Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) in press
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Research Article

https://doi.org/10.1631/jzus.B2400341

Metapangenomics reveals an increased proportion of an *Escherichia coli*-dominated enterotype in older Chinese people

Jinyou LI[1*], Yue WU[1*], Yichen YANG[1], Lufang CHEN[1], Caihong HE[1], Shixian ZHOU[1], Shunmei HUANG[1], Xia ZHANG[1], Yuming WANG[1], Qifeng GUI[1], Haifeng LU[2], Qin ZHANG[1]✉, and Yunmei YANG[1]✉

[1]*Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China*
[2]*State Key Laboratory for Diagnosis and Treatment of Infectious Disease, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China*

Abstract: It is likely that gut microbial communities are remodeled in tandem with accumulated physiological decline during aging, yet there is limited understanding of gut microbiome variation in advanced age. Here, we performed a metagenomics-based enterotype analysis in a geographically homogeneous cohort of 367 enrolled Chinese individuals between the ages of 60 and 94, with the goal of characterizing the gut microbiome of the older individuals and identifying factors linked to enterotype variations. In addition to two adult-like enterotypes dominated by *Bacteroides* (ET-*Bacteroides*) and *Prevotella* (ET- *Prevotella*), we identified a novel enterotype dominated by *Escherichia coli* (ET- *Escherichia*), whose prevalence increased in advanced age. Our data demonstrated that age explained more of the variance in the gut microbiome than previously identified factors such as type 2 diabetes mellitus (T2DM) or diet. We characterized the distinct taxonomic and functional profile of ET- *Escherichia*, and found the strongest cohesion and highest robustness of the microbial co-occurrence network in this enterotype, as well as the lowest species diversity. In addition, we carried out a series of correlation analyses and co-abundance network analyses which showed that several factors were likely linked to the overabundance of Escherichia members, including advanced age, vegetable intake, and fruit intake. Overall, our data revealed an enterotype variation characterized by *E. coli* enrichment in the elderly population. Considering the different age distribution of each enterotype, these findings provide new insights into the changes that occur in the gut microbiome with age and highlight the importance of microbiome-based stratification of older individuals.

unedited

Key words: Gut microbiome; Aging; Elderly population; *Escherichia*-dominated enterotype; Enterotype; *Escherichia coli*

1 Introduction

Human aging and disease are deeply influenced by the gut microbiota. Gut microbiota and their metabolites can drive age-related damage such as chronic inflammation and oxidative stress (Ghosh et al., 2022; Mossad et al., 2022). As people live longer with the help of modern medicine, it has become more urgent to promote healthy aging. Advanced age features a combination of multiple factors that affect gut microbiota and cause dynamic remodeling of microbial communities in older adults (Pang et al., 2023).

✉ Yunmei YANG, 1194070@zju.edu.cn
Qin ZHANG, zhangqin1978@zju.edu.cn
* The two authors contributed equally to this work
Yunmei YANG, https://orcid.org/0000-0002-6646-4954
Qin ZHANG, https://orcid.org/0000-0003-4161-1871

Received July 8, 2024; Revision accepted Sept. 13, 2024;
Crosschecked xxx. xx, 20xx; Published online xxx. xx, 20xx

© Zhejiang University Press 2024

The advantage of enterotype classification is that the large number of individual microbiome profiles can be represented by several typical patterns, which can then be linked to clinical traits to evaluate disease risk. Enterotypes are not always stable. A transition of infant enterotypes to adult-like enterotypes has been observed at an early age (Xiao et al., 2021). Similarly, a transition from the adult-like enterotypes to an aging-associated enterotype in older individuals has been confirmed in a large-scale, cross-sectional investigation on gut microbial alterations during the aging of centenarians (Pang et al., 2023). However, few comprehensive analyses based on metagenomic sequencing data have been performed to date in terms of enterotypes within the elderly population.

In this study, we conducted a metagenomic-based enterotype analysis of 367 elderly people aged 60 to 94 years to investigate variations in the gut microbiome profile with sufficient metadata to explore the relationships between health status and enterotypes. A comprehensive analysis of the gut microbiome was performed to investigate enterotype-specific taxonomic and functional differences. In addition to the two adult-like enterotypes reported previously (ET-*Prevotella* and ET-*Bacteroides*), we identified and characterized an *E.coli*-enriched enterotype in the elderly population. This enterotype exhibits several aging signatures of gut microbiota that have been previously reported, such as impaired core gut microbes such as *Faecalibacterium prausnitzii* (*F. prausnitzii*)(Ling et al., 2022). Moreover, several tools were employed in this study to delineate and investigate gut microbial co-occurrence networks, and identify key factors that may be linked to these variations (Herren and Mcmahon, 2017; Kuntal et al., 2019; Shi et al., 2020). In our cohort, the effect of age on the human gut microbiome was more potent than that of previously reported factors. These factors include chronic noncommunicable diseases such as type 2 diabetes mellitus (T2DM) (Wang et al., 2022) and long-term dietary habits(Gacesa et al., 2022). Our study expands earlier observations on enterotypes which were reported in the PG cohort and Guangxi cohort in the elderly population, and further characterizes the *Escherichia* -dominated enterotype in terms of microbiome composition, functional profile, and microbial co-occurrence networks (Xu and Zhang, 2021; Pang, et al., 2023). The findings provide new insights into aging-associated signatures of the gut microbiome and advance our understanding of meaningful host-microbiota associations. This should prove helpful in identifying potential targets for microbiota-targeted interventions.

unedited

2 Methods

2.1 Study cohorts

In this study, we used metagenomic shotgun sequencing to investigate the microbial community structure of fecal samples from 367 older Chinese individuals (154 males and 213 females) aged 60 to 94 years. Time-matched clinical phenotypic data were collected to investigate external influencing factors associated with enterotype variations in advanced age. Advanced organic disease (end-stage cancer or renal or liver disease) and cognitive impairment were exclusion criteria. Participants could not have taken antibiotics within 6 months prior to enrollment; proton pump inhibitors, H2 receptor antagonists, tricyclic antidepressants, narcotics, anticholinergic medications, laxatives, or antidiarrhea medications within 4 weeks of enrollment; or antacids within 2 weeks of enrollment. Each volunteer completed questionnaires to provide clinical information, including demographic, lifestyle, and diet-frequency data, as well as sleep information. T2DM, dyslipidemia, hypertension, frailty, and sarcopenia were diagnosed by trained doctors. The Chinese volunteers enrolled were living in three adjacent residential communities in Shaoxing City, Zhejiang Province (in the Yangtze River Delta metropolitan area of China). These communities were formed during the process of rural urbanization, with most residents initially making a living by agriculture and gradually shifting to industrial work since the 1980s. They maintained traditional large family settlement patterns in combination with a modern urban lifestyle.

2.2 Sample collection

Fecal samples were meticulously collected from volunteers, using disposable bedpans to prevent con-

tamination from toilet water. Each stool specimen was carefully obtained with a DNA collection kit containing a proprietary chemical DNA stabilizer and transferred to the laboratory in liquid nitrogen. These samples were then stored at −80 °C in a long-term biobank, awaiting extraction of total DNA for sequencing purposes.

2.3 Bioinformatic processing

We used metagenomic shotgun sequencing to analyze the fecal samples. DNA from the fecal samples (~200 mg) was extracted from the supernatant using a QIAamp DNA Stool Mini Kit (Qiagen) according to the manufacturer's instructions. DNA library construction (based on DNA nanospheres and combinatorial probe-anchor synthesis-based shotgun metagenomic sequencing at a depth of 5 Gb per sample for all 367 samples) was performed at MGI (Shenzhen, China) using their routine procedures. An overall accuracy-control strategy (≥0.8) was used to perform quality control on the raw sequencing reads by removing low-quality reads.

2.4 Profiling microbiome composition and function

Sequencing libraries were metagenomically classified using MetaPhlAn3 to obtain gut microbial profiles that included bacteria, archaea, eukaryotes, and viruses. Taxon-specific community functional profiles were generated using Human Microbiome Project (HMP) Unified Metabolic Analysis Network 2 (HUMAnN2) (Franzosa et al., 2018). The National Center for Biotechnology Information (NCBI) database (2014 edition) and HUMAnN2 were used to annotate the nonredundant gene sets. The functional profiles were annotated according to metabolic pathways listed in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Wibowo et al., 2021) and MetaCyc-reactions functional categories, as described in previous studies (Roth-Schulze et al., 2021). The relative abundance of each taxonomic unit was determined by summing the relative abundances of the unit's annotated genes per individual (Qin et al., 2012). For gut microbial diversity analyses, diversity indices (Shannon and Simpson) were calculated based on the abundance profiles at the species level, using the 'diversity' function in the *R* package 'vegan' (v2.5–6).

To investigate the difference in metabolic functional profiles among enterotypes, we performed Wilcoxon rank-sum tests. This allowed us to compare the relative abundance of functional pathways grouped into different metabolic categories. Jensen–Shannon divergence (JSD) distance and partitioning around the medoid (PAM) clustering algorithm were applied to cluster samples based on the microbial functional profile (Chen et al., 2020).

unedited

2.5 Enterotype clustering

Enterotypes were classified in the cohort (n=367) according to methods previously described by Arumugam et al (Arumugam et al., 2011). We applied low-abundance and low-prevalence filtering to exclude taxonomic features whose relative abundance across all samples was lower than 0.1% and whose prevalence was lower than 10%. PAM using the JSD of the normalized species counts was performed to cluster the samples. The analysis was performed using the "ade4," "cluster," and "clusterSim" packages in R. The Calinski–Harabasz (CH) index indicated that the optimal number of clusters for this dataset was 3 (Fig. S1a). The dominant species with both high relative abundance in samples and high mean decrease accuracy were considered the main contributors to each enterotype. The significance of the model and cross-validated R^2 was evaluated with 5000 permutations of the response variable (ecosystem multifunctionality) using the 'A3' R package.

2.6 Independent validation cohorts of metagenomic data from public datasets

Available shotgun metagenomic data public datasets were collected to assess whether the age-related gut microbial characteristics in the Shaoxing cohort could be generalized to other older populations. We performed RDA analysis to visualize the beta-diversity of a Japan CRC cohort of 616 individuals (DRA006684 and DRA008156) and a combined cohort based on the metagenomic data of CRC patients (aged ≥60 years) from European and American countries, which included: (1) 134 metagenomic samples from an Austria CRC cohort

(ERP008729); (2) 129 individuals from a French & German CRC cohort (ERP005534); (3) 61 individuals from an Italy CRC cohort (SRP136711); (4) 45 individuals from a German CRC cohort (PRJEB27928); and (5) 64 individuals from USA CRC cohort (PRJEB12449) (Zeller et al., 2014; Feng et al., 2015; Vogtmann et al., 2016; Thomas et al., 2019; Wirbel et al., 2019; Yachida et al., 2019). All published metagenomic datasets were processed using the same pipeline to reduce technical bias in the bioinformatics analysis.

2.7 Co-occurrence network and diversity analyses

For the co-occurrence network analysis, we computed the bacterial correlations in each enterotype based on the abundance of each species, using the SparCC function of the 'SpiecEasi' package. We performed"100 bootstraps to estimate the *P* value constructed, with the 'SpiecEasi' package in R. Prior to this analysis, species with prevalences lower than 20% were removed to decrease noise. Absolute correlation values higher than 0.4 with a *P* value <0.05 were retained. We calculated the closeness of species in the main microbial community groups to measure node centrality in each network. To identify external influencing factors of microbial co-occurrence networks, 28 variables were considered as potential covariates, including demographic factors, chronic noncommunicable diseases, sleep time per day, habitual long-term diet information, and information about participation in sports or physical work. The correlation coefficients between potential covariates and topological features were calculated using Spearman's rank correlation test. We used multiple regression on distance matrices (MRM) in the *ecodist* package to estimate the importance of potential covariates. The robustness of a network was revealed by the natural connectivity of a complex network(Shi, et al., 2020). Positive and negative cohesion values were calculated as described in previous studies(Herren and Mcmahon, 2017). We calculated the NetShift (NESH) score to quantify community changes in gut microbial co-occurrence networks between enterotypes (https://web.rniapps.net/netshift)(Kuntal, et al., 2019). Species with NESH score values >2 were identified as important microbial taxa that serve as 'drivers' facilitating prominent changes in networks. All the undirected network graphs were visualized using Gephi (https://gephi.org).

2.8 Identification of external variable associations with the three enterotypes

Potential covariates, including clinical characteristics, dietary habits, and participation in sports or physical work, were used as explanatory variables in redundancy analysis (RDA). The analysis was computed based on Bray–Curtis dissimilarity matrices based on the relative abundances of species, using the vegan package. The associations between potential covariates and the overall bacterial composition were determined by *PERMANOVA* at the species level. We used Spearman's rank correlation test to calculate the correlations between potential covariates and the relative abundances of selected species. We also applied the multiple regression model with variation decomposition analysis to estimate the importance of the potential covariates in explaining the dissimilarities in microbial communities and the differences in the relative abundance of selected species and microbial metabolic pathways. This analysis was performed using the 'psych', 'reshape2', 'relaimpo', and 'packfor' packages in *R* (Jiao et al., 2020).

2.9 Statistical analysis

All statistical analyses were conducted and visualized in *R* within RStudio. The statistical significance of differences in the relative abundance of each taxon among enterotypes was determined by the Kruskal–Wallis H-test, with Benjamini–Hochberg false discovery rate (FDR) adjustment. The statistical significance of the differences between enterotypes in the top 25 dominant species with average relative abundance >1% in the cohort was also determined by the Wilcoxon rank-sum test. Correction for multiple testing was based on the Benjamini–Hochberg FDR, with the corrected *P value* cutoff set at 0.05.

Unedited

3 Results

3.1 Three enterotypes were identified in the gut microbiome of elderly individuals

In this study, we performed a metagenomics-based enterotype analysis in a cohort of 367 enrolled Chinese seniors between the ages of 60 and 94 years. After bioinformatic analysis, taxonomic assignment showed that 98.995%, 0.736%, 0.101%, and 0.001% of the sequencing reads corresponded to the domains *Bacteria*, *Viruses*, *Archaea*, and *Eukaryota*, respectively. Using JSD distance and the PAM clustering algorithm, we identified three enterotypes in this cohort (Fig. 1a). The predominant species identified by fivefold cross-validation analysis using a random forest model were *Prevotella copri* and *Prevotella stercorea* in enterotype 1 (designated ET-*Prevotella*), *Bacteroides uniformis* and *Bacteroides coprocola* in enterotype 2 (designated ET-*Bacteroides*), and *E. coli* in enterotype 3 (designated ET- *Escherichia*) (Fig. 1b). In terms of β-diversity analysis, principal coordinate analysis (PCoA) plots demonstrated that sample clustering at the order level did not result in a notable separation between ET-*Prevotella* and ET-*Bacteroides*, but most ET- *Escherichia* samples were well separated from these other enterotypes (Fig. S1a); the separation of all three enterotypes became clearer with clustering at the species level (Fig. S1b). The gut microbial communities of the *Escherichia*-dominated enterotype exhibited a distinctive taxonomic profile compared with those of ET-*Prevotella* and ET-*Bacteroides* (Figs. 1c and 1d, File S2). The 25 most dominant taxa showed significantly different abundance patterns across enterotypes (Fig. S2). Notably, *F. prausnitzii*, a key anti-inflammatory bacterium that is depleted in early frailty (Jackson et al., 2016), was found to be significantly depleted in ET-*Escherichia*, accounting for 8.31% of ET- *Prevotella*, 6.88% of ET-*Bacteroides*, and 3.20% of ET-*Escherichia.* (Fig. S2b). A similar situation also existed with *Eubacterium rectale.* Overall, these results demonstrated enterotype variation in some older persons, characterized by a remarkable enrichment of *E. coli* and depletion of butyrate-producing bacteria in their gut microbiota.

unedited

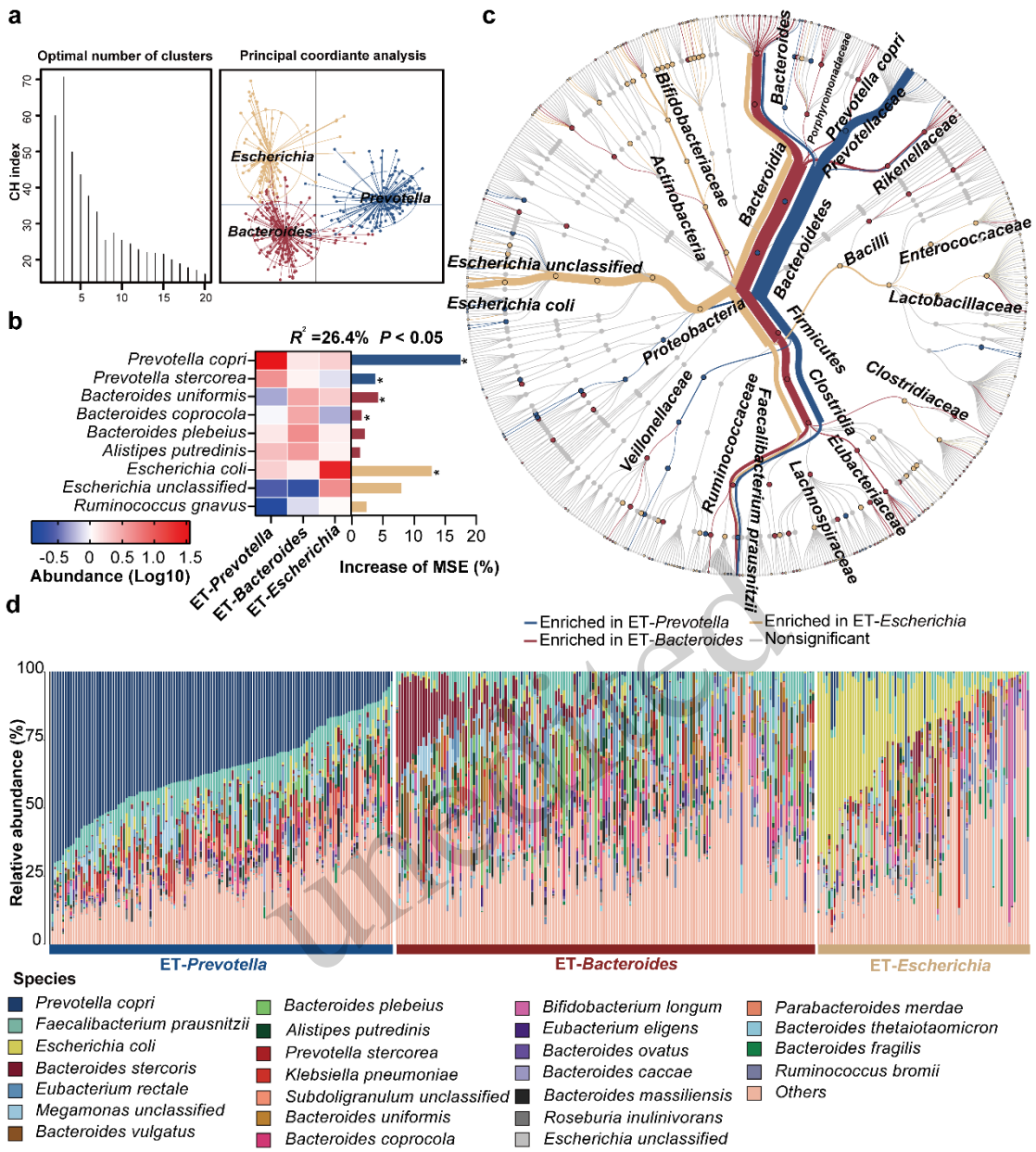


Fig. 1 Stratification of elderly participants into three enterotypes based on their gut microbiomes. (a) Three enterotypes identified in a cohort of 367 Chinese individuals aged 60–94 using Jensen–Shannon distance (JSD) and partitioning around medoid (PAM) clustering based on species profile. The labels ET- *Prevotella*, ET-*Bacteroides*, and ET-*Escherichia* represent the three enterotypes. The Calinski–Harabasz (CH) index was calculated for each set of clusters generated by PAM clustering. The CH index was maximized when gut microbial communities were assigned to three enterotypes. (b) Bacterial contributors to each enterotype identified by fivefold cross-validation analysis using a random forest model. The heatmap and bar chart show the key bacteria that distinguish the enterotypes. The top 10 bacterial species with cross-validated R^2 scores >0.5% in each enterotype were viewed as the bacteria most responsible for the observed clustering in the dataset. *P < 0.05. (c) Taxonomic differences in the bacterial communities. The thickness of a branch indicates the average relative abundance of each taxonomic category that is a main contributor to each enterotype. Each node in the phylogenetic tree represents one microbial taxon. Blue, red, and gold indicate the most abundant microbial taxa for ET- *Prevotella*, ET-*Bacteroides*, and ET-*Escherichia*, respectively. (d) Composition of species-level proportions for all 367 subjects across three enterotypes, as determined by shotgun metagenomic sequencing.

unedited

3.2 Enterotype variation in the older Chinese population reflects older age

To better understand the associations between enterotypes and health status in the elderly population, we examined 28 variables as potentially significant microbiota covariates, including demographic factors (2), chronic noncommunicable diseases (10), sleep time per day, habitual long-term diet information (11), and participation in physical work, daily indoor exercise, or outdoor sports (4). The clinical characteristics of participants have been summarized in Table S2. Among the variables tested in this study, the results of *PERMANOVA* clearly suggested that age was the strongest factor associated with overall microbiota composition, with 3.61% explained variance (Adonis analysis calculated based on Bray–Curtis similarity, $P<0.001$, Fig. 2a). Additional variables achieving significance were, in order, constipation, egg consumption, bone loss, T2DM, vegetable consumption, gastrointestinal disturbance, soy consumption, and stroke. We found significant differences in enterotype distribution in different age groups, but not in gender groups (Figs. 2b and 2c). The RDA showed that age, frailty, sarcopenia, T2DM, and fruit consumption had a higher impact on ET-*Escherichia* samples, while vegetable consumption and other potential covariates had opposite effects on gut community composition (Fig. 2d). Specifically, ET- *Prevotella* was the most prevalent enterotype among people younger than 75 years of age, whereas people over 75 years of age tended to harbor ET-*Escherichia* or ET-*Bacteroides.* The prevalence of the *E. coli*-dominant enterotype tended to increase dramatically with age. We next collected metagenomic data from published projects to validate the associations between increased *E. coli* abundance and age in older populations. In the Japan cohort of patients with colonic cancer (CRC) and a combined cohort based on metagenomic data of CRC patients from European and American countries, enrichment of *E. coli* is closely related to age, although its effect is weaker than that of CRC (Figs. 2e and 2f).

We next integrated the Spearman correlation coefficients with other statistical analyses to quantify the associations between key distinguishing bacterial species and all the variables for the three enterotypes (Fig. 2g). The data showed that potential covariates explained 14.53% of the variation in abundance for *P. copri* across the study population, but only 4.51% for *E. coli*, indicating that the abundance of *P. copri* in the elderly population is more susceptible to external influences. In addition to age, the relative abundance of unclassified *Escherichia* was positively correlated with fruit intake but negatively correlated with vegetable intake and constipation. Our results indicated that the enterotype's prevalence in the elderly population is highly correlated with advanced age and age-related diseases, including frailty, T2DM, arthritis, and metabolic dysfunction. The altered gut microbiome in older age might be a result of long-term adaptation to accumulated changes in dietary habits, gut physiology, and gastrointestinal motility.

Unedited

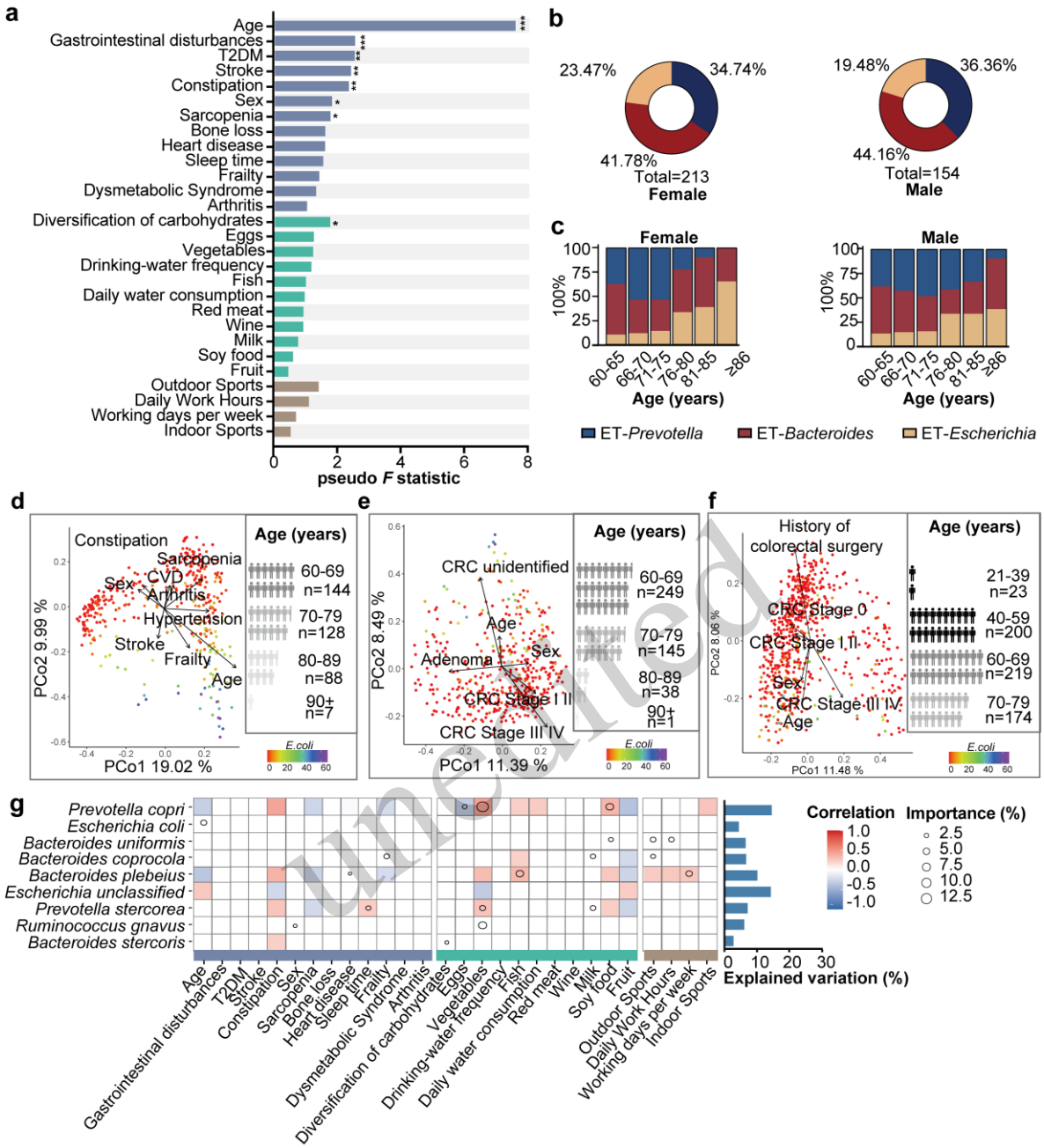


Fig. 2 Associations of enterotypes with age and other external factors. (a) Significant associations between external factors and gut microbiome composition. The amount of variance (R^2) explained by external factors, as assessed using *PERMANOVA* based on Bray–Curtis distance (*adjusted P*<0.05, *n*=367). (b) Pie charts indicate the percentage of individuals with each enterotype in males and females. (c) Bar plots indicate the percentage of individuals with each enterotype in each age range. d-f) Redundancy analysis (RDA) and principal coordinates of the Bray-Curtis distance matrix, calculated based on microbial species of the Shaoxing cohort (d), a combined cohort of 433 older individuals from European and American countries (e), and 616 individuals from Japan (f). Each point represents a sample and the color indicates the relative abundance of *E. coli*. Arrows indicate the influence of diseases and fecal-sample metadata. Color indicates relative abundance of *E. coli*. (g) Contributions of clinical and demographic variables to the differences in relative abundance of key distinguishing bacterial species for the three enterotypes, based on Spearman correlation coefficients and the best multiple regression model. Bar color shows the correlation value, where red indicates a positive association and blue indicates a negative association; only significant correlations are shown. Circle sizes represent the importance value (%).

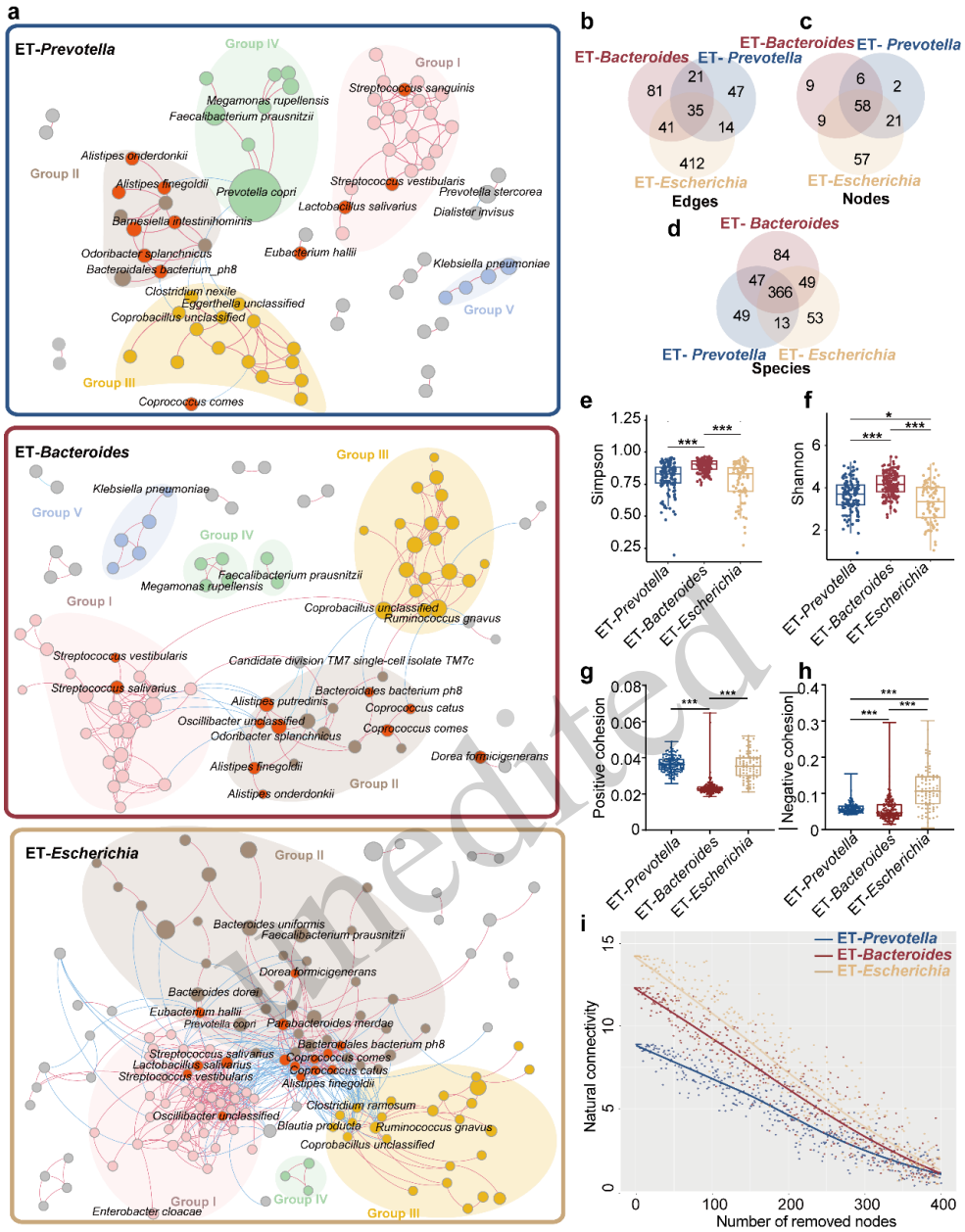
unedited

3.3 Distinctive pattern of the microbial co-occurrence network in ET-*Escherichia*

We performed co-occurrence network analysis to predict biotic interactions such as resource competition and metabolic cross-feeding within the resident microbial community, as revealed by negative edges and positive edges, respectively (Fig. 3a). Approximately 82-145 species (nodes) and 117-502 connections (edges) were retained at a correlation cutoff of 0.4 in the co-occurrence network for an enterotype, and most nodes and edges were specific to each enterotype (Figs. S3a and S3b). To investigate the bacterial consortia in each co-occurrence community, we analyzed the five main microbial community groups (group I to group V) of co-occurring species that appeared across enterotypes. The maximum number of unique nodes and edges was found in the co-occurrence network of ET-*Escherichia*, which was highly complex (Figs. 3b and 3c). The main microbial community groups consisted primarily of taxa from the same or closely related species. Specifically, 13 species appeared in group I across all three enterotypes, mainly from the genera *Streptococcus*, *Veillonella*, *Granulicatella*, *Lactobacillus*, and *Rothia*; 6 species mainly from the genus *Alistipes* in group II; and 7 species mainly from the genera *Clostridium*, *Blautia*, *Coprobacillus*, and *Flavonifractor* in group III. The main microbial community groups in ET-*Bacteroides* and ET-*Prevotella* consisted of similar components, suggesting high community homogeneity in their co-occurrence network when compared with that of ET-*Escherichia*. In line with that hypothesis, the nodes of ET-*Escherichia* showed significantly higher NESH scores than those of ET-*Bacteroides* or ET-*Prevotella* (Wilcoxon test, $P<0.001$; File S3 and Fig. S4a). Notably, there were more oral microbes in co-occurrence group I of the ET-*Escherichia* co-occurrence network.

There was a significant difference among the network-level topological features of the co-occurrence networks of the three enterotypes, including degree assortativity, average path length, and betweenness centralization (Kolmogorov–Smirnov test, $P<0.001$; Table S3). Fourteen nodes with NESH score values >2 were considered the key drivers facilitating the main changes between co-occurrence networks, and most of them showed significantly different enrichment or depletion across enterotypes, except *Lactobacillus salivarius* and *Streptococcus salivarius* (Figs. S4e and S4g). In addition, ET-*Escherichia* exhibited the lowest α-diversity indices among the three enterotypes and the exclusive growth of one single co-occurrence group in an older individual, implying that many more out-group species were suppressed by strong colonization resistance in ET-*Escherichia* (Figs. 3e and 3f). The maximum number of unique nodes and edges was found in the co-occurrence network of ET-*Escherichia*, which exhibited a highly complex co-occurrence network. In line with that, the gut microbial network of ET-*Escherichia* also showed the strongest cohesion value (Figs. 3g and 3h). The stability of the microbial co-occurrence network was quantified by the robustness of the microbial networks via a natural connectivity analysis. The data showed that the gut microbial network of ET-*Escherichia* had the highest robustness and strongest cohesion (Fig. 3i). To identify key factors affecting the diversity and stability of the human microbiome, we assessed external influencing factors based on their correlations with these fourteen key drivers and topological features. For the microbial network of ET-*Bacteroides*, constipation was positively correlated with average path length (Spearman's $r=0.335$, $P<0.01$) and negatively correlated with degree assortativity (Spearman's $r=-0.32$, $P<0.01$). Interestingly, opposite associations with degree assortativity were also observed for vegetable and fruit consumption (Spearman's $r=-0.334$, $P<0.01$, for vegetable consumption and Spearman's $r=0.304$, $P<0.01$, for fruit consumption). Similarly, the relative abundance of *S. salivarius* was positively correlated with age and fruit consumption in seniors and negatively correlated with constipation and vegetable consumption (Spearman's $r=0.31$, $P<0.0001$, for age; Spearman's $r=-0.37$, $P<0.0001$, for constipation; Spearman's $r=-0.34$, $P<0.0001$, for vegetable consumption; and Spearman's $r=0.33$, $P<0.0001$, for fruit consumption). These results suggest that constipation symptoms and a preference for fruits vs. vegetables may contribute to variations in microbial co-occurrence networks in different enterotypes by changing the interactions among several specific microbes, such as *S. salivarius* (Fig. S5).

unedited



Unedited

Fig. 3 Discrepancies in co-occurrence networks of the three enterotypes in older persons. (a) The complex nature of inter-microbial interactions in the ecological community of each enterotype was characterized by co-occurrence networks, using graphs. The main groups of co-occurring species are presented in different colors, and smaller groups are shown in gray. Networks of species in the three enterotypes were identified by positive and negative correlations among the dominant bacteria. The red line indicates a positive correlation, and the blue line indicates a negative correlation (SparCC, pseudo *P* <0.05). Only the bacterial connections (edges) with correlation values >0.4 were retained. Each node in the network indicates a species. The size of each node is proportional to the relative abundance of each species. Nodes in red show driver microbes that significantly contributed to separation of the networks (NESH score value >2). (b and c) Venn diagrams showing the number of unique and shared edges (b) and nodes (c). (d) Venn diagram showing the number of shared and unique species in each enterotype. e and f) Species profile-based *α*-diversity of the gut microbiome. Diversity was measured based on the species profiles for the Simpson (e) and Shannon (f) indices. (g) Significant difference in the positive cohesion of the microbial co-occurrence networks among enterotypes. h) Significant difference in the absolute value of negative cohesion of the microbial co-occurrence networks among enterotypes. i) The highest natural connectivity was observed for the microbial co-occurrence network in ET-*Escherichi. P* values are derived from the Wilcoxon rank-sum test: **P* < 0.05, ***P* < 0.01, and ****P* < 0.001.

3.4 Functional variations across enterotypes are associated with age

Consistent with our previous analysis, PAM clustering of ET-*Escherichia* samples was also supported based on microbial functional profiles. Most ET-*Escherichia* samples grouped into a cluster discrete from the cluster consisting of the adult-like enterotype ET-*Prevotella* and ET-*Bacteroides* samples (Fig. 4a). As shown in the heatmap in Fig. S6, the functional profile of ET-*Escherichia* varied markedly from those of ET-*Bacteroides* and ET-*Prevotella.* Our data showed that 319 of 431 microbial metabolic pathways presented statistically significant differences among the three enterotypes, as identified by the Wilcoxon rank-sum test (Fig. 4b, Additional file 4). Of the 158 pathways with 100% prevalence in the cohort, 128 pathways presented significant differences in abundance among enterotypes (Fig. 4c). Overall, the core microbial functionality that facilitates bacterial survival is generally activated by different functional pathways in ET-*Escherichia* and the adult-like enterotypes, with involvement of all metabolic categories (Fig. S7). We next investigated the patterns of age-associated change in the functional profile in each enterotype. The trend of decreased relative abundance of metabolic pathways (such as glycolysis, glycerol degradation, stachyose degradation, and purine ribonucleoside degradation pathways) with age was similar in ET-*Bacteroides* and ET-*Prevotella*. The relative abundance of several amino-acid biosynthesis pathways involved in L-threonine, L-lysine, L-arginine, and L-ornithine biosynthesis gradually increased with age in ET-*Bacteroides* and ET-*Escherichia* (Fig. 4d). In the whole cohort, significant positive correlations of these pathways with age were confirmed using Spearman's rank correlation test (Fig. 5). Interestingly, inverse associations were observed for two polyamine biosynthesis pathways and two functional pathways involving the biosynthesis of proinflammatory components of bacterial wall components, such as O-antigen and UDP-N-acetyl-D-glucosamine, and their correlations with age and vegetable consumption were also inverted. In addition to two functional pathways involved in cell-structure biosynthesis, most of the functional pathways involved in amino acid biosynthesis (except L-lysine biosynthesis) were positively correlated with age and fruit consumption but negatively correlated with vegetable consumption. Notably, the relative abundance of two functional pathways involved in L-lysine biosynthesis (L-lysine biosynthesis III and L-lysine biosynthesis VI) was negatively associated with age and frailty and positively associated with vegetable, milk, and soy food consumption.

unedited

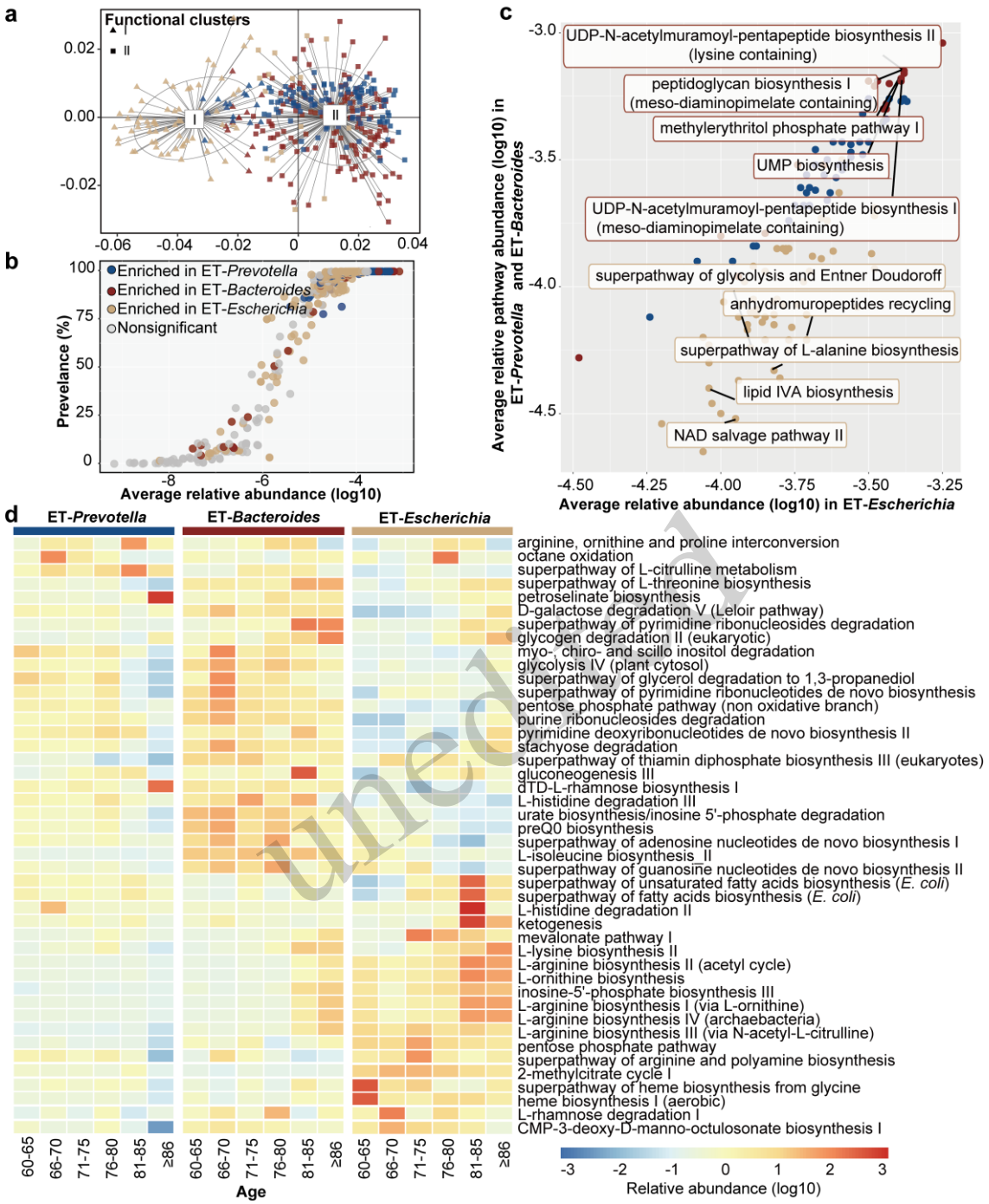


Fig. 4 Dramatic variation in the functional profiles of older persons with the *Escherichia coli*-enriched enterotype. (a) Principal coordinate analysis (PCoA) for visualization of functional clusters based on partitioning around medoids (PAM) of Jensen–Shannon distance (JSD). Point color represents the enterotype (red, ET-*Bacteroides*; gold, ET-*Escherichia*; blue, ET-*Prevotella*), and shape represents the functional cluster (triangle, cluster I; square, cluster II). (b) Functional differences identified by Wilcoxon rank-sum test are based on the metagenomic sequence data. Samples are colored to match their enterotype assignments. (c) 129 microbial metabolic pathways were significantly differentially abundant among the three enterotypes. The scatterplot shows only the pathways with 100% prevalence in our cohort. The ten most significantly differentiated metabolic pathways of the three enterotypes are labeled. (d) Enterotype- and age-specific patterns in the metabolic pathways. The heatmap shows the top 44 most significantly differentiated metabolic pathways of the three enterotypes, with relative abundance varying with age.

unedited

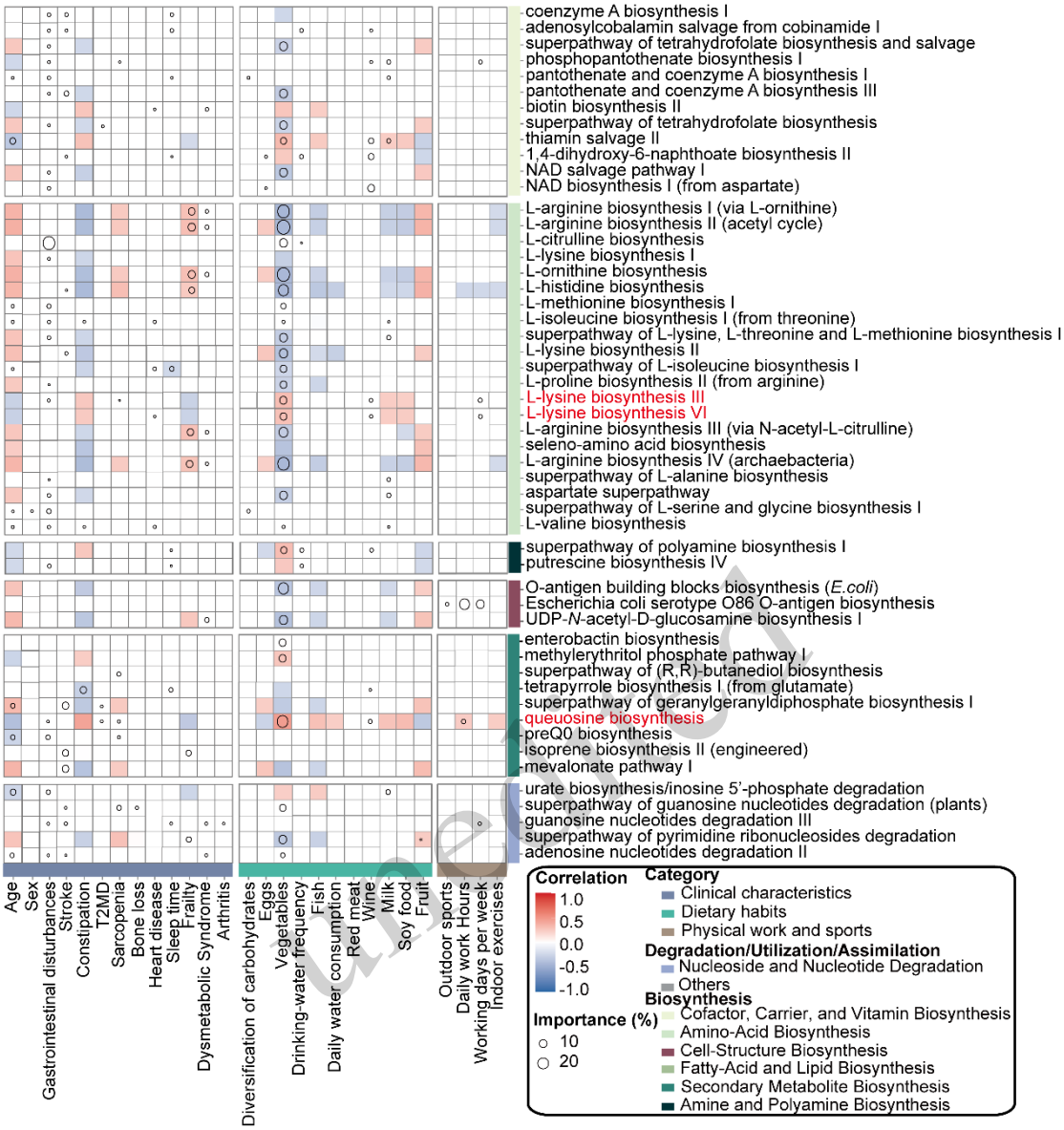


Fig. 5 Correlation between gut microbial functions and environmental factors. Rows correspond to metabolic pathways; columns correspond to clinical and lifestyle factors. Red indicates a positive association, and blue indicates a negative association. Only significant correlations are shown (*FDR*-corrected **P* < 0.05). The circle size represents the variable importance.

4 Discussion

In this study, we performed a metagenomic-based enterotype analysis and found an association between an increased prevalence of ET-*Escherichia* and advanced age in a cohort of older Chinese individuals. Compared with the previously reported effects of single factors on the gut microbiome (Gacesa, et al., 2022), such as chronic noncommunicable disease and long-term dietary habits, we show that the effects of age are much stronger, suggesting that the enterotype variations in older people may be a result of altered environmental factors superimposed on pathological changes and physiological aging. Similar to variations in the infant gut microbiome with respect to enterotype, which indicate different levels of gut maturation during infancy, our results suggest that the aging process could be accompanied by an enterotype switch.

Previous studies have shown that older adults harbor higher levels of *Enterobacteriaceae* and *E. coli* than younger ones (Leite et al., 2022; Pang, et al., 2023). A cross-sectional investigation of 1,575 individuals (20-117 years) from Guangxi Province, China reported an *Escherichia-Shigella*-dominated enterotype in advanced age using 16S rRNA sequencing data (Pang et al., 2023). This observation, which is inconsistent with ours, may be attributable to the high relatedness of DNA from *E. coli* and *Shigella* and the poor accuracy of 16S rRNA gene amplicon methods for *Escherichia/Shigella* (Martinson et al., 2019). *Shigella* species and *E. coli*, both Gram-negative bacteria within the *Enterobacterales* order, exhibit significant genetic similarities, but different levels of infectiousness and clinical outcomes. *E. coli* residency in the human gut has been consistently identified and investigated in the past century, by means of diverse microbiological and molecular techniques. As a constituent of the intestinal microbiome in more than 90% of individuals, *E. coli* is among the initial bacteria to establish colonization in neonates at birth (Martinson and Walk, 2020). Consumption of raw foods and water are recognized as primary pathways for bacterial and *E. coli* exposure. Additional common fomites include currency and cell phones (Pal et al., 2015; Abia and Ubomba-Jaswa, 2019).

Although the reasons for overgrowth of *E. coli* in the gut are not well understood, some available evidence implies that it results from a long-term adaptation of the gut microbiome to an aging host with a selective advantage in the presence of immune dysfunction. For example, it has been observed in some infants that the domination of *E. coli* was abolished as their immune function developed (Salahshouri et al., 2021). Animal studies have shown that nonpathogenic *E. coli* can thrive in IL-10 KO mice and complements factor-D-deficient mice; it appears to belong to the *Firmicutes* phylum rather than *Bacteroidetes* (Lupp et al., 2007; Qi et al., 2020). This suggests a link between *E. coli* overgrowth and immune dysfunction. Intestinal inflammation also produces reactive nitrogen species and reactive oxygen species that promote a growth advantage for *E. coli*. Meanwhile, the disruption of butyrate-producing bacteria during aging diminishes luminal butyrate, prompting intestinal epithelial cells to shift their metabolism towards anaerobic glycolysis, which is a pathway that bypasses oxygen consumption. This metabolic shift leads to an elevated release of oxygen from the colonic surface, which in turn fuels the proliferation of facultative anaerobes like *E. coli* within the gut ecosystem(Moreira De Gouveia et al., 2024). On the other side, overgrowth of *Escherichia* members might occur in a senior's gut when the dominance of core functional microbes, such as *P. copri*, *F. prausnitzii*, and *Megamonas unclassified*, is lost. Inhibition of *E. coli* growth by resident microbial communities has been observed in vitro using a gut microcosm approach (Baumgartner et al., 2020). Gut-environment alterations due to disease-related dysbiosis can facilitate the expansion of *Enterobacteriaceae*, owing to their superior adaptability in metabolizing emerging substrates. In line with this, our findings indicate that ET-*Escherichia* has a richer diversity of metabolic pathways, potentially a way of adapting to an aging host susceptible to various diseases. We also found that the relative abundance of unclassified *Escherichia* was inversely correlated with a preference for fruits or vegetables, although both are foods rich in dietary fiber and carbohydrates. One interpretation is that more live bacteria are consumed when eating fruit than when eating heat-treated vegetables and that the swallowed bacteria cannot be eliminated effectively in older people due to a decline in the gastric acid barrier and impaired immune function. Moreover, low-grade intestinal inflammation disrupts the balance of the normal gut microbiota and may also enhance the colonization of *E coli* (Davies et al., 2022). However, further research is needed to determine the links more definitively.

The highest absolute value of negative cohesion occurs in ET-*Escherichia*, which suggests that the strongest colonization resistance between the species exists in different co-occurrence groups. Negative cohesion in microbial co-occurrence networks may represent competition among co-occurrence groups of species with overlapping ecological niches, which may cause overactivation of host immune function or depletion of nutrient substrates to inhibit other organisms from growing nearby, finally resulting in disruption of intestinal homeostasis (Diard et al., 2013; Ghoul and Mitri, 2016). Accordingly, a predominance of *E. coli* or of a single co-occurrence group of species is observed in most older individuals with ET-*Escherichia*. In contrast, harmonious coexistence of different co-occurrence groups of species is present in the microbial community dominated by *Bacteroides*. Meanwhile α-diversity indices reveal that ET-*Bacteroides* samples show the highest

unedited

diversity, whereas ET-*Escherichia* samples show the lowest. This may explain the inconsistent reports of gut microbial diversity in older individuals across studies, which may be caused by different proportions of enterotypes like ET-*Escherichia* and ET-*Bacteroides* in the observed cohorts (O'toole and Jeffery, 2015; Kong et al., 2019).

The complex microbial interactions of ET-*Escherichia* are also associated with increased ectopic colonization of multiple oral bacteria in the gut, which is considered an important gut microbiota signature of aging and cancer (Larson et al., 2022; Tan et al., 2023). We found that the lactic acid producers *Bifidobacterium spp*. disappeared in co-occurrence group I of the ET-*Escherichia* co-occurrence network, whereas multiple members of the oral microbiota were present. Some of them were opportunistic human pathogens, such as *Atopobium parvulum*. *A. parvulum* has been reported to control the central hub of H_2S-producing bacterial communities, such as *Streptococcus* and *Veillonella*, in colorectal cancer (Mottawea et al., 2016; Yachida, et al., 2019).

A limitation of this study is that there are insufficient high-resolution temporal data based on longitudinal cohorts with large sample sizes to clarify age-dependent changes. Thus, patterns of enterotype transition during aging and the underlying mechanisms remain unclear. Future comprehensive examinations in humanized gnotobiotic mice and human fecal transplantation may provide a better understanding of the relationship of ET-*Escherichia* and healthy aging with regarding to its impacts on nutritional metabolites, host energy hemostasis, and immunosenescence.

In this study we found that the gut bacterial communities dominated by *P. copri* and *Bacteroides spp.* presented quite similar functional profiles. In contrast, the *E. coli*-dominated gut microbiome showed notable differences. Although non-pathogenic strains of *E. coli* that dominate the gut microbiota typically do not cause any noticeable symptoms, rapid turnover and a temporary increase of *E. coli* can lead to latent infections. This further underscores the importance of longitudinal monitoring of the gut microbiota, emphasizing the need to be vigilant about the potential risks associated with the emergence of pathogenic strains, particularly for the elderly population (Han et al., 2024). Our study should also contribute to the development of enterotype-specific nutritional management for healthy aging. Specifically, several functional pathways involved in the amino acid biosynthesis of L-ornithine, L-arginine, and L-lysine showed low abundance in ET-*Prevotella*. Additionally, a relative abundance of the functional pathways involved in polyamine biosynthesis is negatively correlated with age and shows an inverse association with functional pathways involving biosynthesis of the O-antigen building block (*E. coli*), a component of *E. coli* lipopolysaccharide. The anti-aging properties of polyamines have been determined in animal and clinical studies, so older individuals may benefit from dietary supplementation with polyamines (Yu et al., 2023). Live biotherapeutic agents could also be used to replenish the core gut bacteria lost in ET-*Escherichia* during aging, such as *F. prausnitzii*.

unedited

5 Conclusions

In conclusion, our data revealed an enterotype dominated by *E. coli* that was strongly associated with advanced age, indicating that enterotype transitions may occur not only in infants at an early age but also in some individuals of advanced age. Given that the microorganisms living in the intestine play an important role in maintaining host health, the aging-associated microbiome characterized by an overabundance of *E. coli* could profoundly affect age-related changes in metabolism, nutrient absorption, intestinal functions, and the immune system. Our integrated analysis of the gut microbial profile in older individuals will help researchers better understand the biological changes in the gut microbiome that occur with aging, which will assist in development of clinically relevant dietary recommendations or interventions targeting the gut microbiome of older adults.

Data availability statement

Raw sequencing data of all metagenomes have been deposited at the National Genomics Data Center BioProject database under the accession numbers PRJCA026979. All other data are available from the authors upon reasonable request (Jinyou Li, 20918253@zju.edu.cn).

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant Nos. 82101665, 82271588, 82200665, 82100795), the Zhejiang Provincial Natural Science Foundation of China (grant No. LY22H030009), the Zhejiang Provincial Science and Technology Program of Traditional Chinese Medicine (grant No. 2023ZL480), and the Medical and Health Research Project of Zhejiang Province (grant No. 2023RC153).

We would like to thank all of the volunteers who participated in this study, as well as the physicians from the Department of Geriatrics of the First Affiliated Hospital, School of Medicine, Zhejiang University, for their collaboration and help with sample collection. We thank the excellent analysis assistant of Core Facilities, Central Laboratory, the First Affiliated Hospital, Zhejiang University School of Medicine. We also thank the Center for Innovation & Translational Medicine, the First Affiliated Hospital, Zhejiang University School of Medicine, for the technical support.

Author contributions

All authors read and approved the final manuscript. Y-MY and QZ conceived and designed the experiments and carried out data collection, analysis, and interpretation. J-YL, Y-W, Y-CY, L-FC, C-HH, S-XZ, S-MH, XZ, Y-MW, Q-FG, and XL performed the experiments. J-YL and H-FL performed the statistical analyses and drafted the manuscript. J-YL prepared the figures and tables. J-YL wrote the manuscript.

Compliance with ethics guidelines

Jinyou LI, Yue WU, Yichen YANG, Lufang CHEN, Caihong HE, Shixian ZHOU, Shunmei HUANG, Xia ZHANG, Yuming WANG, Qifeng GUI, Haifeng LU, Qin ZHANG, and Yunmei YANG declare that they have no conflict of interest.

No conflict of interest exists in the submission of this manuscript, and the manuscript has been approved by all authors for publication. I would like to declare on behalf of my coauthors that the work described is original research that has not been published previously and is not under consideration for publication elsewhere, in whole or in part.

unedited

References

Abia ALK, Ubomba-Jaswa E, 2019. Dirty money on holy ground: Isolation of potentially pathogenic bacteria and fungi on money collected from church offerings. *Iran J Public Health*, 48(5):849-857.

Arumugam M, Raes J, Pelletier E, et al., 2011. Enterotypes of the human gut microbiome. *Nature*, 473(7346):174-180. https://doi.org/10.1038/nature09944

Baumgartner M, Bayer F, Pfrunder-Cardozo KR, et al., 2020. Resident microbial communities inhibit growth and antibiotic-resistance evolution of escherichia coli in human gut microbiome samples. *PLoS Biol*, 18(4):e3000465. https://doi.org/10.1371/journal.pbio.3000465

Chen L, Collij V, Jaeger M, et al., 2020. Gut microbial co-abundance networks show specificity in inflammatory bowel disease and obesity. *Nat Commun*, 11(1):4018. https://doi.org/10.1038/s41467-020-17840-y

Davies M, Galazzo G, Van Hattem JM, et al., 2022. Enterobacteriaceae and bacteroidaceae provide resistance to travel-associated intestinal colonization by multi-drug resistant escherichia coli. *Gut Microbes*, 14(1):2060676. https://doi.org/10.1080/19490976.2022.2060676

Diard M, Garcia V, Maier L, et al., 2013. Stabilization of cooperative virulence by the expression of an avirulent phenotype. *Nature*, 494(7437):353-356. https://doi.org/10.1038/nature11913

Feng Q, Liang S, Jia H, et al., 2015. Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat Commun*, 6:6528. https://doi.org/10.1038/ncomms7528

Franzosa EA, Mciver LJ, Rahnavard G, et al., 2018. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods*, 15(11):962-968. https://doi.org/10.1038/s41592-018-0176-y

Gacesa R, Kurilshikov A, Vich Vila A, et al., 2022. Environmental factors shaping the gut microbiome in a dutch population. *Nature*, 604(7907):732-739. https://doi.org/10.1038/s41586-022-04567-7

Ghosh TS, Shanahan F, O'toole PW, 2022. The gut microbiome as a modulator of healthy ageing. *Nat Rev Gastroenterol Hepatol*, 19(9):565-584. https://doi.org/10.1038/s41575-022-00605-x

Ghoul M, Mitri S, 2016. The ecology and evolution of microbial competition. *Trends Microbiol*, 24(10):833-845. https://doi.org/10.1016/j.tim.2016.06.011

Han N, Peng X, Zhang T, et al., 2024. Rapid turnover and short-term blooms of escherichia coli in the human gut. *J Bacteriol*, 206(1):e0023923. https://doi.org/10.1128/jb.00239-23

Herren CM, Mcmahon KD, 2017. Cohesion: A method for quantifying the connectivity of microbial communities. *ISME J*, 11(11):2426-2438. https://doi.org/10.1038/ismej.2017.91

Jackson MA, Jeffery IB, Beaumont M, et al., 2016. Signatures of early frailty in the gut microbiota. *Genome Med*, 8(1):8. https://doi.org/10.1186/s13073-016-0262-7

Jiao S, Yang Y, Xu Y, et al., 2020. Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern china. *ISME J*, 14(1):202-216. https://doi.org/10.1038/s41396-019-0522-9

Kong F, Deng F, Li Y, et al., 2019. Identification of gut microbiome signatures associated with longevity provides a promising modulation target for healthy aging. *Gut Microbes*, 10(2):210-215. https://doi.org/10.1080/19490976.2018.1494102

Kuntal BK, Chandrakar P, Sadhu S, et al., 2019. 'Netshift': A methodology for understanding 'driver microbes' from healthy and disease microbiome datasets. *ISME J*, 13(2):442-454. https://doi.org/10.1038/s41396-018-0291-x

Larson PJ, Zhou W, Santiago A, et al., 2022. Associations of the skin, oral and gut microbiome with aging, frailty and infection risk reservoirs in older adults. *Nat Aging*, 2(10):941-955. https://doi.org/10.1038/s43587-022-00287-9

Leite G, Pimentel M, Barlow GM, et al., 2022. The small bowel microbiome changes significantly with age and aspects of the ageing process. *Microb Cell*, 9(1):21-23. https://doi.org/10.15698/mic2022.01.768

Ling Z, Liu X, Cheng Y, et al., 2022. Gut microbiota and aging. *Crit Rev Food Sci Nutr*, 62(13):3509-3534. https://doi.org/10.1080/10408398.2020.1867054

Lupp C, Robertson ML, Wickham ME, et al., 2007. Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of enterobacteriaceae. *Cell Host Microbe*, 2(3):204. https://doi.org/10.1016/j.chom.2007.08.002

Martinson JNV, Pinkham NV, Peters GW, et al., 2019. Rethinking gut microbiome residency and the enterobacteriaceae in healthy human adults. *ISME J*, 13(9):2306-2318. https://doi.org/10.1038/s41396-019-0435-7

Martinson JNV, Walk ST, 2020. Escherichia coli residency in the gut of healthy human adults. *EcoSal Plus*, 9(1) https://doi.org/10.1128/ecosalplus.ESP-0003-2020

Moreira De Gouveia MI, Bernalier-Donadille A, Jubelin G, 2024. Enterobacteriaceae in the human gut: Dynamics and ecological roles in health and disease. *Biology (Basel)*, 13(3) https://doi.org/10.3390/biology13030142

Mossad O, Batut B, Yilmaz B, et al., 2022. Gut microbiota drives age-related oxidative stress and mitochondrial damage in microglia via the metabolite n(6)-carboxymethyllysine. *Nat Neurosci*, 25(3):295-305. https://doi.org/10.1038/s41593-022-01027-3

Mottawea W, Chiang CK, Muhlbauer M, et al., 2016. Altered intestinal microbiota-host mitochondria crosstalk in new onset crohn's disease. *Nat Commun*, 7:13419. https://doi.org/10.1038/ncomms13419

O'toole PW, Jeffery IB, 2015. Gut microbiota and aging. *Science*, 350(6265):1214-1215. https://doi.org/10.1126/science.aac8469

Pal S, Juyal D, Adekhandi S, et al., 2015. Mobile phones: Reservoirs for the transmission of nosocomial pathogens. *Adv Biomed Res*, 4:144. https://doi.org/10.4103/2277-9175.161553

Pang S, Chen X, Lu Z, et al., 2023. Longevity of centenarians is reflected by the gut microbiome with youth-associated signatures. *Nat Aging*, 3(4):436-449. https://doi.org/10.1038/s43587-023-00389-y

Qi H, Wei J, Gao Y, et al., 2020. Reg4 and complement factor d prevent the overgrowth of e. Coli in the mouse gut. *Commun Biol*, 3(1):483. https://doi.org/10.1038/s42003-020-01219-2

Qin J, Li Y, Cai Z, et al., 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*, 490(7418):55-60. https://doi.org/10.1038/nature11450

Roth-Schulze AJ, Penno MaS, Ngui KM, et al., 2021. Type 1 diabetes in pregnancy is associated with distinct changes in the composition and function of the gut microbiome. *Microbiome*, 9(1):167. https://doi.org/10.1186/s40168-021-01104-y

Salahshouri P, Emadi-Baygi M, Jalili M, et al., 2021. A metabolic model of intestinal secretions: The link between human microbiota and colorectal cancer progression. *Metabolites*, 11(7) https://doi.org/10.3390/metabo11070456

Shi Y, Zhang K, Li Q, et al., 2020. Interannual climate variability and altered precipitation influence the soil microbial community structure in a tibetan plateau grassland. *Sci Total Environ*, 714:136794. https://doi.org/10.1016/j.scitotenv.2020.136794

Tan X, Wang Y, Gong T, 2023. The interplay between oral microbiota, gut microbiota and systematic diseases. *J Oral Microbiol*, 15(1):2213112. https://doi.org/10.1080/20002297.2023.2213112

Thomas AM, Manghi P, Asnicar F, et al., 2019. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat Med*, 25(4):667-678. https://doi.org/10.1038/s41591-019-0405-7

Vogtmann E, Hua X, Zeller G, et al., 2016. Colorectal cancer and the human gut microbiome: Reproducibility with whole-genome shotgun sequencing. *PLoS One*, 11(5):e0155362. https://doi.org/10.1371/journal.pone.0155362

Wang H, Gou W, Su C, et al., 2022. Association of gut microbiota with glycaemic traits and incident type 2 diabetes, and modulation by habitual diet: A population-based longitudinal cohort study in chinese adults. *Diabetologia*, 65(7):1145-1156. https://doi.org/10.1007/s00125-022-05687-5

Wibowo MC, Yang Z, Borry M, et al., 2021. Reconstruction of ancient microbial genomes from the human gut. *Nature*, 594(7862):234-239. https://doi.org/10.1038/s41586-021-03532-0

Wirbel J, Pyl PT, Kartal E, et al., 2019. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat Med*, 25(4):679-689. https://doi.org/10.1038/s41591-019-0406-6

Xiao L, Wang J, Zheng J, et al., 2021. Deterministic transition of enterotypes shapes the infant gut microbiome at an early age. *Genome Biol*, 22(1):243. https://doi.org/10.1186/s13059-021-02463-3

Unedited

Xu X, Zhang Z, 2021. Sex- and age-specific variation of gut microbiota in brandt's voles. *PeerJ*, 9:e11434. https://doi.org/10.7717/peerj.11434

Yachida S, Mizutani S, Shiroma H, et al., 2019. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat Med*, 25(6):968-976. https://doi.org/10.1038/s41591-019-0458-7

Yu L, Pan J, Guo M, et al., 2023. Gut microbiota and anti-aging: Focusing on spermidine. *Crit Rev Food Sci Nutr*, :1-19. https://doi.org/10.1080/10408398.2023.2224867

Zeller G, Tap J, Voigt AY, et al., 2014. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol*, 10(11):766. https://doi.org/10.15252/msb.20145645

Supplementary information

Tables S1-S3, Figs. S1-S7, Files S1-S4

unedited