

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

## Jinyou Li

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Qifeng Gui

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Yichen Yang

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Caihong He

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Shunmei Huang

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Xia Zhang

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Jingjin Jiang

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Chunting Peng

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Yue Wu

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Lufang Chen

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

## Haifeng Lu

National Clinical Research Center for Infectious Diseases, Zhejiang University

**Yuanqiang Lu**

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

**Qin Zhang**

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

**Yunmei Yang (✉ [1194070@zju.edu.cn](mailto:1194070@zju.edu.cn))**

Key Laboratory of Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases of Zhejiang Province, the First Affiliated Hospital, School of Medicine, Zhejiang University

---

**Research Article**

**Keywords:** Gut microbiome, Aging, Elderly population, Escherichia coli, Enterotype, Frailty

**Posted Date:** February 11th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1344848/v1>

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

[Read Full License](#)

---

# Abstract

## Background:

Gut microbial communities are likely remodeled along with accumulated physiological decline during aging. 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 Chinese individuals and identifying factors linked to enterotype variations.

## Results:

In addition to two adult-like enterotypes were dominated by *Bacteroides spp.* (ET-*Bacteroides spp.*) and *Prevotella copri* (ET-*P. copri*), we further identified a novel enterotype dominated by *Escherichia coli* (ET-*E. coli*), whose prevalence increased in advanced age. ET-*E. coli* presented some characteristics previously observed in the microbiomes of older people, including low species diversity and diminished abundance of the butyrate producer *Faecalibacterium prausnitzii* (*F. prausnitzii*), and exhibited distinct functions as well as a complex, highly stable microbial co-occurrence network, as evidenced by cohesion and natural connectivity analysis. Furthermore, a series of correlation analyses and co-abundance network analyses were carried out and showed that several factors were likely linked to the overabundance of *Escherichia* members, including advanced age, different vegetable, and fruit intake. Notably, there may be a niche pre-emption of *E. coli* by *P. copri* in ET-*P. copri* via several functional pathways involved in queuosine, braided glycoside, UMP, L-lysine and aromatic amino acid biosynthesis. Overall, 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) and diet. Overgrowth of *Escherichia* members might occur in a senior's gut where the dominance of core functional microbiota was impaired by long-term changes in dietary habits, gut physiology, and gastrointestinal motility during aging.

## Conclusion:

This study demonstrates an enterotype variation featured by *E. coli* enrichment in some older Chinese individuals, indicating a potential dissimilation of the gut microbial community with altered host conditions eventually leading to enterotype transitions in advanced age. These findings provide new insights into the changes in the gut microbiome in older age and clues for mechanistic investigations on the driving forces.

## 1. Background

As people live longer with the help of modern medicine, it has become more urgent to promote healthy aging. Human aging and disease are deeply influenced by the gut microbiome[1-4]. However, the development of dietary recommendations or interventions targeting the gut microbiome of older adults for clinical relevance is challenging because of limited understanding of the large individual variation in

the gut microbiome. 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. However, few comprehensive analyses in terms of enterotypes have been performed within the elderly population to date. Although the transition of infant enterotypes to adult-like enterotypes has been observed at an early age[5-7], a transition from the adult-like enterotypes to an aging-associated enterotype in older individuals has not been confirmed.

Although several common age-related trajectories of the human gut microbiome have been observed in a Han Chinese population-based cohort of 2,338 adults (26–76 years), as well as in Israeli and Dutch adults, some of these characteristics, such as diversity, are inconsistent with the results reported in previous studies on the gut microbiota of the elderly[2, 8, 9]. This discrepancy is probably because dynamic remodeling is highly modifiable during aging, which causes the microbial communities in older individuals to be more complex and variable than those in younger adults[4, 10]. Moreover, previous studies have shown that the large inter- and intrasubject variation across gut microbiome studies performed across different regions and in populations with wide age ranges complicates the signatures of the aging-associated microbiome[11].

Therefore, we performed metagenomic-based enterotype analysis to characterize the aging-associated microbiome in a geographically homogeneous cohort of older individuals to minimize the variance in diet, lifestyle, and genetics. Comprehensive analysis of gut microbiomes was performed to investigate enterotype-specific taxonomic and functional differences. Moreover, several ecological theories and tools[12-14] were employed in this study to delineate competing and cooperating microorganisms, compare stabilities between co-occurrence networks, and identify key factors that may be linked to these variations. In our cohort of Chinese people older than 60 years of age, the effect of age on the human gut microbiome was more potent than that of previously reported factors, such as chronic noncommunicable diseases such as type 2 diabetes mellitus (T2DM)[15] and long-term dietary habits[16]. Our study expands earlier observations on enterotypes based on the PG cohort reported by *Zhang* et al. to include a population older than 76 years[8]. In addition to the two adult-like enterotypes (ET-*P. copri* and ET-*Bacteroides spp.*) reported previously, we first identified and characterized an *Escherichia coli*-enriched enterotype in the elderly population, which exhibits several aging signature as previously reported, such as low diversity[1] and low abundance of *F. prausnitzii*[17, 18]. Our results provide new insights into the aging-associated microbiome and enhance our understanding of how age and other external influencing factors contribute to enterotype variations.

## 2. Methods

### 2.1. Study cohorts

In this study, we used metagenomic shotgun sequencing to investigate the microbial community structures 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

that are 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 prior to enrollment. Each volunteer completed questionnaires to provide clinical information, including demographic, lifestyle, and diet frequency data and 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 while maintaining traditional large family settlement patterns with modern urban lifestyles.

## 2.2. Sample collection

Fecal samples of all 367 participants were collected after defecation at the visit and transported to the laboratory in liquid nitrogen. 223 participants were willing to donate 5 ml blood sample for our research. Fasting blood samples were taken between 7:00 and 9:00. Blood and fecal samples were stored at  $-80^{\circ}\text{C}$  for long-term biobanking until total DNA was extracted for sequencing.

## 2.3. Bioinformatic processing

We used metagenomic shotgun sequencing to analyze the fecal samples. DNA from the fecal samples ( $\sim 200$  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 were performed at MGI (Shenzhen, China) using their routine procedures. The overall accuracy ( $\geq 0.8$ ) control strategy was used to perform quality control on the raw sequencing reads to remove low-quality reads. Sequencing libraries were metagenomically classified using MetaPhlAn2[19] 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 (HUMAN2)[20]. The National Center for Biotechnology Information (NCBI) database (2014 edition) and HUMAN2 were used to annotate the nonredundant gene sets. The functional profiles were annotated according to Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathways[21] and MetaCyc-reactions functional categories as described in previous studies[22]. The relative abundances of each taxonomic unit were generated by summing the relative abundances of their annotated genes per individual[15].

For gut microbial diversity analyses, diversity indexes (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 profile among enterotypes, Wilcoxon rank-sum tests were performed to compare the relative abundance of functional pathways grouped into different metabolic categories. Jensen–Shannon divergence (JSD) distance and the partitioning around medoid (PAM) clustering algorithm were applied to cluster samples based on the microbial functional profile, and sparse inverse covariance estimation for ecological association inference (SPIEC-EASI) using the glasso model was applied to reduce the pathway dimension to the module[23]. Spearman’s rank correlations were used to determine the relationship between external variables and functional pathways and to determine the co-abundance that exhibits the relationship between species and functional pathways[24, 25]. All the undirected network graphs were visualized using Gephi (<https://gephi.org/>).

## 2.4. Enterotype clustering

Enterotypes were classified in the cohort (n = 367) according to methods previously described by Arumugam et al[26]. Low-abundance and low-prevalence filtering was applied 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 (Supplementary Figure 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  $R^2$  were evaluated with 5,000 permutations of the response variable (ecosystem multifunctionality) using the ‘A3’ R package.

## 2.5. Co-occurrence network and diversity analyses

For the co-occurrence network analysis, the bacterial correlations in each enterotype were computed based on the abundance of each species using the SparCC function of the ‘SpiecEasi’ package with 100 bootstraps to estimate the  $P$  value constructed using 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. The closeness of species in the main microbial community groups was calculated to measure node centralities 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. Multiple regression on distance matrices (MRM) in the *ecodist* package was used to estimate the importance of potential covariates. The robustness of a network was revealed by the natural connectivity of a complex network[13]. Positive and negative cohesion values were calculated as described in previous studies[14]. We calculated the NetShift (NESH) score to quantify community changes in gut microbial co-occurrence networks between enterotypes

(<https://web.mnapps.net/netshift>)[12]. Species with NESH score values > 2 were identified as important microbial taxa that serve as 'drivers' facilitating prominent changes in networks.

## 2.6. 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) computed based on Bray–Curtis dissimilarity matrices from 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. Spearman's rank correlation test was used to calculate the correlations between potential covariates and the relative abundances of selected species as the major contributor to each enterotype identification. 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 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*[27].

## 2.7. Piecewise linear regression analysis of the relationships between frailty risk markers and age in each enterotype

We assessed peripheral concentrations of fasting plasma glucose (FPG), hemoglobin A1c (HbA1c), creatinine, high-sensitivity C-reactive protein (hs-CRP), P, Ca, folic acid, and albumin levels using standard in-hospital methods. The relationships between age and peripheral levels of various risk markers were visualized with loess regression models. The breakpoint (age threshold) was determined using a piecewise linear regression analysis with the *R* package *SiZer* v.0.1.5[28]. For the analysis of FPG, the data from patients with T2DM were excluded due to their abnormal blood glucose levels.

## 2.8. 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 most 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.

# 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 analyses, 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 (**Supplementary Figure S1A**). 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-*P. copri*), *Bacteroides uniformis* and *Bacteroides coprocola* in enterotype 2 (designated ET-*Bacteroides spp.*) and *E. coli* in enterotype 3 (designated ET-*E. coli*) (**Figure 1A and 1B**). The gut microbial communities of the *E. coli*-dominated enterotype exhibited a distinctive taxonomic profile compared with those of ET-*P. copri* and ET-*Bacteroides spp.* (**Figure 1C-1F**). For  $\beta$ -diversity analysis, principal coordinate analysis (PCoA) plots demonstrated that sample clustering at the order level did not result in a notable separation between ET-*P. copri* and ET-*Bacteroides spp.*, but most ET-*E. coli* samples were well separated from these other enterotypes (**Supplementary Figure S1B**), the separation of all three enterotypes became clearer with clustering at the species level (**Supplementary Figure S1C**). The 25 most dominant taxa showed significantly different abundance patterns across enterotypes (**Supplementary Figure S2**). Notably, *F. prausnitzii*, a key anti-inflammatory bacterium that is depleted in early frailty[17], was found to be significantly enriched in ET-*P. copri* but depleted in ET-*E. coli*, accounting for 8.31% of ET-*P. copri*, 6.88% of ET-*Bacteroides spp.* and 3.20% of ET-*E. coli* (**Supplementary Figure S2B**). The taxonomic differences across enterotypes are further summarized in **Additional file 1**. Overall, these results demonstrated enterotype variation in some older persons, characterized by a remarkable enrichment of *E. coli* in their gut microbiomes.

### 3.2. Enterotype variation in the older Chinese population is associated with 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 exercising, or outdoor sports (4). The RDA showed that age, frailty, sarcopenia, T2DM, and fruit consumption had a higher impact on ET-*E. coli* samples, while vegetable consumption and other potential covariates had opposite effects on gut community composition (**Figure 2A**). Among the variables tested in this study, the results of *PERMANOVA* clearly suggested that age was the strongest factor associated with the overall microbiota composition, with 3.61% explained variance (Adonis analysis calculated based on Bray–Curtis similarity,  $P < 0.001$ , **Figure 2B**). Additional variables achieving significance, in order, were constipation, egg consumption, bone loss, T2DM, vegetable consumption, gastrointestinal disturbances, soy consumption, and stroke. Our results indicated that the enterotype prevalence in the elderly population is related to age. Specifically, ET-*P. copri* was the most prevalent enterotype among people younger than 75 years of age, whereas people over 75 years of age tended to harbor ET-*E. coli* or ET-*Bacteroides spp.* The prevalence of the *E. coli*-dominant enterotype tended to increase dramatically with age. ET-*Bacteroides spp.* was present at similar rates across all ages studied (**Figure 2C**). Deployment of piecewise linear regression analysis enabled us to visualize a clear difference in the age-related trajectories of representative blood inflammatory and nutritional marker levels among enterotypes. Many of them showed generally different trends of associations with age in the older adults with ET-*Bacteroides spp.* or ET-*E. coli*, characterized by different

slopes and age breakpoints of the relationships (**Figure 2D**), suggesting the inconsistent pace of aging in the individuals with different enterotypes.

We next integrated the Spearman correlation coefficients with other statistical analyses to quantify the associations between key distinguishing bacterial species for the three enterotypes with each variable (**Figure 2E**). The data showed that potential covariates explained 14.53% of the variation in abundance for *P. copri* across individuals 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. These results indicated that 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.

### 3.3. Distinctive pattern of the microbial co-occurrence network in ET-*E. coli*.

Most microorganisms living in the gut are auxotrophs whose growth relies both on external nutrients obtained from the diet or metabolism of the host and the exchange of electron donors, amino acids, vitamins, and other cofactors with other microbes[29]. Thus, a robust steady-state bacterial consortium is formed in the co-occurrence community based on cooperative interactions for low metabolic costs and efficiency within member species. In this study, co-occurrence network analysis was performed to predict biotic interactions such as resource competition and metabolic cross-feeding within the residing microbial community, as revealed by negative edges and positive edges, respectively. We investigated the difference in co-occurrence communities of the three enterotypes to look for potential influencing factors linked to transitions between the enterotypes (**Figure 3**). 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 (**Supplementary Figure S3**). The maximum number of unique nodes and edges was found in the co-occurrence network of ET-*E. coli*, which exhibited a highly complex co-occurrence network (**Figure 3A**).

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,  $P < 0.001$ , **Additional file 2**). 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* (**Supplementary Figure S4**). To identify key factors affecting the diversity and stability of the human microbiome, external influencing factors were assessed based on their correlations with these fourteen key drivers and topological features. For the microbial network of ET-*Bacteroides spp.*, 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 preference between fruits or 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* (**Supplementary Figure S5**).

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, whose total members accounted for 36% of the nodes in ET-*P. copri*, 44% in ET-*Bacteroides spp.*, and 45% in ET-*E. coli* (**Supplementary Figure S6A**). We found that these main microbial community groups consisted mainly of taxa from the same or closely related species (**Supplementary Figure S6B**). 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 spp.* and ET-*P. copri* consisted of similar components, suggesting high community homogeneity in their co-occurrence network when compared with that of ET-*E. coli*. In line with that hypothesis, the nodes of ET-*E. coli* showed significantly higher NESH scores than those of ET-*Bacteroides spp.* or ET-*P. copri* (Wilcoxon test, *P* < 0.001, **Additional file 3** and **Supplementary Figure S7**). Notably, there were more oral microbes in co-occurrence group I of ET-*E. coli* co-occurrence network.

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-*E. coli* had the highest robustness and the strongest cohesion (**Figure 3B-3E**). In addition, ET-*E. coli* exhibited the lowest  $\alpha$ -diversity indexes 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 a strong colonization resistance in ET-*E. coli* (**Figure 3F-3H** and **Supplementary Figure S8**). However, within ET-*E. coli* samples, *Spearman's* test showed a significantly positive correlation between the positive cohesion value and Shannon index, suggesting that the more microbes that benefited from cooperative interactions within a group, the lower the potential risk of extinction due to environmental degradation (**Figure 3I**).

### 3.4. Gut microbial function differences among enterotypes.

For this analysis, genes were grouped based on functional pathways. We found that PAM clustering of ET-*E. coli* samples was also supported based on microbial functional profiles. Most ET-*E. coli* samples grouped into a cluster discrete from the cluster consisting of the adult-like enterotype ET-*P. copri* and ET-*Bacteroides spp.* samples (**Figure 4A**). The distinct functionality of the gut microbiome dominated by *E.*

*coli* is shown in the heatmap in **Supplementary Figure S9**. Obviously, the functional profile of ET-*E. coli* varied markedly from those of ET-*Bacteroides spp.* and ET-*P. copri*. 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 (**Figure 4B, Additional file 4**). Of the 158 pathways with 100% prevalence in the cohort, 128 pathways presented significant differences in their abundance among enterotypes (**Figure 4C**). Overall, the core microbial functionality that facilitates bacterial survival is performed by generally different functional pathways in ET-*E. coli* and the adult-like enterotypes, with involvement of all metabolic categories (**Supplementary Figure S10**).

We next investigated the patterns of age-associated changes in the functional profile in each enterotype. The trend of the decreased relative abundance of metabolic pathways, such as glycolysis, glycerol degradation, stachyose degradation, and purine ribonucleoside degradation, with age was similar in ET-*Bacteroides spp.* and ET-*P. copri*. The relative abundance of several amino acid biosynthesis pathways involved in L-threonine, L-lysine, L-arginine, and L-ornithine biosynthesis diminished with age in ET-*P. copri* but gradually increased with age in ET-*Bacteroides spp.* and ET-*E. coli* (**Figure 4D**).

The SPIEC-EASI method was utilized to identify functional modules that were constructed by coabundant functional pathways (**Figure 4E**). Module I consisted of the highly abundant and highly prevalent functional pathways that were significantly enriched in the two adult-like enterotypes (ET-*P. copri* and ET-*Bacteroides spp.*), suggesting their particular importance for the core functionality of the human gut microbiome, whereas module II was composed mainly of the functional pathways significantly enriched in ET-*E. coli*, featuring many functional pathways involving cofactor, carrier, and vitamin biosynthesis (**Supplementary Figure S9**). Interestingly, vast negative correlations existed between most ET-*P. copri*-enriched pathways and five ET-*Bacteroides spp.*-enriched pathways in module I with some ET-*E. coli*-enriched pathways in module II, which indicated the potential participation of these microbial functions in the competition between the predominant bacteria of the enterotypes. The 336 negative correlations (edges) are listed in **Additional file 5**.

### 3.5. Factors associated with enterotype variations in the older population.

To investigate the factors associated with enterotype variations, especially for the overgrowth of *E. coli* in older individuals, we further analyzed differences in species–pathway and species–host associations among the three enterotypes. Application of Spearman correlation coefficient analysis showed that members of co-occurrence group I (mainly from *Streptococcus* and *Veillonella*) generally might have a niche overlap with *Enterobacteriaceae* members (e.g., *Klebsiella pneumoniae*, *Enterobacter cloacae*, *E. coli*, and *unclassified Escherichia*), who presented growth optima involving various metabolic pathways (**Figure 5**). Their close relationship existed in all enterotypes. In contrast, only in ET-*P. copri*, it was found that nine metabolic biosynthesis pathways showed a positive correlation with *P. copri* but a negative correlation with *E. coli*, suggesting their potential roles in the competition between microbes. These nine pathways were including queuosine, UMP, L-lysine (L-lysine biosynthesis III and L-lysine biosynthesis VI), and aromatic amino acid biosynthesis. We also observed similar relationships of *E. coli*

with *F. prausnitzii* linked by genes involved in L-isoleucine biosynthesis I and *Megamonas unclassified* linked by genes involved in Calvin-Benson-Bassham cycle, pentose phosphate pathway (non oxidative branch), and stachyose degradation. These results indicated the growth of *E. coli* might be suppressed in ET-*P. copri* via a niche pre-emption generated by *P. copri*, *F. prausnitzii*, and *Megamonas unclassified*.

We further analyzed the correlation between the relative abundances of metabolic pathways and age and other external influencing factors (**Figure 6**). Consistent with the apparent age-associated increase in the relative abundance of several amino acid biosynthesis pathways involved in L-threonine, L-lysine, L-arginine, and L-ornithine biosynthesis in the enterotypes dominated by *Bacteroides* spp. and *E. coli*, significant positive correlations of these pathways with age were confirmed in the whole cohort using Spearman's rank correlation test. Interestingly, inverse associations were observed in 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.

## 4. Discussion

Although a previous study showed that some older people harbor higher levels of *Enterobacteriaceae* (the family including *E. coli*) than younger adults[2, 30, 31], the dominance of *E. coli* in the gut microbiota might not have been observed in previous studies due to the poor accuracy of 16S rRNA sequencing for *Enterobacteriaceae* and *Escherichia/Shigella*[32]. This work included a metagenomic-based enterotype analysis and first reported the association of an increased prevalence of ET-*E. coli* with advanced age in a cohort of older Chinese individuals with geographic homogeneity. Compared with the previously reported effects of single factors on gut microbiomes[16, 33], such as chronic noncommunicable diseases 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. Moreover, we found that different enterotypes showed inconsistent age-related trajectories of representative blood inflammatory and nutritional markers, such as creatinine, whose increasing serum level indicates a higher risk of sarcopenia and low renal function in older individuals[34]. These findings in our study partly support the hypothesis that interindividual differences in the gut microbiota may indicate different degrees of aging for older individuals, just as differences in the intestinal bacterial community of infants in terms of enterotype represent different stages of maturation in early life[6].

We found that the relative abundance of unclassified *Escherichia* was inversely correlated with preference for fruits or vegetables, although both are foods rich in dietary fiber and carbohydrates. One interpretation is that more live bacteria may be 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. Although the reasons for the evident overgrowth of *E. coli* in the gut are not well understood, some available evidence implies that it may result 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[5]. Animal studies have shown that nonpathogenic *E. coli* could thrive in IL-10 KO mice and complement factor D-deficient mice, appearing to the bacteria from the *Firmicutes* phylum rather than *Bacteroidetes*[35, 36]. Consistent with this finding, we found that *E. coli* grew particularly abundantly in elderly people (>75 years) with significantly reduced relative abundances of the anti-inflammatory *F. prausnitzii* and *Ruminococcus* members. This observation suggests an association between *E. coli* overgrowth and immune senescence and inflammatory aging. On the other side, overgrowth of *Escherichia* members might occur in a senior's gut where the dominance of core functional microbes, such as *P. copri*, *F. prausnitzii*, and *Megamonas unclassified*, was lost. Inhibition of *E. coli* growth by resident microbial communities has been observed in vitro using a gut microcosm approach[37]. Our results support the mechanisms by which such suppression of invaders may be generated by resource competition. We found notable negative relationships between some signature pathways of the adult-like enterotypes and those of ET-*E. coli*. It is likely that the dominant growth of *P. copri* may deplete these substrates, such as L-aspartate for the biosynthesis of UMP and L-lysine, in the microenvironment and exclude other competitors overlapping the ecological niche. However, the metabolic intermediates and end-products in microbial resource competition have not been confirmed. It would be helpful to investigate the metabolic mechanisms involved in the ecological niche competition between *E. coli* and *P. copri* or other species by well-designed cell culture analyses in the future.

The redundancy of members in the co-occurrence network could mitigate the threat of extinction of gut microbes. In this cohort of older individuals, we found the highest robustness in the co-occurrence network of ET-*E. coli*, suggesting a greater capacity to resist fluctuations driven by the loss of members and a better adaptation to a senescent host[38]. Conversely, the weak robustness of ET-*P. copri* indicates the vulnerability of this ecosystem, which might explain why this enterotype disappears substantially with age in the population[29]. On the other side, 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[39, 40]. The highest absolute value of negative cohesion in ET-*E. coli* suggested the strongest colonization resistance between the species in different co-occurrence groups. Accordingly, the predominance of *E. coli* or a single co-occurrence group of species was observed in most older individuals with ET-*E. coli*. In contrast, harmonious coexistence of different co-occurrence groups of

species was present in the microbial community dominated by *Bacteroides spp.* And  $\alpha$ -diversity indexes revealed that ET-*Bacteroides spp.* samples showed the highest diversity, whereas ET-*E. coli* samples showed the lowest diversity. Thus, it is likely that in the gut microbial community dominated by *E. coli*, the "friend or foe" style survival strategy adopted by a few groups might result in a hostile and deteriorating environment for out-group members and reduce species diversity. That may explain the inconsistent reports of the gut microbial diversity in older individuals across studies, which may be caused by different proportions of enterotypes like ET-*E. coli* and ET-*Bacteroides spp.* in their observed cohorts[1, 30].

The complex microbial interactions of ET-*E. coli* are also associated with the increased ectopic colonization of multiple oral bacteria in the gut, which has been considered as an important gut microbiota signature of aging and cancer[9, 41]. We found that the lactic acid producers *Bifidobacterium spp.* disappeared in co-occurrence group I of the ET-*E. coli* co-occurrence network, whereas multiple members of the oral microbiota are present. Some of them are opportunistic human pathogens, such as *Atopobium parvulum*. *A. parvulum* has been reported to control the central hub of H<sub>2</sub>S-producing bacterial communities, such as *Streptococcus* and *Veillonella*, in colorectal cancer[42, 43].

Our study would contribute to the development of enterotype-specific nutritional management for healthy aging. The gut bacterial communities dominated by *P. copri* and *Bacteroides spp.* presented quite similar functional profiles. In contrast, loss of core module I and gain of module II make the *E. coli*-dominated gut microbiome functionally different from the two adult-like enterotypes. Specifically, several functional pathways involved in the amino acid biosynthesis of L-ornithine, L-arginine, and L-lysine showed low abundance in ET-*P. copri*. Additionally, the relative abundance of the functional pathways involved in polyamine biosynthesis is negatively correlated with age and shows an inverse association with functional pathways involving the biosynthesis of the O-antigen building block (*E. coli*), a component of *E. coli* lipopolysaccharide. As the anti-aging properties of polyamines have been determined in animal and clinical studies[44], older individuals may benefit from dietary supplementation with polyamines. In addition, live biotherapeutic agents could be used to replenish the lost core gut bacteria in ET-*E. coli* during aging, such as *F. prausnitzii*.

A limitation of this study is that there are insufficient well-controlled longitudinal human studies to support the hypothesis that the enterotype transition occurs in an individual during aging or that the mortality rates of older persons vary by enterotype. In addition, patterns of enterotype transition during aging and the underlying mechanisms remain unclear due to a lack of high-resolution temporal data based on longitudinal cohorts with large sample sizes to clarify age-dependent changes. Moreover, it is difficult to extrapolate fundamental aspects of nonpathogenic *E. coli* ecology and their effects on host health status from studies of pathogenic and/or antibiotic-resistant strains from an unhealthy person, whereas most nonpathogenic *E. coli* strains living in the human gut may be common residents[45]. Therefore, caution should be used when assuming that ET-*E. coli* is detrimental, although overgrowth of *E. coli* may present a threat of fatal endogenous infections to older patients with immunosenescence and diminished gut barrier function because of its abundant antibiotic resistance genes and high tolerance to the inflammatory environment[46]. Future comprehensive examinations in

humanized gnotobiotic mice and human fecal transplantation may provide a better understanding of the relationship of ET-*E. coli* and healthy aging with regarding to its impacts on nutritional metabolites, host energy hemostasis, and immunosenescence.

## 5. Conclusion

Our data revealed an enterotype dominated by *E. coli* 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* might profoundly affect age-related changes in metabolism, diet 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. This has important implications for the prediction of aging and health loss, as well as investigation of the occurrence of mechanisms that can result in an aging-associated gut microbiome.

## Abbreviations

ET-*Bacteroides spp.*: enterotype dominated by *Bacteroides spp.*, ET-*P. copri*: enterotype dominated by *P. copri*, ET-*E. coli*: enterotype dominated by *E. coli*, *Faecalibacterium prausnitzii*: *F. prausnitzii*, T2DM: type 2 diabetes mellitus, PG cohort: Pinggu cohort, JSD: Jensen–Shannon divergence, PAM: partitioning around medoid, SPIEC-EASI: sparse inverse covariance estimation for ecological association inference, CH: Calinski–Harabasz, MRM: multiple regression on distance matrices, NESH: NetShift, RDA: redundancy analysis, FPG: fasting plasma glucose, HbA1c: hemoglobin A1c, hs-CRP: high-sensitivity C-reactive protein, *FDR*: false discovery rate, PCoA: principal coordinate analysis.

## Declarations

### Ethics approval and consent to participate

This research was reviewed and approved by the Clinical Research Ethics Committee of the First Affiliated Hospital, College of Medicine, Zhejiang University, under reference number 2019IIT(1276). Informed consent was provided by all volunteers themselves.

### Consent for publication

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.

### Availability of data and materials

Raw sequencing data of all metagenomes have been deposited at the European Nucleotide Archive under the accession numbers PRJEB48848.

## Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Funding

This work was supported by the National Key Research and Development Program of China (No. 2018YFC2000301), the National Natural Science Foundation of China (grant No. 82101665,81771498), the National Key Research and Development Program of China (2018YFC2000500), the Zhejiang Provincial Natural Science Foundation of China (LY22H03009), the National Clinical Key Specialty Construction Project of Geriatrics, and the Key Disciplines Construction Plan of Zhejiang Province Traditional Chinese Medicine (2017-XK-A31).

## Authors' contributions

All authors read and approved the final manuscript. Y-MY, QZ, and Y-QL conceived and designed the experiments and carried out data collection, analysis, and interpretation. J-YL, Q-FG, Y-CY, C-HH, J-JJ, Y-W, C-TP, and S-ML 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.

## Acknowledgments

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. We thank editors from Springer Nature Author Services (<https://www.springernature.com/gp/authors/research-data>), for editing the English text of a draft of this manuscript.

## References

1. O'Toole PW, Jeffery IB. Gut microbiota and aging. *Science*. 2015;350(6265):1214-5, doi: 10.1126/science.aac8469.
2. O'Toole PW, Jeffery IB. Microbiome-health interactions in older people. *Cell Mol Life Sci*. 2018;75(1):119-28, doi: 10.1007/s00018-017-2673-z.

3. Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S, et al. Gut microbiota composition correlates with diet and health in the elderly. *Nature*. 2012,488(7410):178-84, doi: 10.1038/nature11319.
4. Maynard C, Weinkove D. The Gut Microbiota and Ageing. *Subcell Biochem*. 2018,90:351-71, doi: 10.1007/978-981-13-2835-0\_12.
5. Li X, Stokholm J, Breyndrod A, Vestergaard GA, Russel J, Trivedi U, et al. The infant gut resistome associates with *E. coli*, environmental exposures, gut microbiome maturity, and asthma-associated bacterial composition. *Cell Host Microbe*. 2021,29(6):975-87 e4, doi: 10.1016/j.chom.2021.03.017.
6. Xiao L, Wang J, Zheng J, Li X, Zhao F. Deterministic transition of enterotypes shapes the infant gut microbiome at an early age. *Genome Biol*. 2021,22(1):243, doi: 10.1186/s13059-021-02463-3.
7. Zhong H, Penders J, Shi Z, Ren H, Cai K, Fang C, et al. Impact of early events and lifestyle on the gut microbiota and metabolic phenotypes in young school-age children. *Microbiome*. 2019,7(1):2, doi: 10.1186/s40168-018-0608-z.
8. Xu X, Zhang Z. Sex- and age-specific variation of gut microbiota in Brandt's voles. *PeerJ*. 2021,9:e11434, doi: 10.7717/peerj.11434.
9. Zhang X, Zhong H, Li Y, Shi Z, Ren H, Zhang Z, et al. Sex- and age-related trajectories of the adult human gut microbiota shared across populations of different ethnicities. *Nature Aging*. 2021,1(1):87-100, doi: 10.1038/s43587-020-00014-2.
10. Lu N, Hu Y, Zhu L, Yang X, Yin Y, Lei F, et al. DNA microarray analysis reveals that antibiotic resistance-gene diversity in human gut microbiota is age related. *Sci Rep*. 2014,4:4302, doi: 10.1038/srep04302.
11. He Y, Wu W, Zheng HM, Li P, McDonald D, Sheng HF, et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat Med*. 2018,24(10):1532-5, doi: 10.1038/s41591-018-0164-x.
12. Kuntal BK, Chandrakar P, Sadhu S, Mande SS. 'NetShift': a methodology for understanding 'driver microbes' from healthy and disease microbiome datasets. *ISME J*. 2019,13(2):442-54, doi: 10.1038/s41396-018-0291-x.
13. Shi Y, Zhang K, Li Q, Liu X, He JS, Chu H. Interannual climate variability and altered precipitation influence the soil microbial community structure in a Tibetan Plateau grassland. *Sci Total Environ*. 2020,714:136794, doi: 10.1016/j.scitotenv.2020.136794.
14. Herren CM, McMahon KD. Cohesion: a method for quantifying the connectivity of microbial communities. *ISME J*. 2017,11(11):2426-38, doi: 10.1038/ismej.2017.91.
15. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*. 2012,490(7418):55-60, doi: 10.1038/nature11450.
16. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science*. 2011,334(6052):105-8, doi: 10.1126/science.1208344.

17. Jackson MA, Jeffery IB, Beaumont M, Bell JT, Clark AG, Ley RE, et al. Signatures of early frailty in the gut microbiota. *Genome Med.* 2016,8(1):8, doi: 10.1186/s13073-016-0262-7.
18. Mueller S, Saunier K, Hanisch C, Norin E, Alm L, Midtvedt T, et al. Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study. *Appl Environ Microbiol.* 2006,72(2):1027-33, doi: 10.1128/AEM.72.2.1027-1033.2006.
19. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods.* 2015,12(10):902-3, doi: 10.1038/nmeth.3589.
20. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods.* 2018,15(11):962-8, doi: 10.1038/s41592-018-0176-y.
21. Wibowo MC, Yang Z, Borry M, Hubner A, Huang KD, Tierney BT, et al. Reconstruction of ancient microbial genomes from the human gut. *Nature.* 2021,594(7862):234-9, doi: 10.1038/s41586-021-03532-0.
22. Roth-Schulze AJ, Penno MAS, Ngui KM, Oakey H, Bandala-Sanchez E, Smith AD, et al. Type 1 diabetes in pregnancy is associated with distinct changes in the composition and function of the gut microbiome. *Microbiome.* 2021,9(1):167, doi: 10.1186/s40168-021-01104-y.
23. Chen L, Collij V, Jaeger M, van den Munckhof ICL, Vich Vila A, Kurilshikov A, et al. Gut microbial co-abundance networks show specificity in inflammatory bowel disease and obesity. *Nat Commun.* 2020,11(1):4018, doi: 10.1038/s41467-020-17840-y.
24. Li D, Ni H, Jiao S, Lu Y, Zhou J, Sun B, et al. Coexistence patterns of soil methanogens are closely tied to methane generation and community assembly in rice paddies. *Microbiome.* 2021,9(1):20, doi: 10.1186/s40168-020-00978-8.
25. Zheng J, Zhou Z, Wei Y, Chen T, Feng W, Chen H. High-throughput profiling of seasonal variations of antibiotic resistance gene transport in a peri-urban river. *Environ Int.* 2018,114:87-94, doi: 10.1016/j.envint.2018.02.039.
26. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. *Nature.* 2011,473(7346):174-80, doi: 10.1038/nature09944.
27. Jiao S, Yang Y, Xu Y, Zhang J, Lu Y. Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *ISME J.* 2020,14(1):202-16, doi: 10.1038/s41396-019-0522-9.
28. Hu A, Wang J, Sun H, Niu B, Si G, Wang J, et al. Mountain biodiversity and ecosystem functions: interplay between geology and contemporary environments. *ISME J.* 2020,14(4):931-44, doi: 10.1038/s41396-019-0574-x.
29. Zengler K, Zaramela LS. The social network of microorganisms - how auxotrophies shape complex communities. *Nat Rev Microbiol.* 2018,16(6):383-90, doi: 10.1038/s41579-018-0004-5.
30. Kong F, Deng F, Li Y, Zhao J. Identification of gut microbiome signatures associated with longevity provides a promising modulation target for healthy aging. *Gut Microbes.* 2019,10(2):210-5, doi: 10.1080/19490976.2018.1494102.

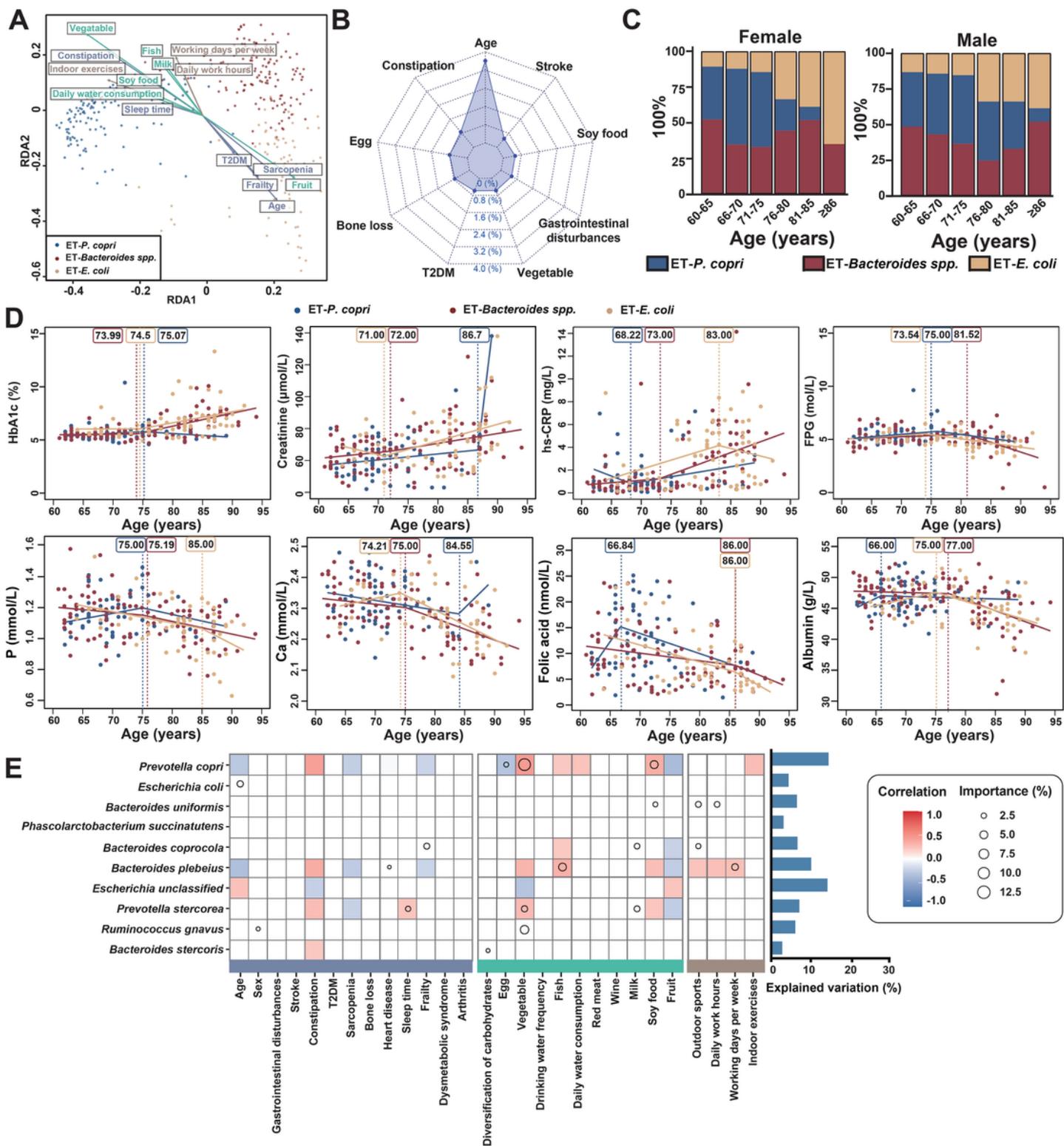
31. Park SH, Kim KA, Ahn YT, Jeong JJ, Huh CS, Kim DH. Comparative analysis of gut microbiota in elderly people of urbanized towns and longevity villages. *BMC Microbiol.* 2015,15:49, doi: 10.1186/s12866-015-0386-8.
32. Martinson JNV, Pinkham NV, Peters GW, Cho H, Heng J, Rauch M, et al. Rethinking gut microbiome residency and the Enterobacteriaceae in healthy human adults. *ISME J.* 2019,13(9):2306-18, doi: 10.1038/s41396-019-0435-7.
33. Jackson MA, Verdi S, Maxan ME, Shin CM, Zierer J, Bowyer RCE, et al. Gut microbiota associations with common diseases and prescription medications in a population-based cohort. *Nat Commun.* 2018,9(1):2655, doi: 10.1038/s41467-018-05184-7.
34. He Q, Jiang J, Xie L, Zhang L, Yang M. A sarcopenia index based on serum creatinine and cystatin C cannot accurately detect either low muscle mass or sarcopenia in urban community-dwelling older people. *Sci Rep.* 2018,8(1):11534, doi: 10.1038/s41598-018-29808-6.
35. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, et al. Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. *Cell Host Microbe.* 2007,2(3):204, doi: 10.1016/j.chom.2007.08.002.
36. Qi H, Wei J, Gao Y, Yang Y, Li Y, Zhu H, et al. Reg4 and complement factor D prevent the overgrowth of *E. coli* in the mouse gut. *Commun Biol.* 2020,3(1):483, doi: 10.1038/s42003-020-01219-2.
37. Baumgartner M, Bayer F, Pfrunder-Cardozo KR, Buckling A, Hall AR. Resident microbial communities inhibit growth and antibiotic-resistance evolution of *Escherichia coli* in human gut microbiome samples. *PLoS Biol.* 2020,18(4):e3000465, doi: 10.1371/journal.pbio.3000465.
38. Rottjers L, Faust K. From hairballs to hypotheses-biological insights from microbial networks. *FEMS Microbiol Rev.* 2018,42(6):761-80, doi: 10.1093/femsre/fuy030.
39. Diard M, Garcia V, Maier L, Remus-Emsermann MN, Regoes RR, Ackermann M, et al. Stabilization of cooperative virulence by the expression of an avirulent phenotype. *Nature.* 2013,494(7437):353-6, doi: 10.1038/nature11913.
40. Ghoul M, Mitri S. The Ecology and Evolution of Microbial Competition. *Trends Microbiol.* 2016,24(10):833-45, doi: 10.1016/j.tim.2016.06.011.
41. Whitmore SE, Lamont RJ. Oral bacteria and cancer. *PLoS Pathog.* 2014,10(3):e1003933, doi: 10.1371/journal.ppat.1003933.
42. Mottawea W, Chiang CK, Muhlbauer M, Starr AE, Butcher J, Abujamel T, et al. Altered intestinal microbiota-host mitochondria crosstalk in new onset Crohn's disease. *Nat Commun.* 2016,7:13419, doi: 10.1038/ncomms13419.
43. Yachida S, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, et al. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat Med.* 2019,25(6):968-76, doi: 10.1038/s41591-019-0458-7.
44. Eisenberg T, Abdellatif M, Schroeder S, Primessnig U, Stekovic S, Pendl T, et al. Cardioprotection and lifespan extension by the natural polyamine spermidine. *Nat Med.* 2016,22(12):1428-38, doi: 10.1038/nm.4222.



## Figure 1

### Stratification of older persons into three enterotypes based on their gut microbiome.

**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-*P. copri*, ET-*Bacteroides spp.* and ET-*E. coli* represent each enterotype. **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 most responsible bacteria 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 as the main contributors 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-*P. copri*, ET-*Bacteroides spp.* and ET-*E. coli*, respectively. **D-E:** The bar graph indicates the comparison of microbial taxa at the order (D) and family levels (E) in each enterotype. **F:** Composition of species-level proportions for all 367 subjects across three enterotypes determined using shotgun metagenomic sequencing.

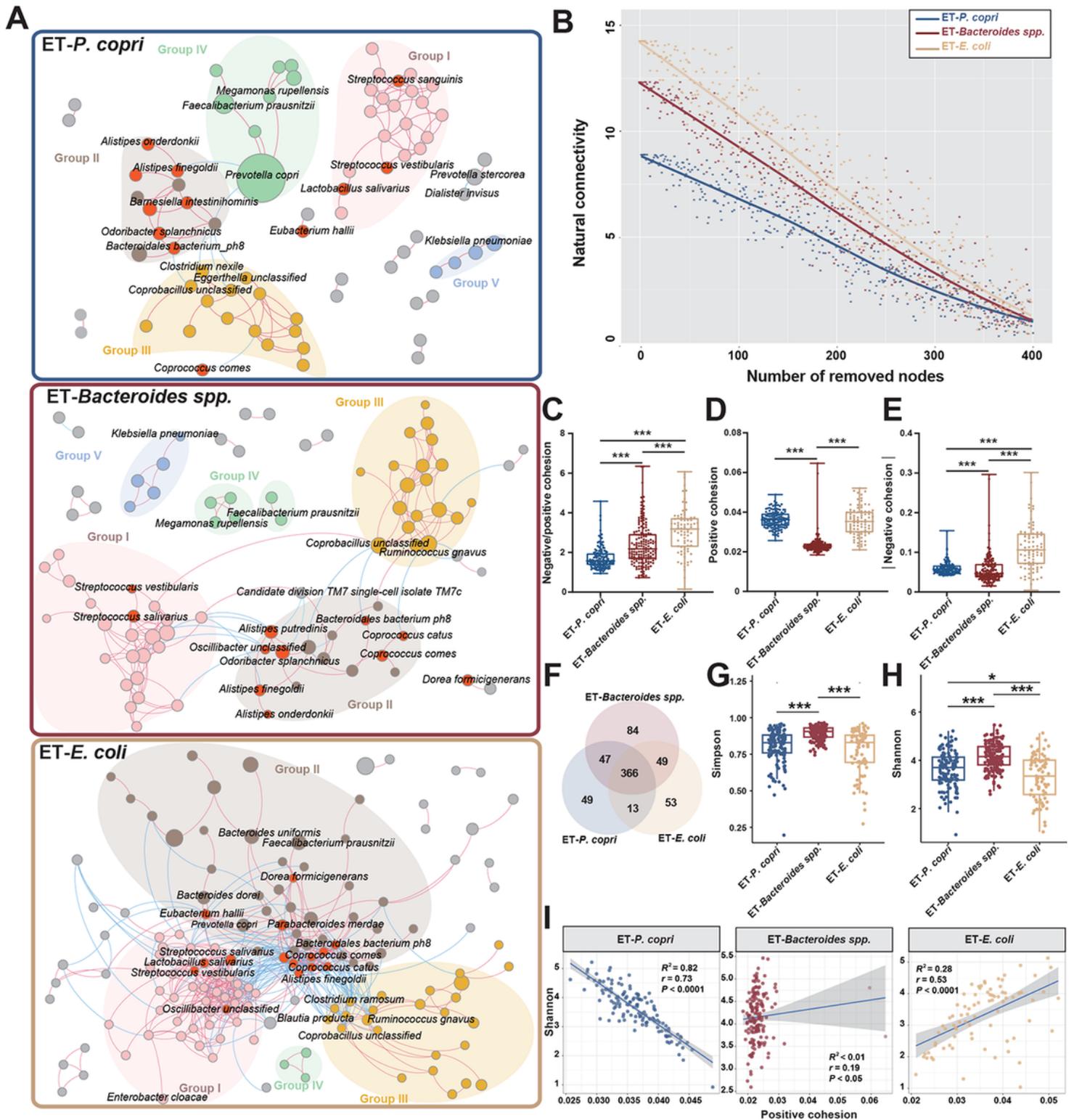


**Figure 2**

**Associations of enterotypes with age and other external factors.**

**A:** Redundancy analysis (RDA) of the clinical characteristics (blue), dietary habits (green), and participation in physical work or having a habit of exercising (brown) in the gut microbiome of the elderly individuals. Each point represents a sample, and the colors represent the enterotypes. **B:** 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). **C:** Bar plots indicate the percentage of individuals with each enterotype in each age range. **D:** Age-related trajectories of eight representative blood inflammatory and nutritional markers in different enterotypes revealed by a piecewise linear regression model with age breakpoint estimation (the dashed lines), including HbA1c, creatinine, hs-CRP, FPG, P, Ca, folic acid, and albumin. **E:** Contributions of clinical and demographic variables to the differences in relative abundances of key distinguishing bacterial species for three enterotypes based on Spearman correlation coefficients and best multiple regression model. The color of the bar 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 (%).

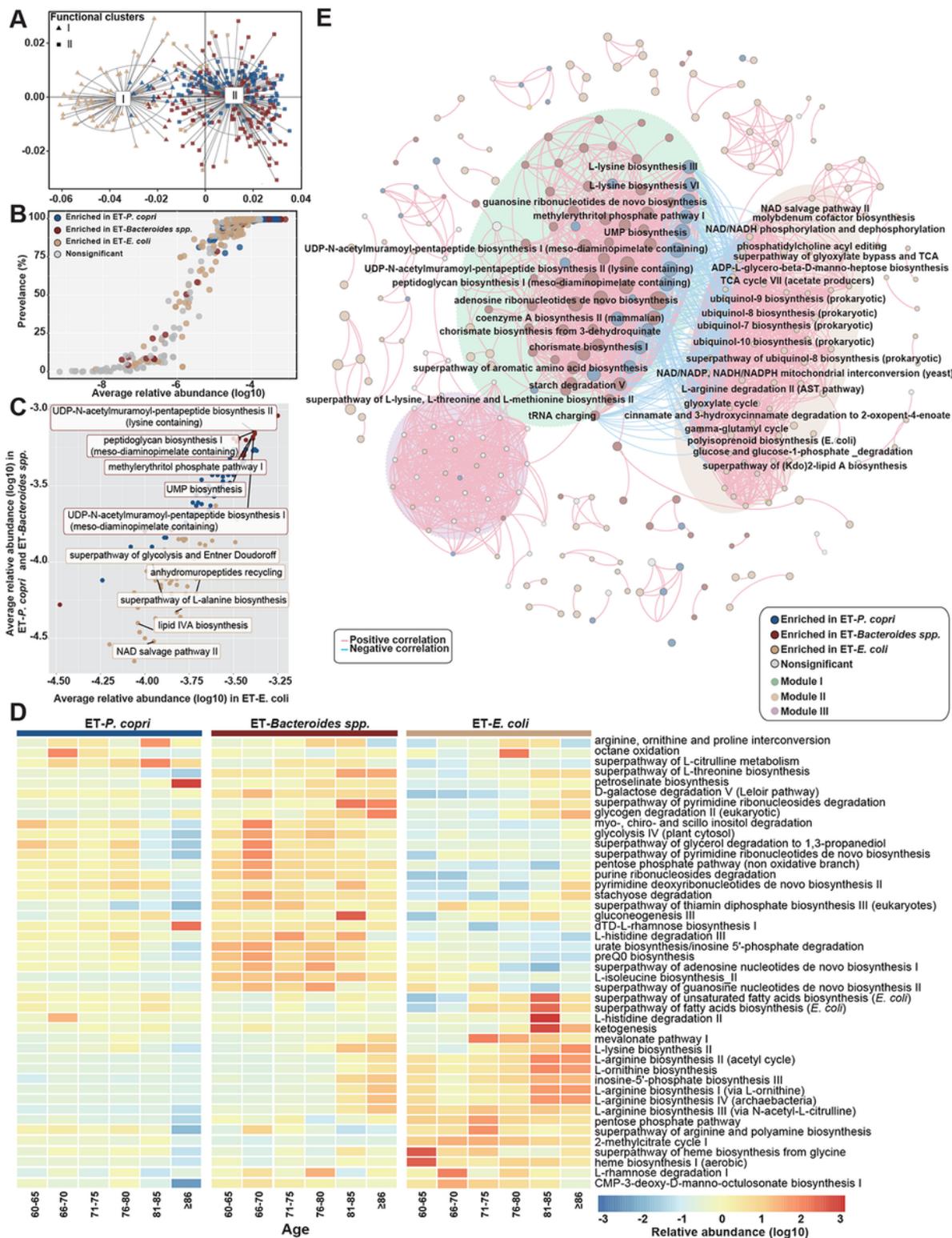


**Figure 3**

**Discrepancies of 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 the separation of the networks (NESH score value  $> 2$ ). **B**: The highest natural connectivity was observed for the microbial co-occurrence network in ET-*E. coli*. **C**: The highest negative:positive cohesion ratio in the co-occurrence network was observed for the gut microbial community dominated by *E. coli*. **D**: The lowest positive cohesion in the microbial co-occurrence network of ET-*Bacteroides spp.* **E**: Significant difference in the absolute value of negative cohesion across the three enterotypes. **F**: Venn diagram showing the number of shared and unique species in each enterotype. **G-H**: Species profile-based  $\alpha$ -diversity of the gut microbiome. The diversities were measured based on the species profiles for Shannon (**H**) and Simpson (**I**) indexes.  $P$  values are derived from the Wilcoxon rank-sum test: \* $P < 0.05$ , \*\* $P < 0.01$ , and \*\*\* $P < 0.001$ . **I**: The correlations between  $\alpha$ -diversity and the absolute value of negative cohesion in each enterotype constructed by Spearman's correlation analysis, box plots show inner quartiles and median values.



**Figure 4**

**Dramatic variation in functional profile in 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). The point color represents the enterotype (red, ET-*Bacteroides* spp., gold, ET-*E. coli*, blue, ET-*P. copri*), and the shape represents the functional cluster

(triangle, cluster I, square, cluster II). **B:** Functional differences identified by Wilcoxon rank-sum test are based on the metagenome sequence data. Samples are colored to match their enterotype assignments. **C:** One hundred twenty-nine 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. **E:** The descriptions of the three modules are as follows. Module I: Core functional pathways that are significantly enriched in ET-*P. copri* or ET-*Bacteroides spp.* Module II: ET-*E. coli*-associated functional pathways. Module III: Scarce functional pathways that are present only in the minority with low relative abundance. The size of each node is proportional to the mean relative abundance.

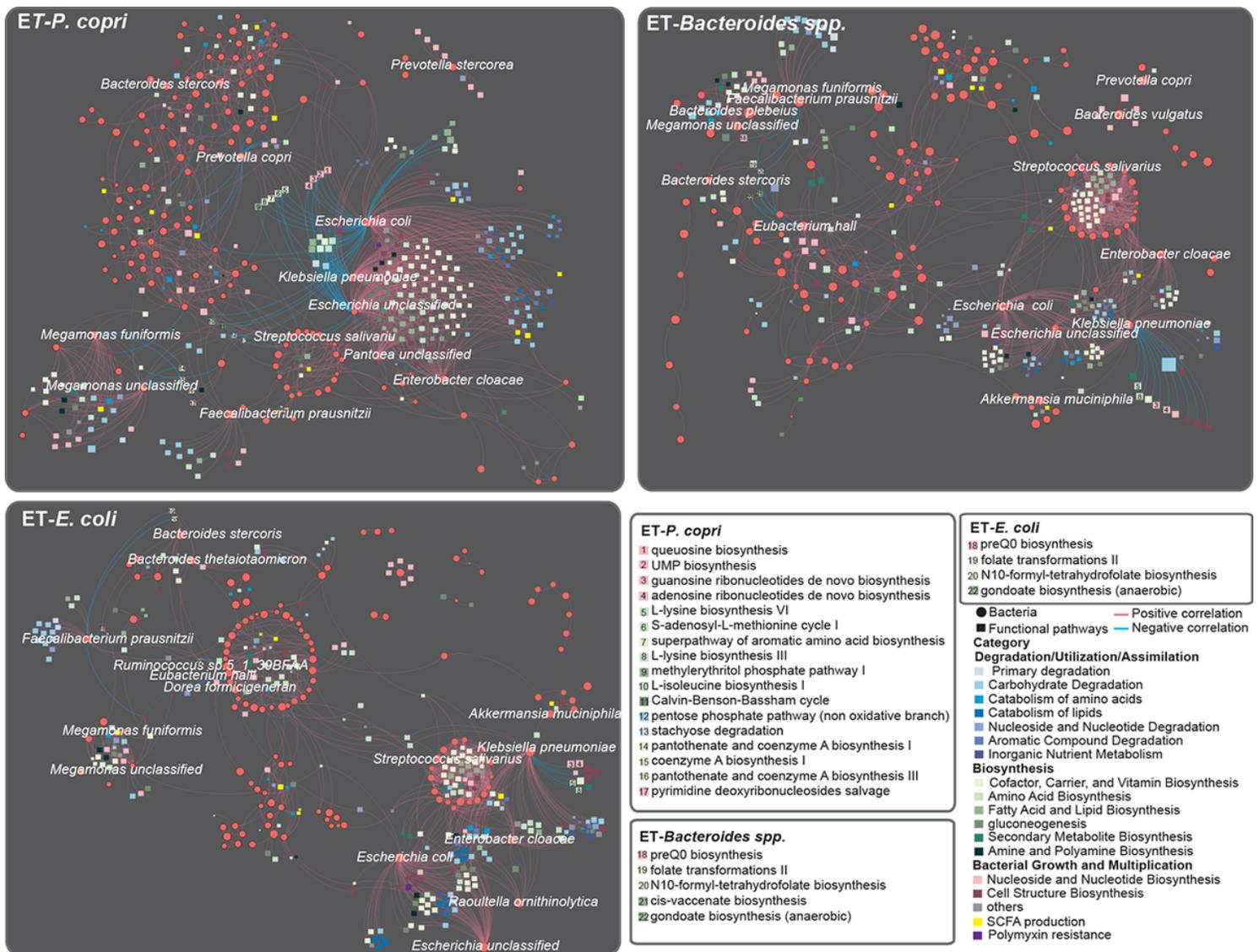
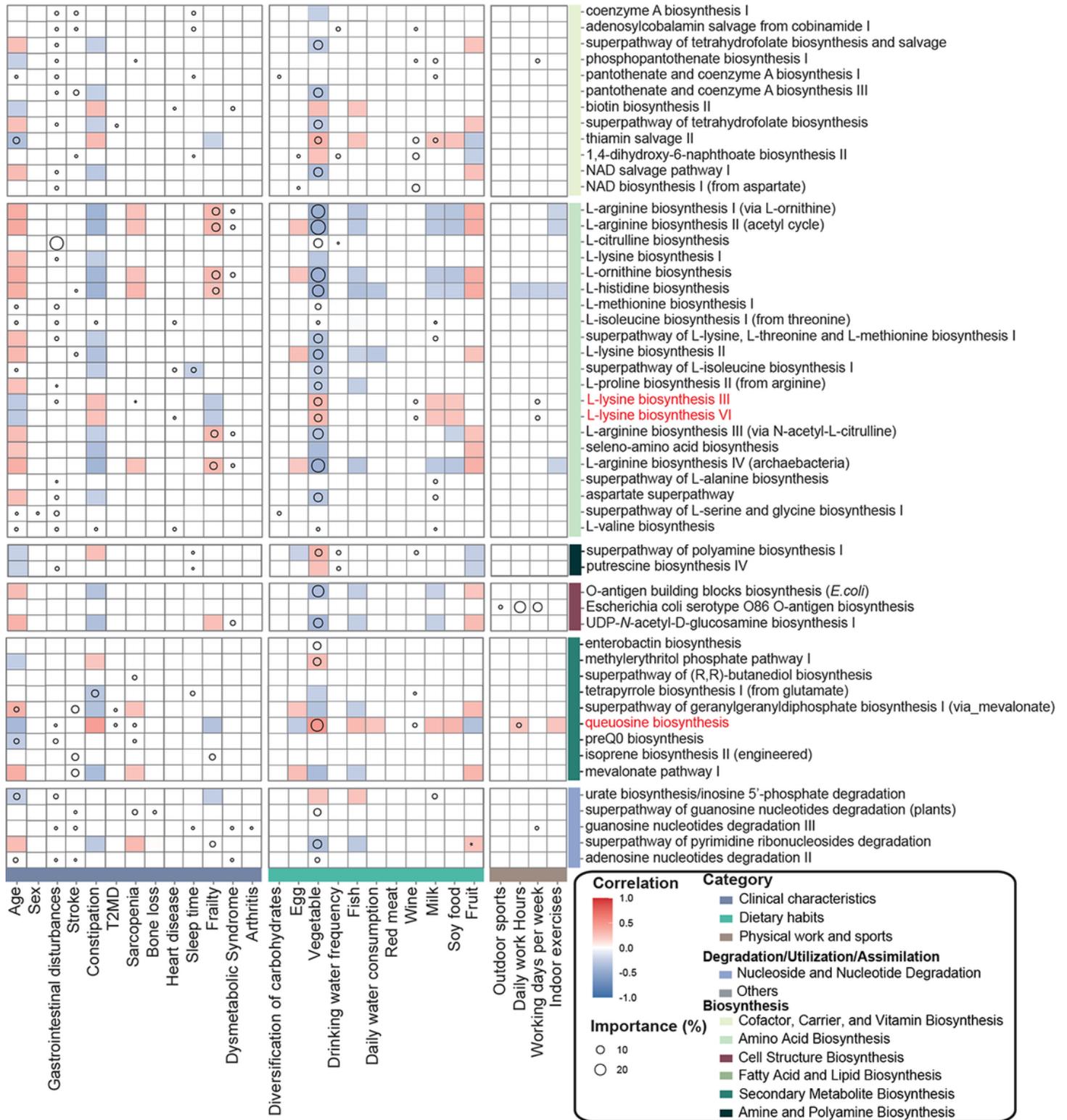


Figure 5

**Functional co-abundance networks of species and functional pathways in each enterotype.** Species and functional pathway co-abundance networks in each enterotype constructed by Spearman's correlation analysis (correlation values >0.3). Red lines indicate a positive correlation, and blue lines indicate a negative correlation. Each dot indicates a species, and each square indicates a functional pathway. The size of each node is proportional to the average relative abundance in each enterotype. The squares were colored according to functional categories.



## Figure 6

**Microbial metabolic pathways associated with the health status, dietary factors, and lifestyle factors in older people.** Rows correspond to bacterial taxa, 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  $*P < 0.05$ ). Circle size represents the variable importance.

## Supplementary Files

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

- [Additionalfile1.docx](#)
- [Additionalfile2.xls](#)
- [Additionalfile3.xls](#)
- [Additionalfile4.xls](#)
- [Additionalfile5.xls](#)