

Bacterial community and composition of soil in world's longevity township Jiaoling, China by 16S rRNA high-throughput sequencing

Lei Wu

Shaanxi University of Science and Technology

Xinqiang Xie

Guangdong Institute of Microbiology

Jun Ma

Shaanxi University of Science and Technology

Jumei Zhang

Guangdong Institute of Microbiology

Yu Ding

Jinan University

Qingping Wu (✉ wuqp203@163.com)

Guangdong Institute of Microbiology

Research Article

Keywords: soil microbiota, bacterial community, longevity, longevity township, high-throughput sequencing

Posted Date: March 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-290155/v1>

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

[Read Full License](#)

Abstract

Healthy longevity is a complicated process, however, the underlying mechanisms between longevity and microbiota warrant investigation. To address this, we characterized a longevity trajectory of environmental microbiota in a longevity township. We used high-throughput sequencing of the 16S rRNA gene to analyse the composition and function of soil microbiota. The composition and diversity of soil microbiota significantly differed between towns. The dominant bacteria at the phylum level included *Proteobacteria*, *Firmicutes*, and *Acidobacteria*. At the genus level, *Chujaibacter*, *Acidipila*, and *Lactobacillus* were dominant. However, *Steroidobacter*, *Comamonas*, and *Pseudoxanthomonas* were only dominant in Xinpu with high centenarian population. Twelve biomarkers were responsible for significant differences between towns, including *Lactobacillus*, *Muribaculaceae*, *Ruminococcaceae*, *Lachnospiraceae*, and *Chitinophagales*, etc. The main species contributing to the differences of towns were *Chujaibacter*, *Acidipila*, *Lactobacillus*, *Rhodanobacter*, *Lysobacter*, *Bryobacter*, *Granulicella*, *Flavobacterium*, and *Mizugakiibacter*. The function of exosome, cysteine and methionine metabolism, amino acid-related enzymes, peptidases, starch and sucrose metabolism, etc., were predicted. Thus, we have revealed significant differences in the composition and diversity of soil microbiota in the world's longevity township, the relationship between soil microbiota and long-lived people. These findings provide a research foundation for the role of soil microbiota in healthy longevity.

Introduction

Throughout the ages, humans have explored the mystery of health and longevity. Longevity is a complex feature. It is related to many factors, such as geography and meteorological, economic, medical, and health conditions. It is also affected by genetic, environmental, and other factors, and ideally would mainly depend on the rate of ageing.

In the biomedical field, the research has focused on searching for molecular and biological factors that promote healthy ageing and longevity. Considerable attention has been paid to analyzing the role of genetic factors in determining healthy ageing and longevity. The main factors that affect longevity and ageing are growth hormone (GH) and insulin/insulin-like growth factor (insulin/IGF-1) pathways¹ in various organisms^{2,3}. Further, overexpression of the Forkhead box O3 gene (*FOXO3*) in model organisms is related to prolonged lifespan⁴⁻⁶. Another protein beneficial for longevity and metabolic regulation is AMP-activated protein kinase (AMPK)⁷. Besides, deacetylase family genes (*Sirtuins*)^{8,9}, the apolipoprotein E gene (*APOE*)¹⁰, Telomerase¹¹, the mammalian target of rapamycin (mTOR) signalling pathway^{12,13}, the tumour suppressor gene *P53*¹⁴, the transcription factor NF- κ B^{15,16}, the autophagy-lysosomal signalling pathway¹⁷⁻¹⁹, long-chain non-coding RNAs^{20,21} play important roles between health and longevity. Finally, methionine sulfoxide is considered to be a marker of biological ageing²². Methionine sulfoxide reductase is a specific antioxidant enzyme that removes this modification of proteins, and at the same time, acts as a general cellular antioxidant to scavenge free radicals and protect the cell from biological oxidative stress²³.

The environment also plays a very important role in determining longevity, as people are sensitive to the environment^{24,25}. For example, dietary restriction, i.e., food intake control to achieve balanced nutrition²⁶, can delay the rate of ageing. Further, appropriate exercise strengthens the body, enhances immunity resistance, and delays ageing. Age positively correlates with free radical levels in humans. Moderate exercise increases the levels of free radical-scavenging enzymes and their activity to delay ageing²⁷. Among the above factors, the natural environment is key to longevity. For human beings to live a long and healthy life, they must establish and maintain a harmonious relationship with the natural environment and living environment. Topography and landforms, climate, soil, water, and other natural geographical environmental factors are the main determinants of longevity²⁸. Soil is the material basis on which organisms rely for survival. The most essential trace elements from the soil affect water quality, plant, and human health through the food chain^{29,30}.

The rapid development of science and technology in the 21st century has enabled researchers to explore the relationship between health and longevity, and the environment. Metagenomics with high-throughput sequencing, combined with multi-omics technologies, is used to analyse the gut microbiota of the elderly, and the correlation and contribution of microorganisms to health and longevity from the perspective of the microbiota. The microbiota co-evolve with human and are vital to human health. However, changes in environmental microbiota composition and function in longevity areas have not been fully studied. To address this question and improve the understanding of whether and how the environment support health and longevity, it is necessary to analyse the environmental microbiota in such areas. To investigate the association between the soil microbiota and longevity in counties of China, we studied Jiaoling, one of the world's longevity townships. The obtained data will help identify commonalities and differences between regional longevity factors; enable correlation analyses between the soil microbiota, health and longevity; and provide a theoretical basis for guiding future research on health and longevity.

Results

Characteristics of the long-lived population and sample sequencing

Jiaoling County is affiliated to Meizhou City, Guangdong Province. It is located in the northeast of Guangdong Province. In eight towns in the Jiaoling County (T1, Guangfu; T2, Nanzhai; T3, Wenfu; T4, Changtan; T5, Jiaocheng; T6, Lanfang; T7, Sanzhen; and T8, Xinpu), the proportion of the population over the age of 80 in the total population over the age of 80 (approximately 8365 people) is 6.91%, 10.57%, 9.07%, 9.12%, 22.30%, 11.73%, 8.67%, and 21.64%, respectively. The elderly population is the largest in Jiaocheng, and the smallest in Guangfu (Fig. 1). According to the proportion of the population aged 80–89 years old, 90–99 years old, and over 100 years old to the total population over-80, Xinpu with the highest proportion of centenarians, followed by Changtan, and the lowest in Nanzhai (Fig. 1). The proportion of centenarians shows a decreasing distribution trend along the Shiku River, from south to north.

Soil samples were collected in the eight towns, bacterial genomic DNA isolated, and the 16S rRNA gene sequenced using the Ion S5™XL sequencing platform (by single-end sequencing), generating small-fragment libraries. By shearing and filtering the reads, an average of 83,564 reads per sample was obtained, with an average of 78,271 valid reads after quality control. The efficiency of quality control was 93.74%. The sequences were then clustered into OTUs, at 97% identity, yielding 10,528 OTUs, which were finally annotated using the Silva132 database.

Composition and structure of the soil microbiota

Based on the OTU analysis and clustering, shared and unique OTUs were identified in different groups. Overall, 1827 OTUs were shared in the T1–T8 samples, and the number of unique OTUs in each group is 58, 96, 422, 351, 129, 401, 51, 153, respectively (Supplementary Figure S1A). Based on species annotation, the top 10 most abundant bacteria at each classification level (phylum, class, order, family, and genus) were identified in each group, and a cumulative bar map of species relative abundance was generated to allow visual assessment of the relative bacterial abundance and their proportions at the different classification levels in each sample. The relative abundance at phylum levels is presented in a histogram (Fig. 2A). The major abundant phyla were *Proteobacteria*, *Firmicutes*, *Acidobacteria*, *Bacteroidetes*, *Actinobacteria*, *Chloroflexi*, *Rokubacteria*, *Planctomycetes*, and *Gemmatimonadetes*. Among these, bacteria from *Proteobacteria*, *Acidobacteria*, and *Bacteroidetes* were dominant at all taxonomic levels.

Based on species annotation and abundance at the family level, 35 most abundant families were identified and clustered at species and sample level, to generate a heat map (Supplementary Figure S1B). The top 10 families in the samples were *Rhodanobacteraceae*, unidentified *Acidobacteriales*, *Lachnospiraceae*, *Burkholderiaceae*, *Lactobacillaceae*, *Pyrinomonadaceae*, *Prevotellaceae*, *Xanthomonadaceae*, *Ruminococcaceae*, and *Muribaculaceae*, with different families being dominant in each group. *Rikenellaceae*, *Ruminococcaceae*, *Bacteroidaceae*, *Muribaculaceae*, *Lactobacillaceae*, *Lachnospiraceae*, *Erysipelotrichaceae*, and *Streptococcaceae* were dominant in the T2 Nanzhai with the lowest proportion of centenarians. *Burkholderiaceae* and *Streptococcaceae* were dominant in the T8 Xinpu with the highest proportion of centenarians.

Based on species annotation and abundance at the genus level, 35 most abundant genera were identified, clustered at the species and sample levels, and a heat map was generated (Fig. 2B). The top 10 genera in the samples were *Chujaibacter*, *Acidipila*, *Lactobacillus*, *Rhodanobacter*, unidentified *Prevotellaceae*, *Lysobacter*, *Bryobacter*, *Granulicella*, *Flavobacterium*, and *Mizugakiibacter*. The dominant genera in the T2 Nanzhai with the lowest proportion of centenarians were unidentified *Prevotellaceae*, *Streptococcus*, *Bacteroides*, unidentified *Ruminococcaceae*, *Lactobacillus*, unidentified *Lachnospiraceae*, unidentified *Burkholderiaceae*, *Ralstonia*, *Collimonas*, *Candidatus Solibacter*, and *Candidatus Koribacter*. The dominant genera in the T8 Xinpu with the highest proportion of centenarians were *Steroidobacter*, *Comamonas*, and *Pseudoxanthomonas*.

To further analyse the phylogenetic relationship of species at the genus level, representative sequences of the top 100 genera were obtained using multiple sequence alignment. A phylogenetic tree constructed using representative sequences of species at the genus level is shown in Fig. 2C. The analysis revealed that *Lactobacillus*, unidentified *Acidobacteria*, *Flavobacterium*, *Lysobacter*, *Chujaibacter*, *Acidibacter*, unidentified *Gammaproteobacteria*, and *Bryobacter* were distributed in the eight groups. For each classification result, species representing the top 10 genera with the highest relative abundance were selected for specific species classification tree analysis (Fig. 2D). The analysis revealed that *Lactobacillus*, a species of special interest, accounted for 1.332% of all species and 15.38% of the selected species. The percentages of *Lactobacillus animalis*, *Lactobacillus delbrueckii*, *Lactobacillus gasseri*, *Lactobacillus mucosae*, and *Lactobacillus reuteri* among the selected species were 7.44%, 0.01%, 4.27%, 0.03%, and 2.76%, respectively.

Alpha diversity and beta diversity analyses

Species diversity curves (dilution curves and hierarchical clustering curves) and species accumulation box plots were used to evaluate the differences in species richness and diversity of microbial communities in each sample. The rarefaction curve (Fig. 3A) tended to be flat, indicating that the sequencing data was sufficient for the analysis. Additional data would generate only a small number of new species (OTUs), and indirectly reflect the abundance of species in the sample. For the rank abundance curve (Fig. 3B) in the horizontal dimension, the span of the curve gradually increased, indicating a higher richness of species; in the vertical dimension, the curve gradually flattened, indicating a more uniform species distribution. The box plot of species accumulation (Fig. 3C), the sample size increased, and the box plot position tended to be flat. This meant that the species richness in the environment would not significantly increase with increasing sample size, indicating that the sample size was sufficient for data analysis. The alpha diversity, assessed by the Chao1 index, was significantly different between the T2 group and the T1, T3, T4, and T5 groups ($p < 0.05$) (Fig. 3D). Wilcoxon rank-sum analysis to test the alpha diversity between any two groups revealed that the alpha diversity in the T2 group was significantly lower than that in the other groups.

Beta diversity is a comparative analysis of microbial community composition in different samples. Beta diversity analysis based on the weighted Unifrac beta-Wilcox rank-sum test index (Fig. 3E) revealed statistically significant differences between the T2 group and the T3, T4, T5, and T6 groups ($p < 0.05$). Significant differences between the T5 group and the T7 and T8 groups were also apparent ($p < 0.05$). To analyse the similarity between samples, the weighted Unifrac distance matrix was used for UPGMA cluster analysis, and the clustering results were integrated and displayed with the relative abundance of each sample at the phylum level (Fig. 3F). The relatively small distance between the T6 and T8 groups, and the T3 and T4 groups, indicated a similar species composition structure. Therefore, samples with similar community structures tended to cluster together, while samples with very different communities were further apart. The T2 and T8 groups had the farthest distance and the largest community difference. The communities in each group exhibit certain inter-group differences, indicating the accuracy of grouping. In the beta diversity analysis, the weighted Unifrac distance index was used to measure the coefficient of difference between the two samples. The smaller the coefficient value, the smaller the

difference in species diversity between the two samples. The distance matrix heat map constructed based on the weighted Unifrac distance (Fig. 3G) revealed a big difference between the T2 group and other groups.

Significance analysis of species differences between groups

Alpha diversity index analysis of differences between groups is used to determine whether the overall community structure in different groups is significantly different. The species responsible for this difference are then identified using LEfSe and ternary plot analysis. The identified species are a group biomarker. Accordingly, we used LEfSe analysis to detect the species differentiating different groups by the rank-sum test of species abundance. A histogram of the different species LDA value distribution is shown in Fig. 4A, and an evolutionary branch diagram of the different species is shown in Fig. 4B. The former revealed biomarker species with LDA score greater than 4. The LEfSe analysis revealed 12 biomarkers, namely, *Lactobacillus*, *Muribaculaceae*, *Ruminococcaceae*, *Lachnospiraceae*, *Chitinophagales*, and others (Fig. 4C–G).

Next, to identify the differences in the dominant species in three groups of samples at each classification level (phylum, class, order, family, genus, and species), the top 10 species with the average abundance in the three groups of samples at the genus level were used to generate a ternary plot (ternary phase diagram) (Supplementary Figure S2). Difference analysis of the dominant species in the T1–T3 samples revealed that *Lactobacillus* was dominant in the T2 group, and *Chujaibacter* and *Granulicella* were dominant in the T1 group (Supplementary Figure S2A). Further, for the T2–T4 samples, *Lactobacillus* was dominant in the T2 group; and *Chujaibacter*, *Acidipila*, *Mizugakiibacter*, and *Rhodanobacter* were dominant in the T4 group (Supplementary Figure S2B). For the T3–T5 samples, unidentified *Prevotellaceae* was dominant in the T3 group, and *Chujaibacter* and *Acidipila* were dominant in the T4 group (Supplementary Figure S2C). For the T4–T6 samples, *Chujaibacter* and *Acidipila* were dominant in the T4 group, and *Lactobacillus* and unidentified *Prevotellaceae* were dominant in the T6 group (Supplementary Figure S2D). For the T5–T7 samples, *Chujaibacter* was dominant in the T7 group, and *Lactobacillus* and unidentified *Prevotellaceae* were dominant in the T6 group (Supplementary Figure S2E). Finally, for the T6–T8 samples, *Flavobacterium* was dominant in the T8 group, and *Chujaibacter* was dominant in the T7 group (Supplementary Figure S2F). Overall, each group had a different dominant species. This species may constitute the characteristic of each group and should be analysed further.

Next, to determine individual species contribution to the above differences, Similarity percentage (Simpser) analysis was performed. Simper is a decomposition of the Bray–Curtis difference index, used to quantify each species contribution to the difference between two groups. The analysis revealed the top 10 species and their abundances, contributing to the difference between the two groups (Supplementary Figure S3). The top 10 species contributing to the difference between two groups and their abundance rankings, from high to low, were: *Chujaibacter*, *Acidipila*, *Lactobacillus*, *Rhodanobacter*, unidentified *Prevotellaceae*, *Lysobacter*, *Bryobacter*, *Granulicella*, *Flavobacterium*, and *Mizugakiibacter*. The contribution of the species to different groups differed, with the highest contribution from *Chujaibacter*, followed by *Acidipila* and *Lactobacillus*.

Co-occurrence network of the soil microbiota

The species co-occurrence network diagram enables intuitive visualisation of the impact of different environmental factors on microbial adaptability and the dominant species and closely interacting populations of species that occupy a dominant position in a specific environment. These dominant species and populations often play a unique and important role in maintaining a stable structure and function of the environmental microbial community. The analysis of the significance of the differences between groups and their contribution revealed that *Lactobacillus* was important for differentiating the groups. Therefore, the co-occurrence network of *Lactobacillus* was next analysed. In the soil microbiota co-occurrence network (Fig. 5), the following were positively correlated with *Lactobacillus*: *Anaerostipes*, *Blautia*, *Dialister*, *Rhodanobacter*, *Pajaroellobacter*, *Sporichthya*, *Edaphobacter*, *Holdemanella*, *Alkanibacter*, unidentified *Acidobacteriales*, *Dorea*, unidentified *Ruminococcaceae*, *Holophaga*, *Roseburia*, unidentified *Cyanobacteria*, *Acidothermus*, *Nocardia*, *Streptococcus*, *Oribacterium*, and *Jatrophihabitans*. The following were negatively correlated with *Lactobacillus*: *Minicystis*, *Rhizorhapis*, *Phycococcus*, *Oligoflexus*, *Gaiella*, *Myxococcus*, *Ensifer*, *Illumatobacter*, *Rhodopirellula*, *Altererythrobacter*, unidentified *Deltaproteobacteria*, *Stenotrophobacter*, *Chiayiivirga*, candidatus *Alysiosphaera*, *Arenimonas*, *Sphingobium*, *Pseudomonas*, unidentified *Dehalococcoidia*, *Fimbriiglobus*, *Amaricoccus*, *Nakamurella*, *Methylobacillus*, *Parafrigoribacterium*, *Solitalea*, *Adhaeribacter*, *Methyloceanibacter*, unidentified *Planctomycetales*, *Azovibrio*, *Gemmata*, unidentified candidatus *Woesebacteria*, unidentified *Gammaproteobacteria*, *Lysobacter*, *Defluviicoccus*, *Nannocystis*, and *Chryseolinea*. Hence, in addition to exerting a probiotic effect, *Lactobacillus* could also exert a synergistic effect via positive and negative regulation of the strains mentioned above.

Predicted functions of the soil microbiota

Using the annotation data (Supplementary Figure S4B), the top 10 functional annotations were retrieved for the most abundant bacteria at each annotation level, and a histogram of relative functional abundance was generated to visualise the relatively most abundant features at different annotation levels (Supplementary Figure S4A). The identified KEGG pathway annotations mainly focused on cellular processes, environmental information processing, genetic information processing, human diseases, metabolism, organic systems. Among them, metabolism had the highest proportion, and organic systems had the lowest proportion. Based on the functional annotation and sample abundance information, the top 35 abundance features and their abundances in each sample were used to generate a heat map, and clustered functional differences. On the horizontal level 1 of the clustering heat map (Fig. 6A), the T2 group had the highest proportion of genetic information processing. On the horizontal level 2 of the clustering heat map (Fig. 6B), immune system, glycan biosynthesis and metabolism, and carbohydrate metabolism had a higher level of function in the T2 group than in others. On the horizontal level 3 of the clustering heat map (Fig. 6C), the abundances of exosome, cysteine and methionine Metabolism, glycolysis/gluconeogenesis, alanine, aspartate and glutamate metabolism, amino acid-related enzymes, peptidases, starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism, etc. were only relatively high in the T2 group. The T2 group had unique features also in the horizontal level k of clustering heat map (Fig. 6D). The abundances of K03406 (methyl-accepting chemotaxis protein),

K01153 (type I restriction enzyme, R subunit [EC:3.1.21.3]), K03043 (DNA-directed RNA polymerase subunit beta [EC:2.7.7.6]), K02337 (DNA polymerase III subunit alpha [EC:2.7.7.7]), K01955, K03046, K06147, K03657, K09687, K02355, K03701, K02003, K01338, K02004, and K02529 were higher in the T2 group than in the other groups. Overall, the T2 group showed unique functional characteristics at levels 1, 2, 3, and k.

Discussion

In this study, we aimed to reveal the composition, diversity, and function of soil microbial community in the world's longevity township, Jiaoling (China), to compare microbiota composition differences between different towns. Prediction of local community compositions and functions³¹ improved the understanding of soil microbial community under current and future changing climate conditions³², which promote research on soil microbiota.

Based on the species annotation analysis, the dominant phyla in the samples were *Proteobacteria*, *Firmicutes*, *Acidobacteria*, and *Actinobacteria*^{33,34}. The dominant genera were *Chujaibacter*, *Acidipila*, and *Lactobacillus*. Interestingly, specific species were only dominant species in specific areas, which may be closely related to the geographical environment³⁵. For example, at the family level, *Rikenellaceae*, *Ruminococcaceae*, *Bacteroidaceae*, *Muribaculaceae*, *Lactobacillaceae*, and *Lachnospiraceae* were only dominant in the T2 group. At the genus level, *Bacteroides*, unidentified *Ruminococcaceae*, *Lactobacillus*, unidentified *Lachnospiraceae*, unidentified *Burkholderiaceae*, *Ralstonia*, *Collimonas*, candidatus *Solibacter*, candidatus *Koribacte* were only dominant in the T2 group. However, the dominant species in the other groups were not dominant in the T2 group. This finding might emphasise the influence of geographical changes on the composition of the local microbial community. Based on the Chao1 index and Wilcoxon rank-sum analysis of the alpha diversity of samples, the alpha diversity of the T2 group was significantly lower than that of the other groups. The beta diversity analysis based on the weighted Unifrac beta-Wilcox rank-sum test index revealed statistically significant differences between the T2 and other groups ($p < 0.05$). Perhaps the diversity of the T2 group was significantly different from other groups was precisely linked to the different geographical localisation of the above species.

The analysis of the differences between the alpha and beta diversity indices revealed that different groups' overall community structures were significantly different. The LEfSe analysis of T1–T8 samples revealed biomarkers, including *Lactobacillus*, *Muribaculaceae*, *Ruminococcaceae*, *Lachnospiraceae*, and *Chitinophagales*. Different species have different contributions to different groups; the highest contribution was that of *Chujaibacter*, followed by *Acidipila* and *Lactobacillus*. The environmental microbiota is complex and diverse. We observed that the dominant strain in the T2 group was *Lactobacillus*. This prompts the urgent questions of how its survival among thousands of microorganisms and in the colony affects the local microbiota and is affected by local microbiota. Using the species co-occurrence network diagram, we intuitively visualise the dominant species and closely interacting populations that are dominant in specific environments. These dominant species and populations often play a unique and important role in maintaining the stable structure and function of the

microbial community in the environment³⁶. In the soil microbial co-occurrence network, approximately 20 strains were positively correlated with *Lactobacillus*, including *Anaerostipes*, *Blautia*, *Dialister*, and *Streptococcus*; and approximately 35 strains negatively correlated with *Lactobacillus*, including *Minicystis*, *Rhizorhapis*, *Phycococcus*, and *Oligoflexus*. In addition to its good viability under various conditions, *Lactobacillus* can also exert a synergistic effect on other microbes by regulating the growth of the above-mentioned impacted strains. *Lactobacillus*, as a general probiotic, is widely distributed, e.g., large numbers of *Lactobacillus* have been isolated from the environment (soil and water), food (fermented dairy products, fermented soy products, etc.)³⁷, and various organisms (human, animals, etc.)³⁸. Their presence greatly affects human life and health. Probiotics are currently used to solve problems in various fields, including agriculture, aquaculture, food processing industry, medical treatment, etc., especially in the food and medicine, and, recently, cosmetics industries.

Molecular methods, such as metagenomics (estimating microbial composition and genome capacity), metatranscriptomics (estimating gene expression), and metaproteomics (estimating protein synthesis), provide insights into the functional profile of the entire soil communities. The expression of most functional genes in the soil is different in different communities, and is driven by climate and soil conditions³⁹. Clustering based on the functional annotation and published abundance information can determine the functional difference levels. On the horizontal level 2 of the clustering heat map, immune system, glycan biosynthesis and metabolism, and carbohydrate metabolism had a higher level of function in the T2 group than in others. The abundances of exosome, cysteine and methionine metabolism, glycolysis/gluconeogenesis, alanine, aspartate and glutamate metabolism, amino acid-related enzymes, peptidases, starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism, etc. were only relatively high in the T2 group on the horizontal level 3 of the clustering heat map. Collectively, T2 group samples have unique functional characteristics at different levels. This may be related to its unique geographical environment, and the underpinning factors require further research. Longevity is comprehensively influenced by environment (soil, water, air, diet, lifestyle, exposome, medication, development, economy), host (genetics, gender, age, stress, psychological factors), microbiome (homeostasis, composition, diversity, function), among other factors. (Fig. 7). The microbiome may play a key role in the connection between the environment and the host. Thus, we may need to ensure a healthy gut microbiota in order to be longevous. The specific bacterial groups found in this study may be related to host ageing and used to promote a healthy microbiota and longevity.

There are some limitations to this study. First, the functions of the soil microbiota were only predicted using high-throughput sequencing. It is difficult to isolate as many types of microorganisms as those analyzed using high-throughput sequencing data. In the future, *in vivo* and *in vitro* functional verification tests on isolated strains will need to be conducted to ensure the accuracy of the high-throughput sequencing analysis results. Second, seasonal changes in the environmental microbiota were not considered, and must be taken into account in future studies. Third, we did not carry out an in-depth longitudinal study of the host microbiota, which should be considered in future studies.

Conclusion

In the current study, 64 bacterial phyla, 74 classes, 148 orders, 276 families, 664 genera, and 513 species were identified in all soil samples. We found that considering the composition and diversity of the soil microbial community in Jiaoling, China, the world's longevity township, at the phylum level, the dominant bacteria were *Proteobacteria*, *Firmicutes*, and *Acidobacteria*. The dominant bacteria at the genus level were *Chujaibacter*, *Acidipila*, and *Lactobacillus*. However, *Bacteroides*, *Lactobacillus*, *Ralstonia*, *Collimonas*, *Candidatus Solibacter*, *Candidatus Koribacter* were only dominant in the T2 group. The alpha diversity and beta diversity of the T2 group were significantly different from those of other groups. That was because of the different geographic locations of these species. LEfSe and ternary plot analysis revealed 12 biomarker microbes that were responsible for significant differences between the groups, mainly including *Lactobacillus*, *Muribaculaceae*, *Ruminococcaceae*, *Lachnospiraceae*, *Chitinophagales*, etc. Simper analysis indicated the species that contributed to the above differences between the groups and their contribution size. The contribution from high to low was *Chujaibacter*, *Acidipila*, *Lactobacillus*, *Rhodanobacter*, unidentified *Prevotellaceae*, *Lysobacter*, *Bryobacter*, *Granulicella*, *Flavobacterium*, and *Mizugakiibacter*. The contributions of different species were different in different groups, with the highest contribution of *Chujaibacter*, followed by *Acidipila* and *Lactobacillus*. The microbiome may play a key role in the connection between the environment and the host. Regional environmental factor soil microbiota as a potential bridge between environment and long-lived People. These data provide a basis for understanding the significant differences in the composition and diversity of soil microbial communities in the long-lived people in longevity township Jiaoling, China.

Methods

Sample collection

August 2018, 52 soil samples were collected from eight towns (T1, Guangfu, 3 samples; T2, Nanzai, 3 samples; T3, Wenfu, 8 samples; T4, Changtan, 11 samples; T5, Jiaocheng, 4 samples; T6, Lanfang, 14 samples; T7, Sanzhen, 3 samples; and T8, Xinpu, 6 samples) belonging to the "world's longevity township" of Jiaoling County (Meizhou City, Guangdong Province, China). Soil samples were collected from the front and back of a long-lived family's house or vegetable garden under ground 10-15cm. After collection, the samples were placed at 4 °C during transportation to the laboratory and then at -80 °C until DNA extraction was performed. The distribution of population over 80 years old is provided by the local Jiaoling County government.

DNA extraction and polymerase chain reaction (PCR) amplification

The genomic DNA was extracted and its purity and concentration were evaluated using agarose gel electrophoresis. Sample DNA was diluted to 1 ng/L in sterile water. PCR primers were designed to target the V4 and V5 hypervariable regions of the bacterial 16S rRNA gene. Phusion high-fidelity PCR Master Mix with GC Buffer (New England Biolabs) were used to ensure the amplification efficiency and accuracy, with the diluted genomic DNA as the template.

16S rRNA gene sequencing

The amplicon library was built using the Ion Plus Fragment Library Kit 48 RXNS (ThermoFisher, Shanghai, China). The library was sequenced using Ion S5™XL (ThermoFisher, Shanghai, China) at Novogene Bioinformatics Technology Co., Ltd. A single-end sequencing method was used to construct a small-fragment library.

Processing of sequencing data

Cutadapt (v1.9.1, <http://cutadapt.readthedocs.io/en/stable/>)⁴⁰ was first used to remove low-quality fragments of reads. Next, the data were filtered for preliminary quality control of the raw reads. Then, to remove the chimera sequences and obtain clean reads, they were processed by comparing with an annotation database as described elsewhere⁴¹ using GitHub (<https://github.com/torognes/vsearch/>)⁴².

Statistical analysis

Uparse software (Uparse v7.0.1001, <http://www.drive5.com/uparse/>)⁴³ was used to cluster all clean reads from the samples. Mothur and the SSUrRNA database⁴⁴ from SILVA132 (<http://www.arb-silva.de/>)⁴⁵ were used for species annotation. MUSCLE⁴⁶ (v3.8.31, <http://www.drive5.com/muscle/>) was used for rapid multiple sequence alignment to determine the phylogenetic relationship of all operational taxonomic units (OTUs). QIIME (v1.9.1) was used to calculate the Shannon index and the Unifrac distance, and to construct the sample-clustering tree using the unweighted pair group method with arithmetic mean (UPGMA). R (v2.15.3) was used for dilution curve, rank abundance curve, and species accumulation curve calculations, and to analyze the differences in alpha and beta diversity indices between the groups using the Wilcoxon test in the agricolae package for parametric test. Linear discriminant analysis effect size (LEfSe) software was used to find biomarkers using a default score setting of linear discriminant analysis (LDA) of 4. Based on species abundance, the correlation coefficient (Spearman correlation coefficient or Pearson correlation coefficient) was calculated between the genera, to obtain a correlation coefficient matrix. For edges, graphviz (v2.38.0) was used to draw the network graph. The Tax4Fun function prediction was executed using the nearest-neighbor method based on the minimum 16S rRNA sequence similarity. The specific method involved extracting the 16S rRNA gene sequences from the whole prokaryotic genome of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and comparing them to those available in the SILVA SSU Ref NR database (BLAST bitscore > 1,500) using the BLASTN algorithm to establish a correlation matrix.

Declarations

Funding

This study was jointly supported by research grants from the Key-Area Research and Development Program of Guangdong Province (2018B020205002), and the Guangdong Province Academy of Sciences Special Project for Capacity Building of Innovation Driven Development (2020GDASYL-20200301002)

Author contributions

LW, QW, YD, XX and JZ designed this study. LW and JM collected the samples and conducted the experiments. LW and JM analysed the data. LW drafted the manuscript. LW, QW, and YD revised the manuscript. QW and YD supervised the project. All authors approved the final version of the manuscript.

Conflict of interest statement

There is no conflict of interest.

References

1. Kenyon, C. The plasticity of aging: Insights from long-lived mutants. *Cell* **120**, 449–460. doi: 10.1016/j.cell.2005.02.002 (2005).
2. Fontana, L., Partridge, L., & Longo, V.D. Extending Healthy Life Span-From Yeast to Humans. *Science* **328**, 321–326. doi: 10.1126/science.1172539 (2010).
3. Junnila, R. K., List, E. O., Berryman, D. E., Murrey, J. W., & Kopchick, J. J. The GH/IGF-1 axis in ageing and longevity. *Nat Rev Endocrinol.* **9**, 366–376. doi: 10.1038/nrendo.2013.67 (2013).
4. Giannakou, M. E., Goss, M., Junger, M. A., Hafen, E., Leivers, S. J., & Partridge, L. Long-lived *Drosophila* with overexpressed dFOXO in adult fat body. *Science* **305**, 361–361. doi: 10.1126/science.1098219 (2004).
5. Bluher, M., Kahn, B.B., & Kahn, C.R. Extended longevity in mice lacking the insulin receptor in adipose tissue. *Science* **299**, 572–574. doi: 10.1126/science.1078223 (2003).
6. Tan, Q., Soerensen, M., Kruse, T. A., Christensen, K., & Christiansen, L. A novel permutation test for case-only analysis identifies epistatic effects on human longevity in the FOXO gene family. *Aging Cell* **12**, 690–694. doi: 10.1111/acel.12092 (2013).
7. Salminen, A., & Kaarniranta, K. AMP-activated protein kinase (AMPK) controls the aging process via an integrated signaling network. *Ageing Res Rev.* **11**, 230–241. doi: 10.1016/j.arr.2011.12.005 (2012).
8. Kenyon, C.J. The genetics of ageing. *Nature* **464**, 504–512. doi: 10.1038/nature08980 (2010).
9. Dang, W. et al. Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* **459**, 802–U802. doi: 10.1038/nature08085 (2009).
10. Mahley, R.W., & Huang, Y. Apolipoprotein E sets the stage: response to injury triggers neuropathology. *Neuron* **76**, 871–885. doi: 10.1016/j.neuron.2012.11.020 (2012).
11. Blackburn, E.H. Switching and signaling at the telomere. *Cell* **106**, 661–673. doi: 10.1016/s0092-8674(01)00492-5 (2001).

12. Jazwinski, S.M., & Yashin, A.I. Aging and health—a systems biology perspective. Introduction. *Interdiscip Top Gerontol.* **40**, 7–12. doi: 10.4081/ejh.2015.2508 (2015).
13. Bjedov, I., & Partridge, L. A longer and healthier life with TOR down-regulation: genetics and drugs. *Biochem Soc Trans.* **39**, 460–465. doi: 10.1042/bst0390460 (2011).
14. Tyner, S.D. et al. p53 mutant mice that display early ageing-associated phenotypes. *Nature* **415**, 45–53. doi: 10.1038/415045a (2002).
15. Baker, D.J. et al. Clearance of p16(Ink4a)-positive senescent cells delays ageing-associated disorders. *Nature* **479**, 232–236. doi: 10.1038/nature10600 (2011).
16. Zhang, G. et al. Hypothalamic programming of systemic ageing involving IKK-beta, NF-kappa B and GnRH. *Nature* **497**, 211–216. doi: 10.1038/nature12143 (2013).
17. Melendez, A., Tallocczy, Z., Seaman, M., Eskelinen, E. L., Hall, D. H., & Levine, B. Autophagy genes are essential for dauer development and life-span extension in C-elegans. *Science* **301**, 1387–1391. doi: 10.1126/science.1087782 (2003).
18. Simonsen, A., Cumming, R. C., Brech, A., Isakson, P., Schubert, D. R., & Finley, K. D. Promoting basal levels of autophagy in the nervous system enhances longevity and oxidant resistance in adult *Drosophila*. *Autophagy* **4**, 176–184. doi: 10.4161/auto.5269 (2008).
19. Pyo, J.O. et al. Overexpression of Atg5 in mice activates autophagy and extends lifespan. *Nat Commun.* **4**, 2300. doi: 10.1038/ncomms3300 (2013).
20. Guttman, M. et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227. doi: 10.1038/nature07672 (2009).
21. Jin, L., Song, Q., Zhang, W., Geng, B., & Cai, J. Roles of long noncoding RNAs in aging and aging complications. *Biochim Biophys Acta Mol Basis Dis.* **1865**, 1763–1771. doi: 10.1016/j.bbadis.2018.09.021 (2019).
22. Vanhooren, V. et al. Protein modification and maintenance systems as biomarkers of ageing. *Mech Ageing Dev.* **151**, 71–84. doi: 10.1016/j.mad.2015.03.009 (2015).
23. Oien, D.B., & Moskovitz, J. Substrates of the methionine sulfoxide reductase system and their physiological relevance. *Curr Top Dev Biol.* **80**, 93–133. doi: 10.1016/S0070-2153(07)80003-2 (2008).
24. Loo, B.P.Y., Lam, W.W.Y., Mahendran, R. & Katagiri, K. How Is the Neighborhood Environment Related to the Health of Seniors Living in Hong Kong, Singapore, and Tokyo? Some Insights for Promoting Aging in Place. *Ann. Am. Assoc. Geog.* **107**, 812–828. doi.org/10.1080/24694452.2016.1271306 (2017).

25. Lv, J.M., Wang, W.Y. & Li, Y.H. Effects of environmental factors on the longevous people in China. *Arch. Gerontol. Geriatr.* **53**, 200–205. doi.org/10.1016/j.archger.2010.10.012 (2011).
26. Colman, R.J. et al. Caloric restriction delays disease onset and mortality in rhesus monkeys. *Science* **325**, 201–204. doi: 10.1126/science.1173635 (2009).
27. Christensen, K., Doblhammer, G., Rau, R., & Vaupel, J. W. Ageing populations: the challenges ahead. *Lancet* **374**, 1196–1208. doi: 10.1016/s0140-6736(09)61460-4 (2009).
28. Huang, Y., Rosenberg, M., Hou, L.L. and Hu, M.J. Relationships among Environment, Climate, and Longevity in China. *Int. J. Env. Res. Pub. He.* **14**. doi.org/10.3390/ijerph14101195 (2017).
29. Favorito, J.E., Grossl, P.R., Davis, T.Z., Eick, M.J. & Hankes, N., Soil-plant-animal relationships and geochemistry of selenium in the Western Phosphate Resource Area (United States): A review. *Chemosphere* **266**, 128959. doi.org/10.1016/j.chemosphere.2020.128959 (2021).
30. van Bruggen, A.H.C., Goss, E.M., Havelaar, A., van Diepeningen, A.D., Finckh, M.R. & Morris, J.G., Jr. One Health - Cycling of diverse microbial communities as a connecting force for soil, plant, animal, human and ecosystem health. *Sci. Total Environ.* **664**, 927–937. doi.org/10.1016/j.scitotenv.2019.02.091 (2019).
31. Fierer, N., & Jackson, R.B. The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A.* **103**, 626–631. doi: 10.1073/pnas.0507535103 (2006).
32. Bahram, M. et al. Structure and function of the global topsoil microbiome. *Nature* **560**, 233–237. doi: 10.1038/s41586-018-0386-6 (2018).
33. Delgado-Baquerizo, M. et al. A global atlas of the dominant bacteria found in soil. *Science* **359**, 320–325. doi: 10.1126/science.aap9516 (2018).
34. Fierer, N. et al. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Natl Acad Sci U S A.* **109**, 21390–21395. doi: 10.1073/pnas.1215210110 (2012).
35. Lauber, C. L., Hamady, M., Knight, R., & Fierer, N. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl And Environ Microbiol.* **75**, 5111–5120. doi: 10.1128/aem.00335-09 (2009).
36. Ma, B. et al. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *ISME J.* **10**, 1891–1901. doi: 10.1038/ismej.2015.261 (2016).
37. Cavaillon, J.M., & Legout, S. Centenary of the death of Elie Metchnikoff: a visionary and an outstanding team leader. *Microbes And Infection* **18**, 577–594. doi: 10.1016/j.micinf.2016.05.008 (2016).

38. Biagi, E. et al. Gut microbiota and extreme longevity. *Curr Biol.* **26**, 1480–1485. doi: 10.1016/j.cub.2016.04.016 (2016).
39. Noronha, M. F., Lacerda, G. V., Gilbert, J. A., & de Oliveira, V. M. Taxonomic and functional patterns across soil microbial communities of global biomes. *Sci Total Environ.* **609**, 1064–1074. doi: 10.1016/j.scitotenv.2017.07.159 (2017).
40. Aßhauer, K. P., Wemheuer, B., Daniel, R., & Meinicke, P. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* **31**, 2882–2884. doi: 10.1093/bioinformatics/btv287 (2015).
41. Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahe, F. VSEARCH: a versatile open source tool for metagenomics. *Peer J.* **4**, e2584. doi: 10.7717/peerj.2584 (2016).
42. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**. doi: 10.14806/ej.17.1.200 (2011).
43. Haas, B.J. et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.* **21**, 494–504. doi: 10.1101/gr.112730.110 (2011).
44. Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol.* **73**, 5261–5267. doi: 10.1128/aem.00062-07 (2007).
45. Edgar, R.C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods.* **10**, 996–998. doi: 10.1038/nmeth.2604 (2013).
46. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596. doi: 10.1093/nar/gks1219 (2013).

Figures

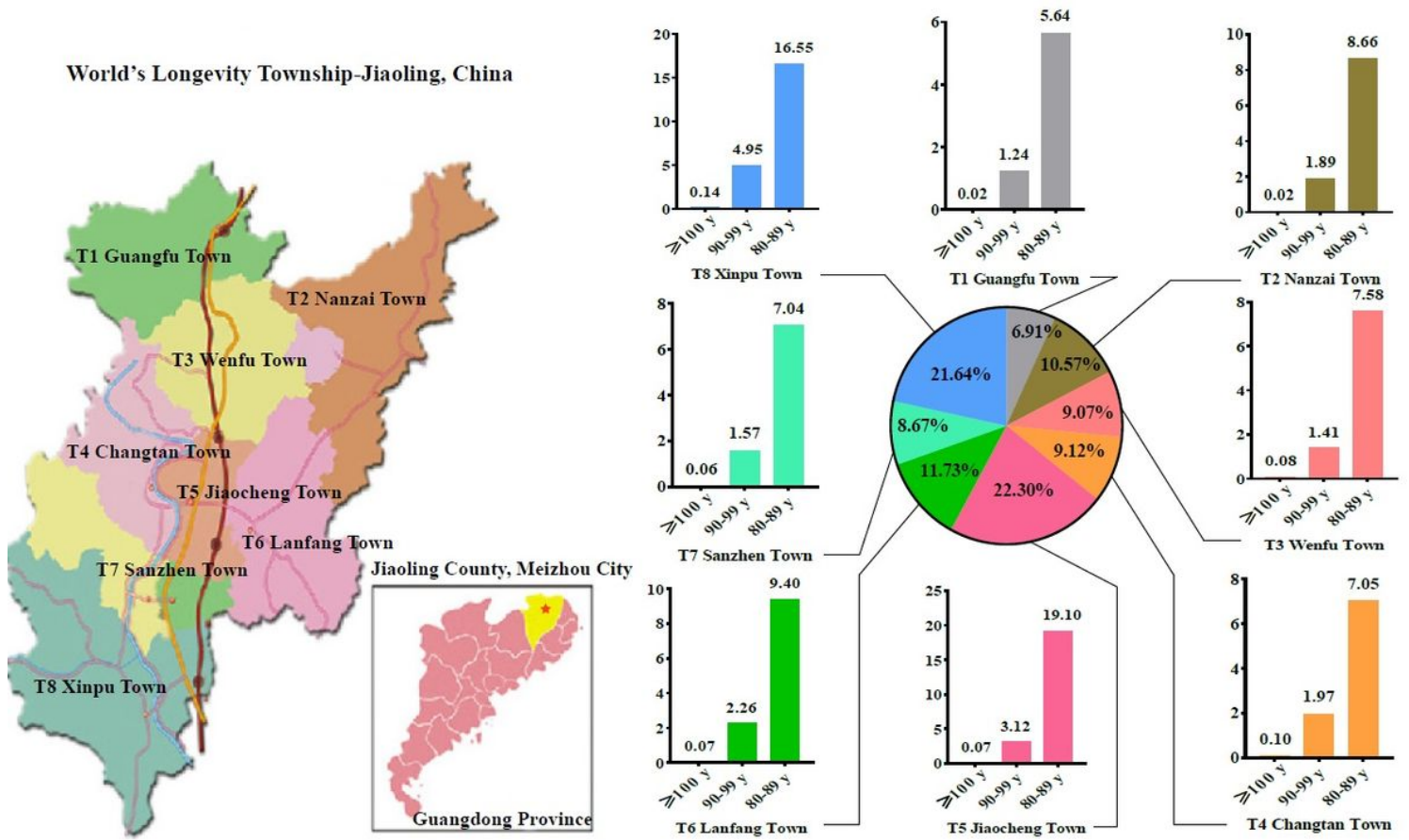


Figure 1

Distribution map of the population over 80 years old in world longevity township-Jiaoling, China. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

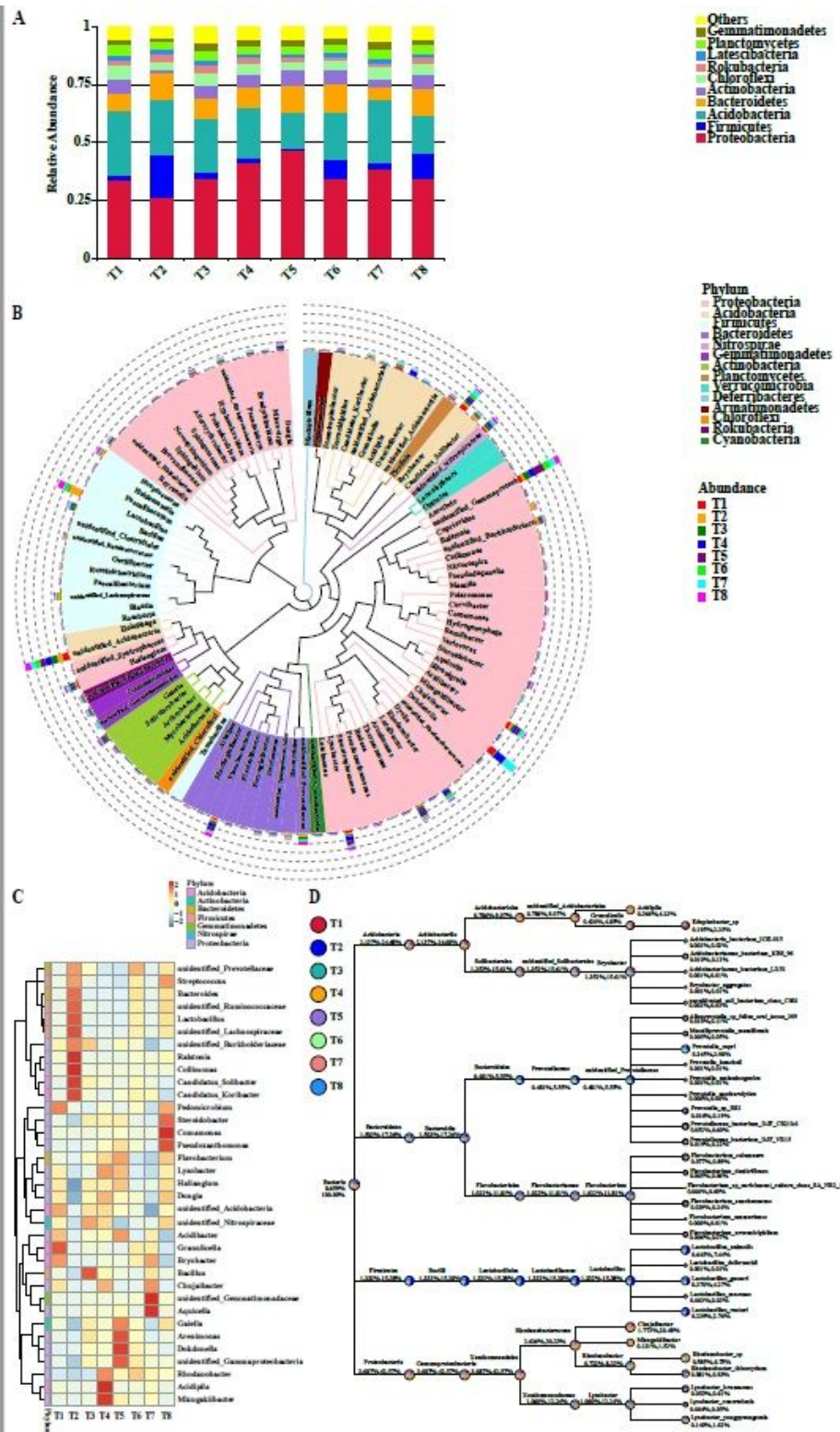


Figure 2

Composition and structure of soil microflora. (A) Columnar cumulative graph of the relative abundance of the top 10 species of microbiota among the different groups at the phylum level; (B) Cluster heat map of the composition abundance of the top 35 species of microbiota among different groups at the genus level; (C) The phylogenetic relationship of species at the genus level; (D) Species classification tree analysis (selected top 10 genera with the largest relative abundance).

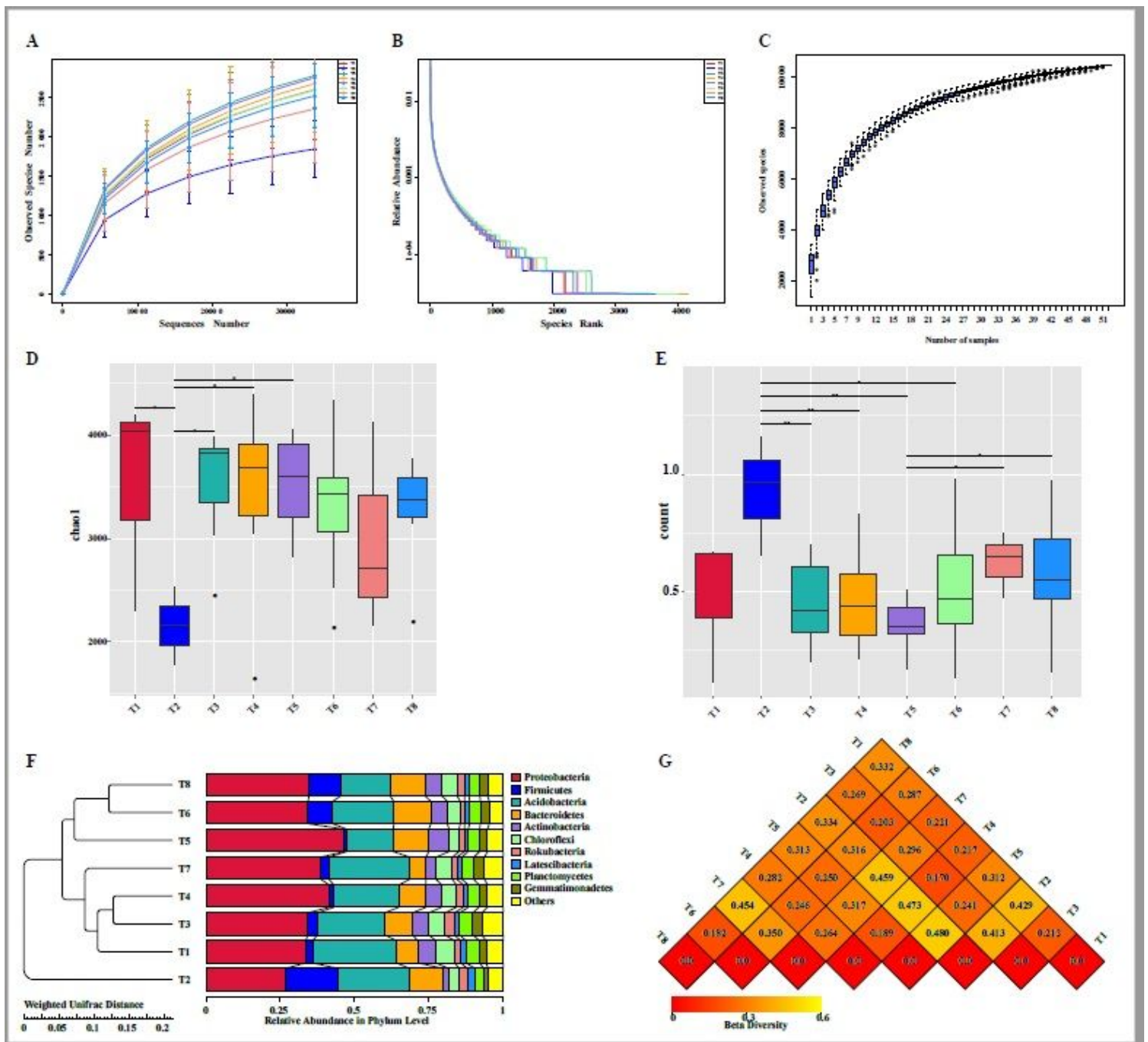


Figure 3

Alpha diversity and beta diversity. (A) The rarefaction curve; (B) Rank abundance curve; (C) Box plot of species accumulation; (D) Analysis of alpha diversity based on the Wilcoxon rank sum test of Chao1 index; (E) Box plot of differences between beta diversity groups based on unweighted unifracs beta-Wilcox index; (F) Beta diversity analysis of UPGMA cluster tree based on unweighted unifracs distance; (G) Heat map of distance matrix drawn with weighted unifracs distance. * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

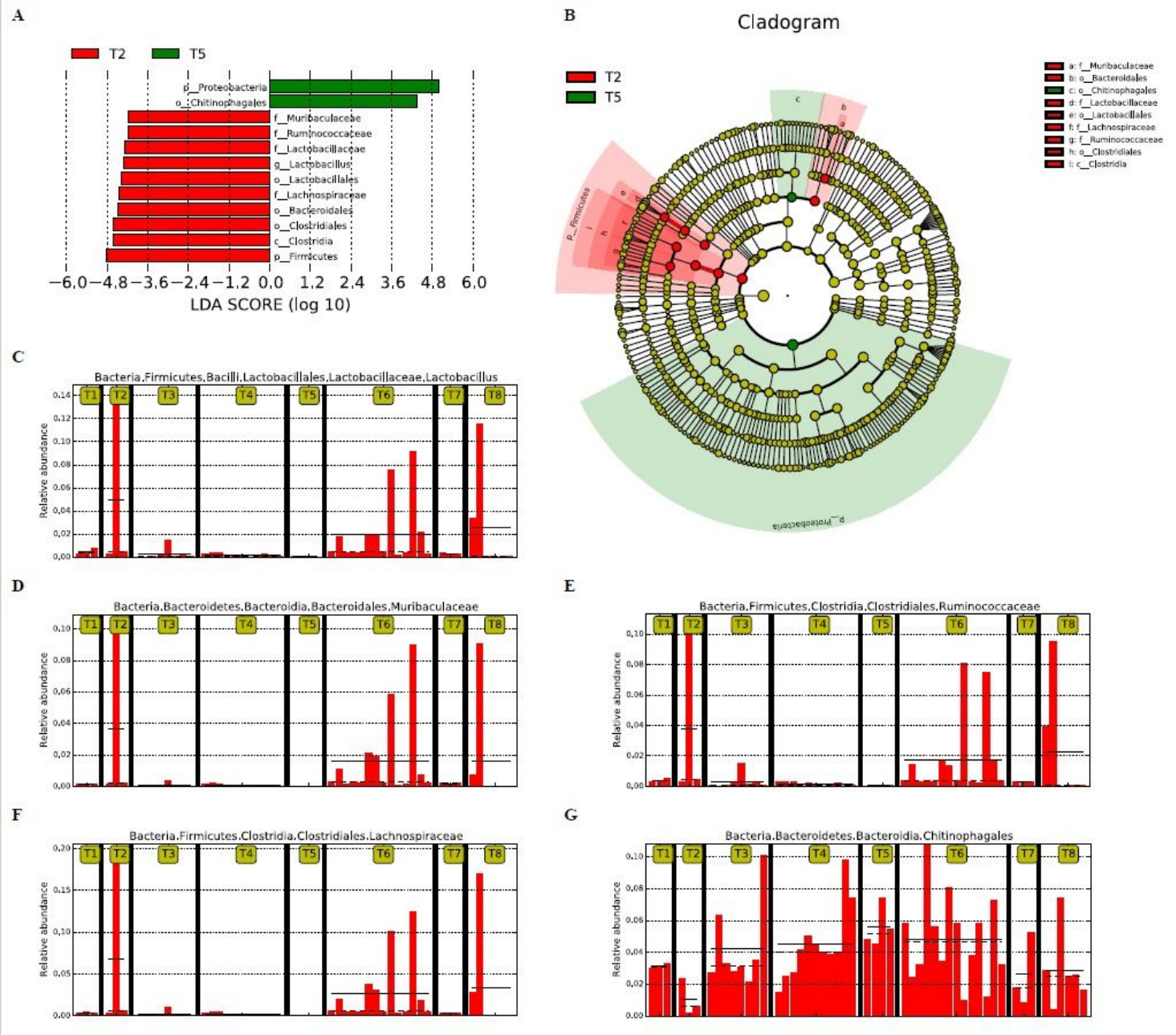


Figure 4

Significance analysis of species differences between groups. (A) Histogram of LDA value distribution of species with significant differences in abundance between groups; (B) Evolutionary clade of species with significant differences in abundance between groups; (C) Species with significant differences in abundance between groups-Lactobacillus; (D-G) Species with significant differences in abundance between groups-Muribaculaceae, Ruminococcaceae, Lachnospiraceae, Chitinophagales.

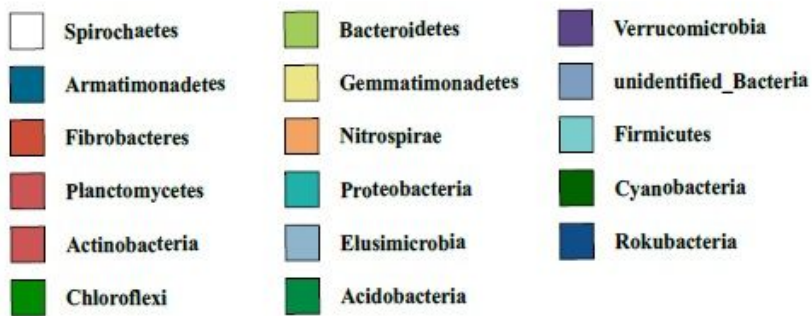
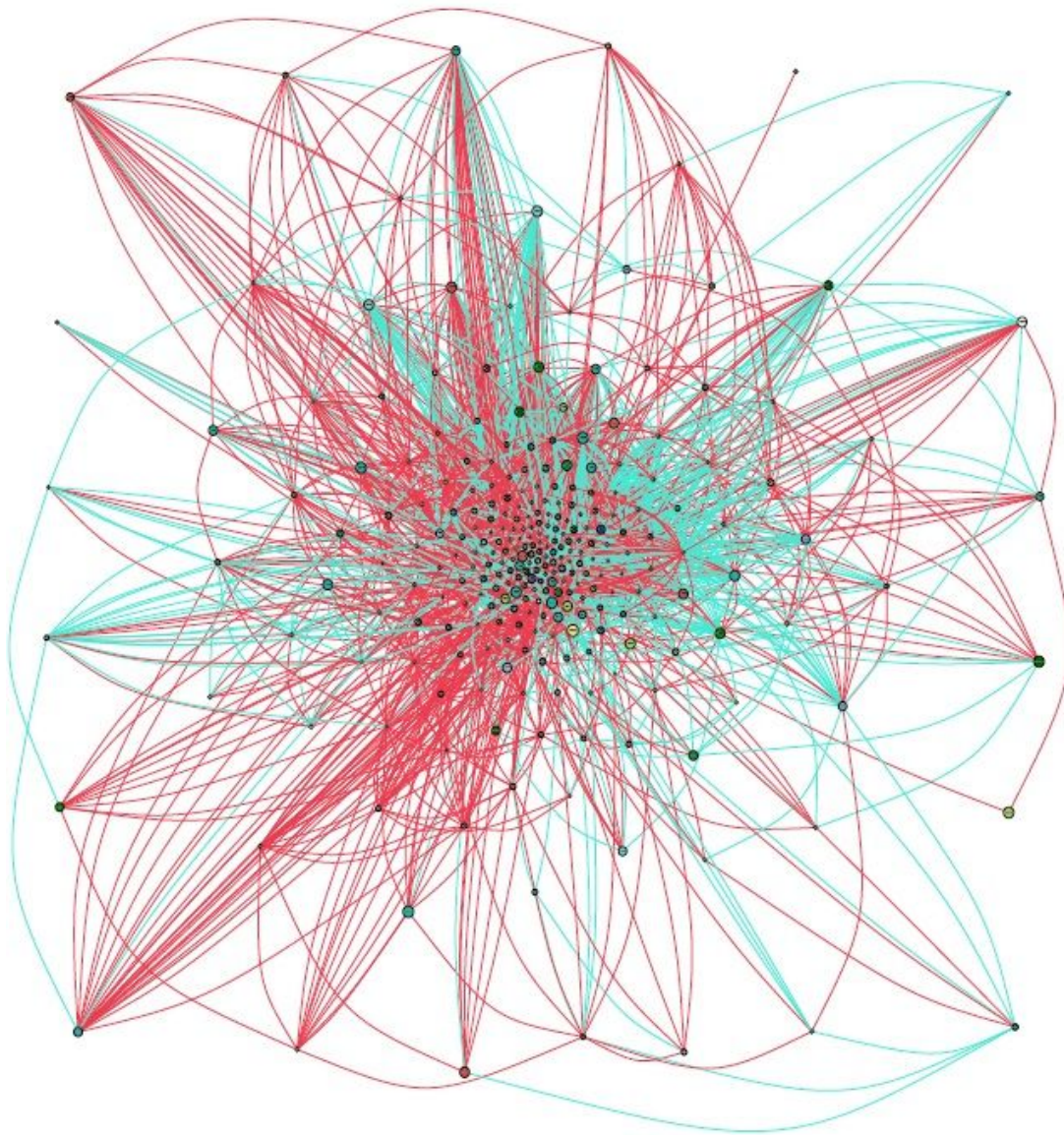


Figure 5

Co-occurrence network of soil microbiota.

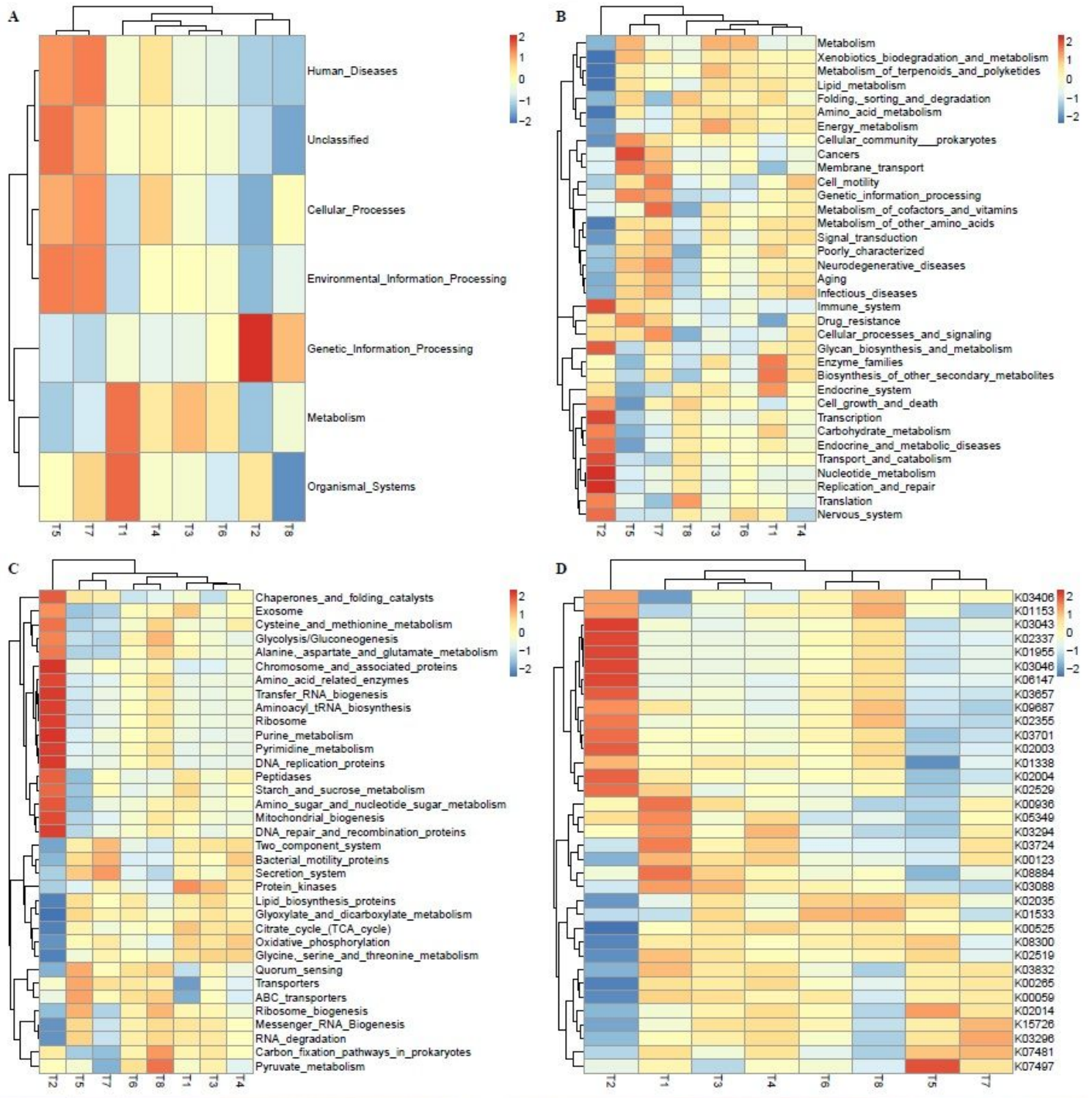


Figure 6

Predicted functions. (A-D) Clustering heat map of functional differences at level 1, level 2, level 3, and level k.

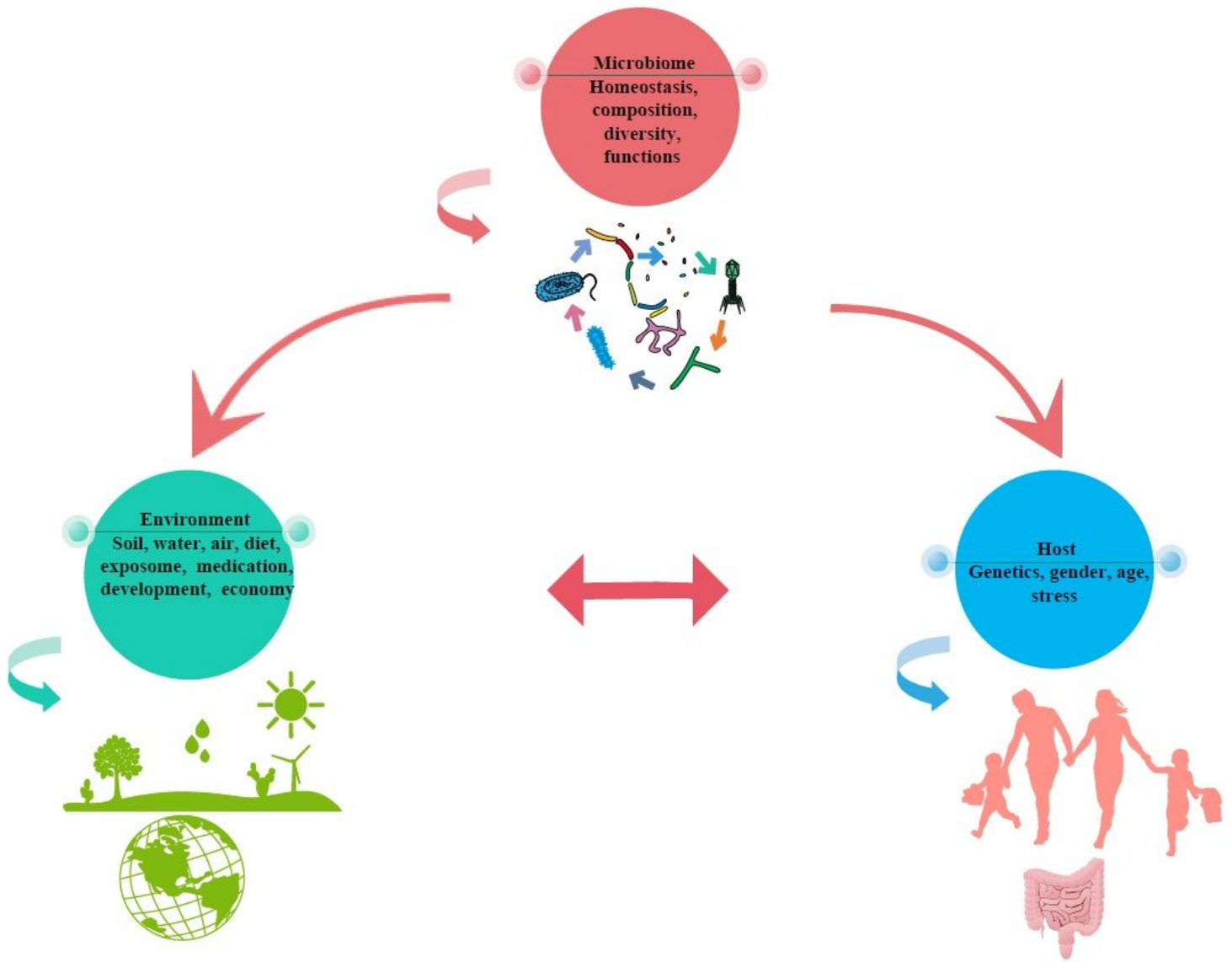


Figure 7

The relationship between environment, host, and microbiota.

Supplementary Files

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

- [FigureS1.pdf](#)
- [FigureS2.pdf](#)
- [FigureS3.pdf](#)
- [FigureS4.pdf](#)