p$-for-trend $= 0.02$) and two *Blautia* species (Additional file 2: Table S13). We did not find differences in diversity or SOM clusters related to sexual maturation in males or associations with testosterone (Additional file 2: Tables S10–S13, S17). Adjustment for age attenuated some associations and strengthened others, but the overall patterns remained the same (Additional file 2: Tables S12–S16).

#### *Concurrent non-dietary factors*

As expected, participants taking antibiotics at the time of stool collection ($n = 4$) had lower Shannon [$\widehat{\beta} = -0.8$

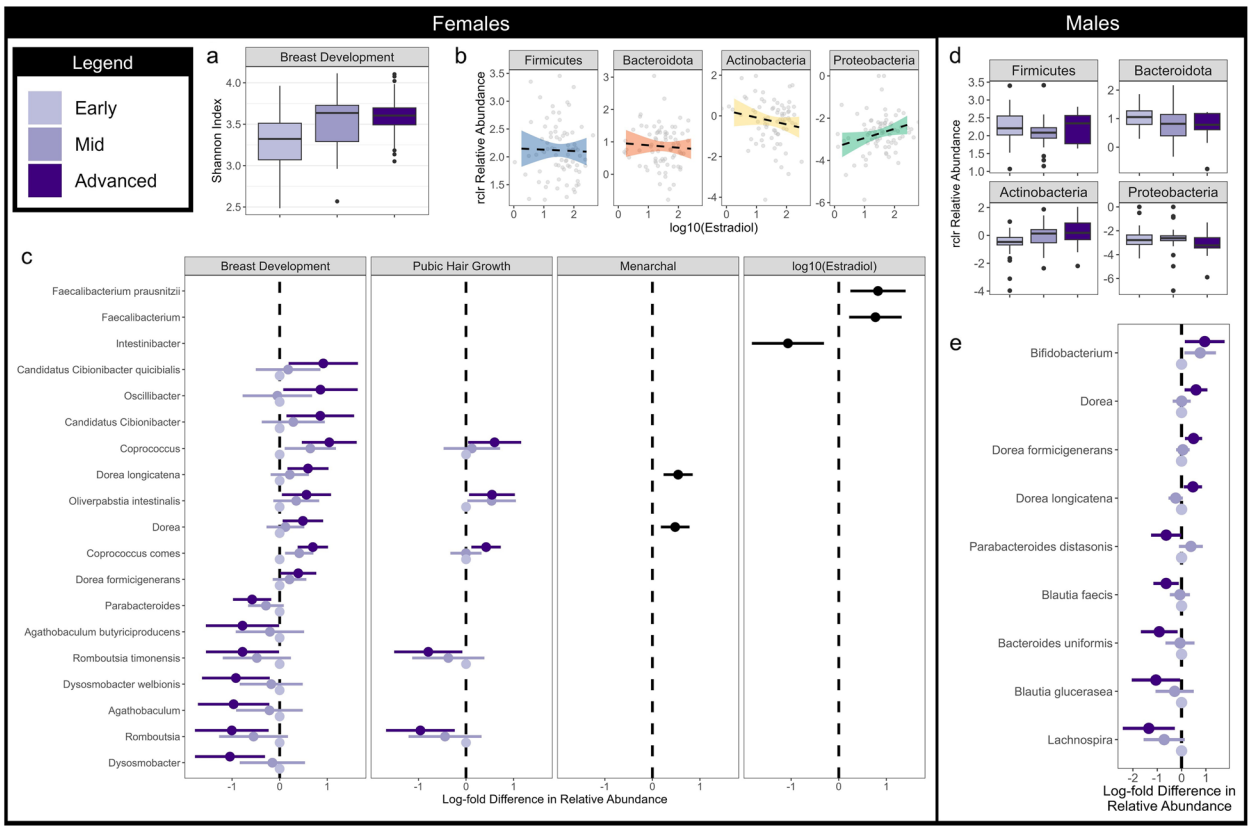


**Fig. 4** Associations between sexual maturation and gut microbiome diversity (Shannon index; **a**), robust centered log ratio (rclr) transformed phyla relative abundances (**b**, **d**), and rclr transformed genus and species relative abundances (**c**, **e**). Estimates in **c** and **e** for breast development and pubic hair growth are compared to early growth, for menarchal are compared to pre-menarchal, and for log10(Estradiol) are per tenfold increase in estradiol. Genus and species estimates with $p < 0.01$ are displayed

(95%CI: −1.1, −0.5), $p < 0.001$] and inverse Simpson diversity [$\widehat{\beta} = -11.3$ (95%CI: −18.9, −3.7), $p = 0.004$] than those not taking antibiotics (Additional file 2: Table S7). Higher household income at the time of the study visit was significantly associated with lower Shannon [$\widehat{\beta} = -0.01$ per $10,000 (95%CI: −0.02, −0.00), $p = 0.023$] and inverse Simpson [$\widehat{\beta} = -0.3$ per $10,000 (95%CI: −0.5, −0.03), $p = 0.028$] diversity, although the effect sizes were small (Fig. 3; Additional file 2: Table S17). Participant BMI *Z*-scores were marginally positively associated with the Shannon index, although this estimate did not reach statistical significance [$\widehat{\beta} = 0.04$ per unit increase in BMI *Z*-score (95%CI: −0.01, 0.1), $p = 0.093$]. Stool consistency (Bristol score) was weakly associated with some differences in diversity, but there were no consistent trends, and once categorized as constipated, healthy, or diarrhetic there were no significant associations (Additional file 2: Table S17). Household income, BMI *Z*-scores, and antibiotic usage were also associated with differential phylum, genus, and species relative abundances (Additional file 2: Tables S11–S13). For example, household income was associated with higher Firmicutes relative abundance [$\widehat{\beta} = 0.02$ log-fold change per $10,000 (95%CI: 0.00, 0.03), $p = 0.011$] and lower Actinobacteria relative abundance [$\widehat{\beta} = -0.04$ log-fold change (95%CI: −0.07, −0.01), $p = 0.004$; Additional file 2: Table S11], as well as lower *Bifidobacterium*, *Dorea*, and *Oscillibacter* (Additional file 2: Table S12). Age at follow-up was not associated with differences in adolescent gut microbiome diversity or composition (Additional file 2: Tables S10–S13, S17).

### *Dietary factors*

Participants had a median (IQR) Healthy Eating Index 2010 (HEI) score of 44 [38, 53], indicating poor adherence to the Dietary Guidelines for Americans (Additional file 1: Fig. S5, Additional file 2: Table S3) [37, 49]. Higher total HEI scores were associated with 0.01 (0.0, 0.01) higher Shannon index ($p = 0.028$; Additional file 2: Table S18), as well as lower Actinobacteria relative abundance (Additional file 2: Table S19), higher *Lachnospira* abundance (Additional file 2: Table S12), and lower Oscillospiraceae bacterium CLA AA H250

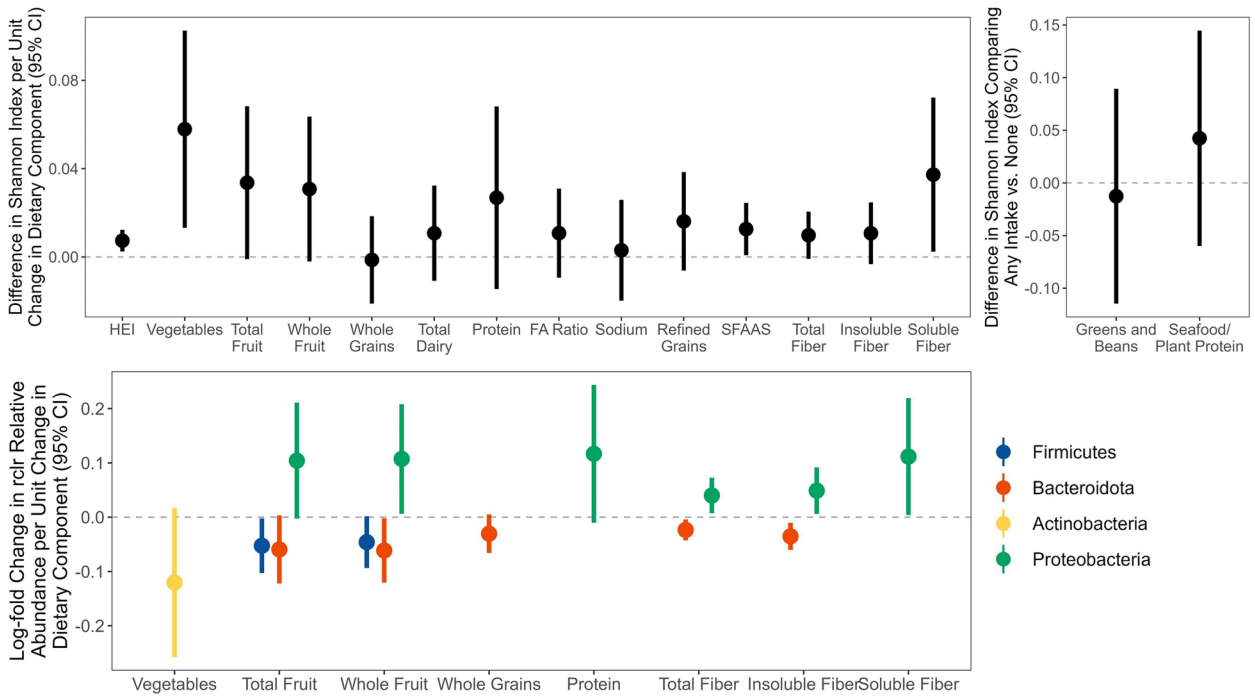


**Fig. 5** Associations between dietary components and gut microbiome diversity (Shannon index) and robust centered log ratio (rclr) transformed phyla relative abundances. Estimates are adjusted for household income. HEI, Health Eating Index; FA Ratio, fatty acid ratio; SFAAS, solid fats and added sugars. Diet-phyla associations with $p < 0.1$ are displayed

(recently named *Hominenteromicrobium mulieris*; Additional file 2: Table S13) [50]. We hypothesized that household income confounded the association between dietary components and microbiome diversity, and thus conducted an adjusted analysis (Fig. 5). Most of these estimates were consistent even after adjusting for income, except for the association with Actinobacteria (Additional file 2: Tables S20, S21). As expected, most individual dietary components were positively correlated with measures of diversity, but only total vegetable intake was significantly associated (Additional file 2: Table S18). After adjusting for household income, these associations were strengthened such that consumption of solid fats and added sugars and fiber variables were also significantly associated with diversity measures (Additional file 2: Table S19). We did not observe differences in SOM cluster assignment related to dietary intake (Additional file 2: Tables S22, S23). Bacteroidota relative abundance was negatively associated with consumption of total and insoluble fiber, especially after adjustment for household income (Additional file 2: Tables S19, S21). In accordance, consumption of fruit and whole grains, sources of dietary fiber, was also negative correlated with Bacteroidota relative abundance. All fiber variables were also positively associated with Proteobacteria relative abundance, as was fruit consumption. Firmicutes relative abundance was positively associated with consumption of total vegetables (unadjusted), total fruit, and whole fruit (adjusted). The only dietary variable associated with Actinobacteria relative abundance was vegetable intake, for which we observed a negative association. We observed some associations between HEI components and relative abundances of genera and species. Most notably, the relative abundance of the species associated with total HEI scores, *Oscillospiraceae bacterium CLA AA H250*, was lower among those with higher consumption of total and insoluble fiber and lower consumption of solid fats and added sugars (Additional file 2: Tables S12, S13).

Our sensitivity analysis excluding participants currently taking antibiotics did not meaningfully alter the associations of early-life and concurrent factors with fecal microbiome diversity and composition (Additional file 2: Tables S12, S13, S24–S35). A notable exception was *Blautia wexlerae*, which was positively associated with sexual maturation in males and females only after excluding participants currently taking antibiotics (Additional file 2: Table S13).

## Discussion

In this study of adolescents' fecal microbiomes, we described prevalent and abundant bacterial species in addition to highlighting predictors of microbial diversity, composition, and variability. On average, adolescent samples had higher relative abundance of Firmicutes than other bacterial phyla, with smaller proportions of Bacteroidetes, Actinobacteria, Proteobacteria, and other phyla. The most prevalent and abundant species and genera were commonly detected human-gut-affiliated microbes, including *Faecalibacterium prausnitzii* [51]. The most consistent predictors of bacterial diversity and composition were household income and self-reported race/ethnicity, suggesting these sociodemographic characteristics may influence the adolescent microbiome. Our study addresses the paucity of studies in the critical developmental window of adolescence, which may be critical for lifelong health.

Prior studies of the adolescent microbiome have been small ($n < 100$) or used arrays or 16S rRNA sequencing. These studies reported that adolescents have less *Coprobacillus* and more *Burkholderiales* than younger children [16], and more *Clostridium* and *Bifidobacterium* than adults [13]. There is also a growing literature on the differences between the late childhood (9–12 years of age) microbiome relative to adults [14, 52]. These studies report that *Bacteroides*, *Faecalibacterium*, and *Bifidobacterium*—all abundant genera in our study—were more abundant in children than adults. Unexpectedly, we found that *Bifidobacterium* relative abundance was higher among males reporting more advanced pubic hair growth, although this was not reflected at the species level. *Blautia*—which was highly abundant in our sample—has previously been found to be more abundant in adults than children [14, 52]. In accordance, we found that *Blautia wexlerae* was associated with more advanced sexual maturation in males and females, but only after excluding individuals taking antibiotics. The genus most consistently positively associated with maturation in the prior literature, *Dorea*, was also associated with more advanced sexual maturation in both males and females in our analyses [14–16]. Together, these studies suggest that some genera or species may be bellwethers of gut maturation.

In line with the hypothesis that the adolescent microbiome is distinct from both pediatric and adult communities, we observed differences in microbial diversity and composition related to sexual maturation and hormone concentrations. Gut bacterial composition may change with sexual maturation in response to a shifting hormonal milieu, as other studies have observed differences in the microbiomes of post-menopausal individuals [10, 53, 54] and those taking hormonal oral contraceptives [55, 56]. Further, ovariectomized or gonadectomized mice experience shifts in their microbiomes that can be reversed by supplementation with the estradiol and dihydrotestosterone, respectively [57, 58]. Certain bacterial enzymes can metabolize unconjugated steroid hormones or facilitate enterohepatic cycling of conjugated steroid hormones, suggesting there may be bidirectional feedback loops between gut bacteria and circulating hormone concentrations [59]. Our findings suggest that puberty may be a key window of microbiome development that should be investigated in relation to chronic health outcomes.

Our findings suggest that household income and race/ethnicity are important factors to consider in future studies of the microbiome, including as a covariate in studies of environmental exposures or health outcomes. Income is likely a proxy for other factors such as household size [60] and housing conditions [61], including residential mold [62] and lead paint or pipe exposure [63, 64], proximity to major roadways and affiliated air pollution [65, 66], and other environmental toxicants that have been linked to the microbiome [67]. Income may also be related to personal [68] and oral [69] hygiene. Income is also related to diet quality, food choices/availability, and lifestyle, which in turn are factors that may be more strongly related to the microbiome than the HEI components assessed in this study [70–75].

Race, which is a social construct, was also associated with the microbiome in our study. While race may serve as a proxy for genetic ancestry, which has a small but significant contribution to microbiome composition [76–78], it also reflects dietary difference between populations [70, 75, 79, 80], either due to cultural differences or systemic racism limiting the availability of nutritious food options in neighborhoods with majority minoritized populations (i.e., food deserts) [81, 82]. Other facets of systemic racism, such as psychological stress [83] and poor environmental quality linked to historic redlining [84–86], may drive the race-microbiome association in our study [87]. Larger epidemiologic studies to elucidate the nuances of this association are needed to understand how social factors influence the microbiome and downstream health consequences.

We hypothesized that diet would be a significant driver of gut bacterial diversity and composition, but did not find strong associations with HEI components. Most HOME Study participants reported poor adherence to US dietary guidelines [49], which is not unusual for studies of the general US population [88–90], particularly adolescents [89]. Consistent with prior studies of diet and the microbiome, we observed a small positive association between total HEI scores and bacterial diversity [88, 91]. Associations of individual food components with

bacterial diversity and composition were weak. Operationalization of diet as HEI components may not capture features that are most relevant to microbiome diversity and composition, although they have previously been reported to be predictors of the microbiome [91]. For example, we did not observe associations between total dairy consumption and most components of the microbiome. Fermented dairy consumption may be a stronger predictor of microbiome composition. Similarly, while there is some evidence to support our findings of associations between fiber and microbiome composition [92], individual dietary fibers have highly specific associations with individual bacterial species [93]; thus, our use of total fiber may have obscured some findings. A more comprehensive investigation of longer-term and more detailed dietary patterns with the adolescent microbiome is warranted.

Based on prior associations of birth mode [18, 94] and human milk consumption [95] with childhood or adult gut microbiome composition, we hypothesized that we would observe associations in adolescence. We found no evidence, however, of an association of birth mode with adolescent microbiome diversity or composition. Most studies of the adult microbiome collect data on birth mode and early diet retrospectively, which could be biased. It is also possible that our study was underpowered to detect small differences in the microbiome according to birth mode. Prior studies have primarily highlighted differences in individual species or genera, rather than compositional profiles or diversity measures [95, 96]. We found some evidence that consumption of human milk was linked to differences in the adolescent gut microbiome, although there was evidence some of these findings were confounded by socioeconomic status. The bacteria associated with early-life consumption of exclusively human milk in our study were different from those associated with longer duration of any human milk consumption. This may reflect true differences in the impact of early-life dietary patterns on the adolescent microbiome, or may reflect sociodemographic differences in ability to continue human milk feeding beyond 6 months. Further, participants who gave birth at one hospital in the cohort catchment area were offered formula immediately; thus, these participants did not meet the definition of being exclusively fed human milk early in life, even though the majority of their calories may have come from human milk. The human milk microbiome is dominated by Firmicutes, Actinobacteria, and Proteobacteria [97–99]. Additional longitudinal studies examining the transfer of milk microbes to offspring could elucidate our findings further.

Similar to prior studies, we found that most of the variability in fecal bacterial diversity and composition was explained by intra-individual variation, with a smaller proportion of variance explained by days between sample collection and stool consistency [100–102]. Even though we selected the subset of participants included in this analysis to have maximum within-subject variability based on stool consistency, this was dwarfed by between-subject variability. Thus, we are reassured that a single sample is sufficient to estimate gut microbial composition over a short period for most epidemiologic studies, although longitudinal studies may provide additional insight into health outcomes. Approximately 12.5% of the variability in microbiome composition remains unexplained by the examined factors. This variability may be attributable to environmental exposures or health status, but future studies are required to elucidate important, modifiable components.

This study had some limitations. The use of fecal occult blood test (FOBT) cards has been validated for metagenomic sequencing, but some differences exist in bacterial relative abundances compared with gold standard methods (e.g., immediately frozen whole stool) [103]. Likewise, sequence processing pipelines like MetaPhlAn can be biased in their alignment to known sequences. We used up-to-date libraries of known sequences, which include species-level genome bins (SGBs) that have not yet been aligned to known taxa, to reduce this bias [28]. Our assessment of factors that may predict microbiome diversity and composition, while broad, was not all-encompassing, and we may not have collected or analyzed all variables that are most predictive of the microbiome. For example, the use of three 24-h dietary recalls, while effective at capturing recent nutrient intake, may not reflect longer-term dietary patterns that may be more relevant to microbiome diversity and composition. Most of our models were unadjusted for other predictors; thus, our results should not be interpreted as causal. Lastly, while we assessed predictors of prevalent phyla, genera, and species, less prevalent features may be uniquely linked to early-life and concurrent factors.

This is one of the first investigations into predictors of the adolescent microbiome and sources of its variability. Our detailed and comprehensive collection of early-life and concurrent features that have been associated with the microbiome coupled with 12 years of prospective follow-up in a moderately large sample size allowed us to conduct a detailed statistical analysis of factors influencing the gut microbiome. In addition to the value of information garnered here, this investigation will be useful for identifying confounders in future analyses of the adolescent gut microbiome and exposures or health outcomes. We employed cutting-edge machine-learning methodologies such as SOMs to identify common microbial profiles. Previously, we have used this approach to

characterize the fecal metabolome [104], but this is the first time that SOMs have been applied to human microbial taxonomies. This method enhances the characterization and visualization of complex data and improves reproducibility across studies because the map from one study can be projected onto data from another [44, 45]. Overall, these strengths contribute to the rigor and reliability of the study findings, advancing our understanding of the adolescent gut microbiome and its determinants.

## Conclusions

In our study of the adolescent gut microbiome, we found that sociodemographic variables were the most consistently associated with bacterial diversity and composition. Our findings shed light on the intricate interplay between early-life exposures, sexual maturation, dietary habits, and socioeconomic factors in shaping the adolescent gut microbiome. Understanding these associations with predictors is crucial for identifying potential confounders in future epidemiologic studies of the adolescent microbiome and health outcomes. This study adds to the growing research on the microbiome and provides some insights into how the adolescent microbiome is a unique biomarker that may be relevant to lifelong health.

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13073-025-01481-1.

Additional file 1. Supplementary figures (Figures S1–S5).

Additional file 2. Supplementary tables (Tables S1–S35).

**Acknowledgements**

The authors would like to thank participants in the Health Outcomes and Measures of the Environment (HOME) Study and their families, as well as cohort staff.

**Authors' contributions**

HEL and JMB designed the work. HEL, JMB, KMC, KY, BPL, AC, MRK, JPB, AFF, and JCM acquired funding. HEL, FW, YX, MCM, and ADW acquired, analyzed, and interpreted the data. HEL drafted the original manuscript. All authors read, revised, and approved the final manuscript.

**Funding**

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The results reported herein correspond to specific aims of grant K99/R00 ES034086 to HEL from the National Institute of Environmental Health Sciences (NIEHS). This work was also supported by NIEHS grants P01 ES011261, R01 ES0272244, R01 ES025214, the National Center for Advancing Translational Sciences grant UL1TR001425, and the National Institute of General Medical Sciences grant R35 GM133420.The New York University Langone Genome Technology Center is supported by the National Cancer Institute through P30CA016087. No funders had a role in conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

**Data availability**

Raw sequences were uploaded to the National Center for Biotechnology Information Sequence Read Archive (Accession: PRJNA1139690) after the removal of human reads, which can be found at https://dataview.ncbi.nlm.nih.gov/object/PRJNA1139690 [24]. Health Outcomes and Measures of the Environment (HOME) Study participants did not consent to sharing metadata in a public repository. Requests to access relevant metadata should be made through the HOME Study Portal https://homestudy.research.cchmc.org/contact, where the cohort data sharing plan can be found. Data will be granted upon reasonable request after approval of the HOME Study Research and Publications Committee. Researchers seeking additional information about working with HOME Study data should contact HOMEStudyDataRequest@cchmc.org. Code is available at github.com/HEL548/HOME_Microbiome [105].

## Declarations

**Ethics approval and consent to participate**

Caregivers provided written, informed consent and adolescents provided written, informed assent. The institutional review boards (IRBs) of Cincinnati Children's Hospital Medical Center (CCHMC) and participating delivery hospitals approved this study (IRB IDs: 2020–0941, 2024–0560). Dartmouth College, Brown University, and the University of Massachusetts Amherst deferred to the CCHMC IRB as the IRB of record. All research protocols conformed to the principles of the Helsinki Declaration.

**Consent for publication**

Not applicable.

**Competing interests**

JMB has served as an expert witness for plaintiffs in litigation related to PFAS-contaminated drinking water. The remaining authors declare that they have no competing interests.

**Author details**

[1]Department of Biostatistics and Epidemiology, University of Massachusetts Amherst School of Public Health and Health Sciences, 715 N. Pleasant Street, Arnold House 429, Amherst, MA 01003, USA. [2]Department of Biostatistics, University of Washington Hans Rosling Center for Population Health, 3980 15 Avenue NE, Box 351617, Seattle, WA 98195-1617, USA. [3]Department of Environmental Health Sciences, Columbia Mailman School of Public Health, 630 W 168th St, P&S 16-416, New York, NY 10032, USA. [4]Division of General and Community Pediatrics, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH 45229, USA. [5]Department of Pediatrics, University of Cincinnati College of Medicine, 3333 Burnet Ave, Cincinnati, OH 45229, USA. [6]Department of Epidemiology, Dartmouth Geisel School of Medicine, One Medical Center Dr Lebanon, Lebanon, NH 03756, USA. [7]Department of Psychiatry, Dartmouth Hitchcock Medical Center, Lebanon, NH, USA. [8]Center for Interdisciplinary and Population Health Research, Maine Institute for Research, Westbrook, ME, USA. [9]Pediatric Endocrinology and Diabetes, Maine Medical Center, 887 Congress St, Portland, ME, USA. [10]Faculty of Health Sciences, Simon Fraser University, Blusson Hall, 8888 University Dr, Burnaby, BC, Canada. [11]Department of Radiology, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, 3333 Burnet Ave, Cincinnati, OH 45229, USA. [12]Department of Environmental and Public Health Sciences, University of Cincinnati, 3333 Burnet Ave, Cincinnati, OH 45229, USA. [13]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania Perelman School of Medicine, 423 Guardian Drive, Philadelphia, PA 19104, USA. [14]Department of Epidemiology, University of North Carolina at Chapel Hill, 2106-B McGavran-Greenberg Hall CB#7435, Chapel Hill, NC 27599, USA. [15]Department of Epidemiology, Brown University, 121 S Main St, Providence, RI, USA.

Received: 9 October 2024 Accepted: 25 April 2025
Published online: 08 May 2025

**References**

1. Young VB. The role of the microbiome in human health and disease: an introduction for clinicians. BMJ. 2017Mar;15(356): j831.
2. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The Human Microbiome Project. Nature. 2007Oct;449(7164):804–10.

3. Lloyd-Price J, Abu-Ali G, Huttenhower C. The healthy human microbiome. Genome Med. 2016Apr;27(8):51.
4. Yassour M, Vatanen T, Siljander H, Hämäläinen AM, Härkönen T, Ryhänen SJ, et al. Natural history of the infant gut microbiome and impact of antibiotic treatments on strain-level diversity and stability. Sci Transl Med. 2016;8(343):343ra81.
5. Bokulich NA, Chung J, Battaglia T, Henderson N, Jay M, Li H, et al. Antibiotics, birth mode, and diet shape microbiome maturation during early life. Sci Translat Med. 2016;8(343):343ra82–343ra82.
6. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012Jun;486(7402):222–7.
7. Koren O, Goodrich JK, Cullender TC, Spor A, Laitinen K, Kling Bäckhed H, et al. Host remodeling of the gut microbiome and metabolic changes during pregnancy. Cell. 2012Aug 3;150(3):470–80.
8. Edwards SM, Cunningham SA, Dunlop AL, Corwin EJ. The maternal gut microbiome during pregnancy. MCN Am J Matern Child Nurs. 2017;42(6):310–7.
9. Smid MC, Ricks NM, Panzer A, McCoy AN, Azcarate-Peril MA, Keku TO, et al. Maternal gut microbiome biodiversity in pregnancy. Am J Perinatol. 2018Jan;35(1):24–30.
10. Santos-Marcos JA, Rangel-Zuñiga OA, Jimenez-Lucena R, Quintana-Navarro GM, Garcia-Carpintero S, Malagon MM, et al. Influence of gender and menopausal status on gut microbiota. Maturitas. 2018Oct;1(116):43–53.
11. Gusberti FA, Mombelli A, Lang NP, Minder CE. Changes in subgingival microbiota during puberty. A 4-year longitudinal study. J Clin Periodontol. 1990;17(10):685–92.
12. Oh J, Conlan S, Polley EC, Segre JA, Kong HH. Shifts in human skin and nares microbiota of healthy children and adults. Genome Med. 2012;4(10):77.
13. Agans R, Rigsbee L, Kenche H, Michail S, Khamis HJ, Paliy O. Distal gut microbiota of adolescent children is different from that of adults. FEMS Microbiol Ecol. 2011Aug;77(2):404–12.
14. Radjabzadeh D, Boer CG, Beth SA, van der Wal P, Kiefte-De Jong JC, Jansen MAE, et al. Diversity, compositional and functional differences between gut microbiota of children and adults. Sci Rep. 2020Jan 23;10(1):1040.
15. Korpela K, Kallio S, Salonen A, Hero M, Kukkonen AK, Miettinen PJ, et al. Gut microbiota develop towards an adult profile in a sex-specific manner during puberty. Sci Rep. 2021Dec;2(11):23297.
16. Yuan X, Chen R, Zhang Y, Lin X, Yang X. Gut microbiota: effect of pubertal status. BMC Microbiol. 2020Nov;3(20):334.
17. Ding T, Schloss PD. Dynamics and associations of microbial community types across the human body. Nature. 2014May 15;509(7500):357–60.
18. Goedert JJ, Hua X, Yu G, Shi J. Diversity and composition of the adult fecal microbiome associated with history of cesarean birth or appendectomy: analysis of the American Gut Project. EBioMedicine. 2014Nov 8;1(2–3):167–72.
19. Braun JM, Kalloo G, Chen A, Dietrich KN, Liddy-Hicks S, Morgan S, et al. Cohort profile: the Health Outcomes and Measures of the Environment (HOME) study. Int J Epidemiol. 2016;46(1):24
20. Braun JM, Buckley JP, Cecil KM, Chen A, Kalkwarf HJ, Lanphear BP, et al. Adolescent follow-up in the Health Outcomes and Measures of the Environment (HOME) Study: cohort profile. BMJ Open. 2020 May 7;10(5). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7228515/. Cited 2021 May 15
21. 16 CFR part 1303 -- ban of lead-containing paint and certain consumer products bearing lead-containing paint. Available from: https://www.ecfr.gov/current/title-16/part-1303. Cited 2025 Apr 20
22. Thompson WG, Heaton KW. Fast facts: irritable bowel syndrome. S.Karger AG; 2003. Available from: https://karger.com/books/book/19/Fast-Facts-Irritable-Bowel-Syndrome. Cited 2024 Jan 16
23. Shen Y, Laue HE, Shrubsole MJ, Wu H, Bloomquist TR, Larouche A, et al. Associations of childhood and perinatal blood metals with children's gut microbiomes in a Canadian gestation cohort. Environ Health Perspect. 2022Jan 17;130(1): 017007.
24. Laue HE, Willis AD, Wang F, MacDougall MC, Xu Y, Karagas MR, et al. Health Outcomes and Measures of the Environmental (HOME) Study adolescent microbiome data. National Center for Biotechnology Information Sequence Read Archive; Available from: https://dataview.ncbi.nlm.nih.gov/object/PRJNA1139690
25. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife. 2021May;4:10.
26. McIver LJ, Abu-Ali G, Franzosa EA, Schwager R, Morgan XC, Waldron L, et al. bioBakery: a meta'omic analysis environment. Bioinformatics. 2018Apr 1;34(7):1235–7.
27. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014Aug 1;30(15):2114.
28. Blanco-Míguez A, Beghini F, Cumbo F, McIver LJ, Thompson KN, Zolfo M, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat Biotechnol. 2023Nov;41(11):1633–44.
29. Romano ME, Xu Y, Calafat AM, Yolton K, Chen A, Webster GM, et al. Maternal serum perfluoroalkyl substances during pregnancy and duration of breastfeeding. Environ Res. 2016Aug;149:239–46.
30. Campbell OMR, Cegolon L, Macleod D, Benova L. Length of stay after childbirth in 92 countries and associated factors in 30 low- and middle-income countries: compilation of reported data and a cross-sectional analysis from nationally representative surveys. PLoS Med. 2016Mar 8;13(3): e1001972.
31. Braun JM, Daniels JL, Poole C, Olshan AF, Hornung R, Bernert JT, et al. A prospective cohort study of biomarkers of prenatal tobacco smoke exposure: the correlation between serum and meconium and their association with infant birth weight. Environ Health. 2010Aug;27(9):53.
32. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, Flegal KM, Guo SS, Wei R, et al. CDC growth charts: United States. Adv Data. 2000Jun 8;314:1–27.
33. Tanner JM. Growth at adolescence: with a general consideration of the effects of hereditary and environmental factors upon growth and maturation from birth to maturity. Blackwell Scientific Publications; 1962. Available from: https://books.google.com/books?id=h8sJAQAAMAAJ
34. Yayah Jones NH, Khoury JC, Xu Y, Newman N, Kalkwarf HJ, Braun JM, et al. Comparing adolescent self staging of pubertal development with hormone biomarkers. J Pediatr Endocrinol Metab. 2021Dec 20;34(12):1531–41.
35. Laue HE, Lanphear BP, Calafat AM, Cecil KM, Chen A, Xu Y, et al. Time-varying associations of gestational and childhood triclosan with pubertal and adrenarchal outcomes in early adolescence. Environ Epidemiol. 2024Apr;8(2): e305.
36. Chumpitazi BP, Self MM, Czyzewski DI, Cejka S, Swank PR, Shulman RJ. Bristol stool form scale reliability and agreement decreases when determining Rome III stool form designations. Neurogastroenterol Motil. 2016Mar;28(3):443–8.
37. Guenther PM, Kirkpatrick SI, Reedy J, Krebs-Smith SM, Buckman DW, Dodd KW, et al. The Healthy Eating Index-2010 is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans. J Nutr. 2014Mar;144(3):399–407.
38. version) JP (S, to 2007) DB (up, to 2002) SD (up, to 2005) DS (up, authors (src/rs.f) E, sigma) SH (Author fixed, et al. nlme: linear and nonlinear mixed effects models. 2023. Available from: https://cran.r-project.org/web/packages/nlme/index.html. Cited 2024 May 31
39. Mixed-effects models in S and S-PLUS. New York: Springer-Verlag; 2000. (Statistics and Computing). Available from: http://link.springer.com/https://doi.org/10.1007/b98882. Cited 2024 May 31
40. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. vegan: community ecology package. 2022. Available from: https://CRAN.R-project.org/package=vegan
41. Bray JR, Curtis JT. An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr. 1957;27(4):325–49.
42. Shannon C, Weaver W. The mathematical theory of communication. Urbana: The University of Illinois Press; 1964.
43. Simpson EH. Measurement of diversity. Nature. 1949Apr;163(4148):688.
44. Pearce JL, Waller LA, Chang HH, Klein M, Mulholland JA, Sarnat JA, et al. Using self-organizing maps to develop ambient air quality classifications: a time series example. Environ Health. 2014Jul;3(13):56.
45. Pearce JL, Waller LA, Sarnat SE, Chang HH, Klein M, Mulholland JA, et al. Characterizing the spatial distribution of multiple pollutants and

populations at risk in Atlanta. Georgia Spat Spatiotemporal Epidemiol. 2016Aug;18:13–23.
46. Kohonen T. Essentials of the self-organizing map. Neural Netw. 2013Jan;1(37):52–65.
47. Martino C, Morton JT, Marotz CA, Thompson LR, Tripathi A, Knight R, et al. A novel sparse compositional technique reveals microbial perturbations. mSystems. 2019;4(1):e00016–19.
48. Doherty BT, Pearce JL, Anderson KA, Karagas MR, Romano ME. Assessment of multipollutant exposures during pregnancy using silicone wristbands. Front Public Health. 2020Sep;29(8): 547239.
49. Krebs-Smith SM, Pannucci TE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, et al. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet. 2018Sep;118(9):1591–602.
50. Afrizal A, Hitch TCA, Viehof A, Treichel N, Riedel T, Abt B, et al. Anaerobic single-cell dispensing facilitates the cultivation of human gut bacteria. Environ Microbiol. 2022;24(9):3861–81.
51. Lopez-Siles M, Duncan SH, Garcia-Gil LJ, Martinez-Medina M. Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. ISME J. 2017Apr;11(4):841–52.
52. Hollister EB, Riehle K, Luna RA, Weidler EM, Rubio-Gonzales M, Mistretta TA, et al. Structure and function of the healthy pre-adolescent pediatric gut microbiome. Microbiome. 2015Aug 26;3(1):36.
53. Flores R, Shi J, Fuhrman B, Xu X, Veenstra TD, Gail MH, et al. Fecal microbial determinants of fecal and systemic estrogens and estrogen metabolites: a cross-sectional study. J Transl Med. 2012Dec;21(10):253.
54. Fuhrman BJ, Feigelson HS, Flores R, Gail MH, Xu X, Ravel J, et al. Associations of the fecal microbiome with urinary estrogens and estrogen metabolites in postmenopausal women. J Clin Endocrinol Metab. 2014Dec;99(12):4632–40.
55. Hua X, Cao Y, Morgan DM, Miller K, Chin SM, Bellavance D, et al. Longitudinal analysis of the impact of oral contraceptive use on the gut microbiome. J Med Microbiol. 2022;71(4):001512.
56. Mihajlovic J, Leutner M, Hausmann B, Kohl G, Schwarz J, Röver H, et al. Combined hormonal contraceptives are associated with minor changes in composition and diversity in gut microbiota of healthy women. Environ Microbiol. 2021;23(6):3037–47.
57. Gao A, Su J, Liu R, Zhao S, Li W, Xu X, et al. Sexual dimorphism in glucose metabolism is shaped by androgen-driven gut microbiome. Nat Commun. 2021Dec 6;12(1):7080.
58. Kaliannan K, Robertson RC, Murphy K, Stanton C, Kang C, Wang B, et al. Estrogen-mediated gut microbiome alterations influence sexual dimorphism in metabolic syndrome in mice. Microbiome. 2018Nov 13;6(1):205.
59. Sisk-Hackworth L, Kelley ST, Thackray VG. Sex, puberty, and the gut microbiome. 2023 Feb 1; Available from: https://rep.bioscientifica.com/view/journals/rep/165/2/REP-22-0303.xml. Cited 2024 Jul 25
60. B19019: median household income in ... - Census Bureau table. Available from: https://data.census.gov/table/ACSDT1Y2023.B19019?q=median+income+household+size&g=040XX00US21,39. Cited 2025 Mar 16
61. Rolfe S, Garnham L, Godwin J, Anderson I, Seaman P, Donaldson C. Housing as a social determinant of health and wellbeing: developing an empirically-informed realist theoretical framework. BMC Public Health. 2020Jul 20;20(1):1138.
62. Panthagani KM, Hoffman KL, Oluyomi A, Sotelo J, Stewart C, Armstrong G, et al. Butyrate-producing gut bacteria are associated with protection from allergic symptoms after Hurricane Harvey. medRxiv; 2021 [cited 2024 May 28]. p. 2021.12.09.21267553. Available from: https://www.medrxiv.org/content/https://doi.org/10.1101/2021.12.09.21267553v1
63. Gao B, Chi L, Mahbub R, Bian X, Tu P, Ru H, et al. Multi-omics reveals that lead exposure disturbs gut microbiome development, key metabolites, and metabolic pathways. Chem Res Toxicol. 2017Apr 17;30(4):996–1005.
64. Eggers S, Midya V, Bixby M, Gennings C, Torres-Olascoaga LA, Walker RW, et al. Prenatal lead exposure is negatively associated with the gut microbiome in childhood. Front Microbiol. 2023 Jun 22 [cited 2024 May 28];14. Available from: https://www.frontiersin.org/journals/microbiology/articles/https://doi.org/10.3389/fmicb.2023.1193919/full
65. Fouladi F, Bailey MJ, Patterson WB, Sioda M, Blakley IC, Fodor AA, et al. Air pollution exposure is associated with the gut microbiome as revealed by shotgun metagenomic sequencing. Environ Int. 2020May;1(138): 105604.
66. Salim SY, Kaplan GG, Madsen KL. Air pollution effects on the gut microbiota: a link between exposure and inflammatory disease. Gut Microbes. 2014Mar 1;5(2):215–9.
67. Campana AM, Laue HE, Shen Y, Shrubsole MJ, Baccarelli AA. Assessing the role of the gut microbiome at the interface between environmental chemical exposures and human health: current knowledge and challenges. Environ Pollut. 2022Dec;15(315): 120380.
68. Esschert KV, Barrett CE, Collier SA, Garcia-Williams AG, Hannapel E, Yoder JS, et al. Demographic differences in use of household tap water in a representative sample of US adults, FallStyles 2019. J Water Health. 2021Dec;19(6):1014–20.
69. Polk DE, Weyant RJ, Manz MC. Socioeconomic factors in adolescents' oral health: are they mediated by oral hygiene behaviors or preventive interventions? Commun Dent Oral Epidemiol. 2010;38(1):1–9.
70. Desbouys L, Méjean C, De Henauw S, Castetbon K. Socio-economic and cultural disparities in diet among adolescents and young adults: a systematic review. Public Health Nutr. 23(5):843–60.
71. Drouillet-Pinard P, Dubuisson C, Bordes I, Margaritis I, Lioret S, Volatier JL. Socio-economic disparities in the diet of French children and adolescents: a multidimensional issue. Public Health Nutr. 2017Apr;20(5):870–82.
72. Finger JD, Varnaccia G, Tylleskär T, Lampert T, Mensink GBM. Dietary behaviour and parental socioeconomic position among adolescents: the German Health Interview and Examination Survey for Children and Adolescents 2003–2006 (KiGGS). BMC Public Health. 2015May;19(15):498.
73. Grosso G, Marventano S, Buscemi S, Scuderi A, Matalone M, Platania A, et al. Factors associated with adherence to the Mediterranean diet among adolescents living in Sicily, southern Italy. Nutrients. 2013Dec 4;5(12):4908–23.
74. Yannakoulia M, Lykou A, Kastorini CM, Saranti Papasaranti E, Petralias A, Veloudaki A, et al. Socio-economic and lifestyle parameters associated with diet quality of children and adolescents using classification and regression tree analysis: the DIATROFI study. Public Health Nutr. 2016Feb;19(2):339–47.
75. Drewnowski A, Rehm CD. Socioeconomic gradient in consumption of whole fruit and 100% fruit juice among US children and adults. Nutr J. 2015Jan;5(14):3.
76. Boulund U, Bastos DM, Ferwerda B, van den Born BJ, Pinto-Sietsma SJ, Galenkamp H, et al. Gut microbiome associations with host genotype vary across ethnicities and potentially influence cardiometabolic traits. Cell Host Microbe. 2022Oct 12;30(10):1464-1480.e6.
77. 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. 2018Oct;24(10):1526–31.
78. Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. Cell. 2019Mar 21;177(1):26–31.
79. Northstone K, Smith ADAC, Newby PK, Emmett PM. Longitudinal comparisons of dietary patterns derived by cluster analysis in 7- to 13-year-old children. Br J Nutr. 2013Jun;109(11):2050–8.
80. Northstone K, Smith AD, Cribb VL, Emmett PM. Dietary patterns in UK adolescents obtained from a dual-source FFQ and their associations with socio-economic position, nutrient intake and modes of eating. Public Health Nutr. 2014Jul;17(7):1476–85.
81. Bower KM, Thorpe RJ, Rohde C, Gaskin DJ. The intersection of neighborhood racial segregation, poverty, and urbanicity and its impact on food store availability in the United States. Prev Med. 2014Jan;58:33–9.
82. Cho CY, Clark JK. Disparities in access to supplemental nutrition assistance program retailers over time and space. Popul Res Policy Rev. 2020Feb 1;39(1):99–118.
83. Paradies Y, Ben J, Denson N, Elias A, Priest N, Pieterse A, et al. Racism as a determinant of health: a systematic review and meta-analysis. PLoS ONE. 2015Sep 23;10(9): e0138511.
84. Estien CO, Wilkinson CE, Morello-Frosch R, Schell CJ. Historical redlining is associated with disparities in environmental quality across California. Environ Sci Technol Lett. 2024Feb 13;11(2):54–9.

85. Mehdipanah R, McVay KR, Schulz AJ. Historic redlining practices and contemporary determinants of health in the Detroit metropolitan area. Am J Public Health. 2023Jan;113(S1):S49-57.
86. Li M, Yuan F. Historical redlining and food environments: a study of 102 urban areas in the United States. Health Place. 2022May;1(75): 102775.
87. Mallott EK, Sitarik AR, Leve LD, Cioffi C, Camargo CA, Hasegawa K, et al. Human microbiome variation associated with race and ethnicity emerges as early as 3 months of age. PLoS Biol. 2023Aug 17;21(8): e3002230.
88. Baldeon AD, McDonald D, Gonzalez A, Knight R, Holscher HD. Diet Quality and the fecal microbiota in adults in the American Gut Project. J Nutr. 2023Jul 1;153(7):2004–15.
89. Shams-White MM, Pannucci TE, Lerman JL, Herrick KA, Zimmer M, Mathieu KM, et al. Healthy Eating Index-2020: review and update process to reflect the Dietary Guidelines for Americans, 2020–2025. J Acad Nutr Diet. 2023Sep 1;123(9):1280–8.
90. Reedy J, Lerman JL, Krebs-Smith SM, Kirkpatrick SI, Pannucci TE, Wilson MM, et al. Evaluation of the Healthy Eating Index-2015. J Acad Nutr Diet. 2018Sep;118(9):1622–33.
91. Bowyer RCE, Jackson MA, Pallister T, Skinner J, Spector TD, Welch AA, et al. Use of dietary indices to control for diet in human gut microbiota studies. Microbiome. 2018Apr 25;6(1):77.
92. Cronin P, Joyce SA, O'Toole PW, O'Connor EM. Dietary fibre modulates the gut microbiota. Nutrients. 2021May 13;13(5):1655.
93. Delzenne NM, Bindels LB, Neyrinck AM, Walter J. The gut microbiome and dietary fibres: implications in obesity, cardiometabolic diseases and cancer. Nat Rev Microbiol. 2025Apr;23(4):225–38.
94. Salminen S, Gibson GR, McCartney AL, Isolauri E. Influence of mode of delivery on gut microbiota composition in seven year old children. Gut. 2004Sep;53(9):1388–9.
95. Johnson KVA. Gut microbiome composition and diversity are related to human personality traits. Human Microbiome Journal. 2020Mar;1(15): 100069.
96. Dinan TG, Kennedy PJ, Morais LH, Murphy A, Long-Smith CM, Moloney GM, et al. Altered stress responses in adults born by caesarean section. Neurobiol Stress. 2021Dec;28(16): 100425.
97. Notarbartolo V, Giuffrè M, Montante C, Corsello G, Carta M. Composition of human breast milk microbiota and its role in children's health. Pediatr Gastroenterol Hepatol Nutr. 2022May;25(3):194–210.
98. Lundgren SN, Madan JC, Karagas MR, Morrison HG, Hoen AG, Christensen BC. Microbial communities in human milk relate to measures of maternal weight. Front Microbiol. 2019Dec;20(10):2886.
99. Lundgren SN, Madan JC, Karagas MR, Morrison HG, Christensen BC, Hoen AG. Human milk-associated bacterial communities associate with the infant gut microbiome over the first year of life. Front Microbiol. 2023Apr;17(14):1164553.
100. Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, et al. Temporal variability is a personalized feature of the human microbiome. Genome Biol. 2014Dec 3;15(12):531.
101. Mohr AE, Ahern MM, Sears DD, Bruening M, Whisner CM. Gut microbiome diversity, variability, and latent community types compared with shifts in body weight during the freshman year of college in dormitory-housed adolescents. Gut Microbes. 15(2):2250482.
102. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. Science. 2016Apr 29;352(6285):560–4.
103. Byrd DA, Sinha R, Hoffman KL, Chen J, Hua X, Shi J, et al. Comparison of methods to collect fecal samples for microbiome studies using whole-genome shotgun metagenomic sequencing. mSphere. 2020 Feb 26;5(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7045388/. Cited 2021 May 14
104. Laue HE, Bauer JA, Pathmasiri W, Sumner SCJ, McRitchie S, Palys TJ, et al. Patterns of infant fecal metabolite concentrations and social behavioral development in toddlers. Pediatr Res. 2024; 96(1):253–60.
105. Laue HE, Willis AD, Wang F, MacDougall MC, Xu Y, Karagas MR, et al. Github. HOME Study microbiome code. Available from: https://github.com/HEL548/HOME_Microbiome


## Publisher's Note