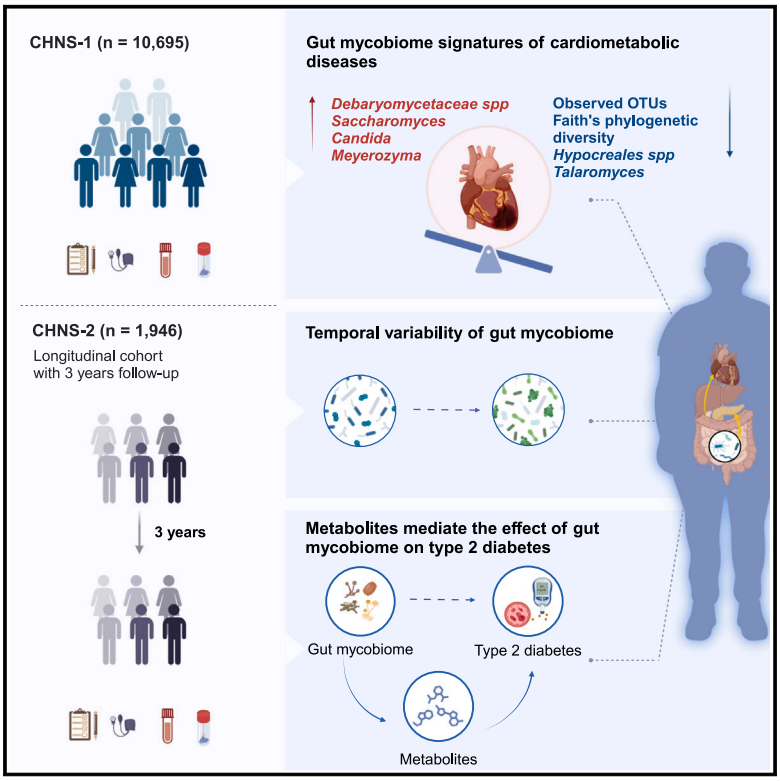
Cell Reports
Medicine

Article

The temporal dynamics of the gut mycobiome and its association with cardiometabolic health in a nationwide cohort of 12,641 Chinese adults

Graphical abstract



Highlights

- We characterize gut fungal composition and dynamics in 12,641 Chinese participants
- Gut fungal diversity and 19 genera are associated with cardiometabolic health
- The influence of gut fungi on type 2 diabetes may be mediated by serum metabolites

Authors

Wanglong Gou, Huijun Wang, Chang Su, ..., Ke Zhang, Bing Zhang, Ju-Sheng Zheng

Correspondence

zhangbing@chinacdc.cn (B.Z.), zhengjusheng@westlake.edu.cn (J.-S.Z.)

In brief

The association between the gut mycobiome and cardiometabolic health remains underappreciated. In a cohort of 12,641 Chinese participants, Gou et al. characterize the gut mycobiome's composition and dynamics, demonstrating its connection to cardiometabolic health. They also identify several metabolites potentially mediating the association of the gut mycobiome with type 2 diabetes.

Gou et al., 2024, Cell Reports Medicine 5, 101775
October 15, 2024
https://doi.org/10.1016/j.xcrm.2024.101775

CellPress

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

The temporal dynamics of the gut mycobiome and its association with cardiometabolic health in a nationwide cohort of 12,641 Chinese adults

Wanglong Gou,[1,2,7] Huijun Wang,[3,4,7] Chang Su,[3,4,7] Yuanqing Fu,[1,2,5,7] Xinyu Wang,[1,2] Chang Gao,[1,2] Menglei Shuai,[1,2] Zelei Miao,[1,2,5] Jiguo Zhang,[3,4] Xiaofang Jia,[3,4] Wenwen Du,[3,4] Ke Zhang,[1,2] Bing Zhang,[3,4,*] and Ju-Sheng Zheng[1,2,5,6,8,*]

[1]Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, China
[2]Zhejiang Key Laboratory of Multi-Omics in Infection and Immunity, Center for Infectious Disease Research, School of Medicine, Westlake University, Hangzhou, China
[3]National Institute for Nutrition and Health, Chinese Center for Disease Control and Prevention, Beijing, China
[4]NHC Key Laboratory of Public Nutrition and Health, Beijing, China
[5]Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, China
[6]Research Center for Industries of the Future, School of Life Sciences, Westlake University, Hangzhou, China
[7]These authors contributed equally
[8]Lead contact
*Correspondence: zhangbing@chinacdc.cn (B.Z.), zhengjusheng@westlake.edu.cn (J.-S.Z.)
https://doi.org/10.1016/j.xcrm.2024.101775

SUMMARY

The dynamics of the gut mycobiome and its association with cardiometabolic health remain largely unexplored. Here, we employ internal transcribed spacer (ITS) sequencing to capture the gut mycobiome composition and dynamics within a nationwide human cohort of 12,641 Chinese participants, including 1,946 participants with repeated measurements across three years. We find that the gut mycobiome is associated with cardiometabolic diseases and related biomarkers in both cross-sectional and dynamic analyses. Fungal alpha diversity indices and 19 mycobiome genera are the major contributors to the mycobiome-cardiometabolic disease link. Particularly, *Saccharomyces* emerges as an effect modifier of traditional risk factors in promoting type 2 diabetes risk. Further integration of multi-omics data reveals key metabolites such as γ-linolenic acid and L-valine linking the gut mycobiome to type 2 diabetes. This study advances our understanding of the potential roles of the gut mycobiome in cardiometabolic health.

INTRODUCTION

The fungal microbiome, known as the mycobiome, is omnipresent in our living environment and has played a central role in the evolution of life.[1] Despite fungi comprising a small proportion of the human gut microbiota, alterations in the fungal community are suggested to be correlated with various health conditions and diseases, such as host immunity,[2] gastrointestinal diseases,[3] and cancer.[4] Yet, vast majority of the prior clinical evidence about human microbiome and health has been focused on gut bacteria but not fungi. The gut mycobiome research remains in its infancy stage, representing a rather underestimated field. Whether and how the gut mycobiome may influence the disease progression and health status in humans are largely unknown.

Cardiometabolic diseases, such as type 2 diabetes (T2D), hypertension, and dyslipidemia, are highly prevalent globally.[5] Recent interest has focused on understanding the role of the gut mycobiome in these diseases.[6] Studies in mice have shown that disturbances in the intestinal fungi, induced by antifungal treatments, may contribute to both intestinal and extra-intestinal pathologies.[7] In humans, there are several studies examining the association of the gut mycobiome with cardiometabolic diseases, but with small sample sizes and conflicting results.[8–12] For example, some studies report alterations in the mycobiome in individuals with T2D[9,10] and hypertension[11,12] while others suggest these changes may be limited.[8] Furthermore, gut fungal composition varies across geographic regions.[13] Previous studies, limited by geographic scope and sample size, may not have captured the full variation in fungal composition, which may confound the relationship between the gut mycobiome and cardiometabolic health. Therefore, it is essential to conduct large-scale mycobiome profiling among participants with diverse dietary and geographic backgrounds to understand the role of the gut mycobiome in cardiometabolic health.

Here, we conducted a nationwide population-based mycobiome mapping involving a deeply phenotyped cohort of 12,641 adult participants from 15 major provinces/municipalities across China. We aimed to provide a comprehensive understanding of the intricate association of the gut mycobiome with multiple host environmental phenotypes, particularly focusing on cardiometabolic health. To capture the temporal dynamics of the gut mycobiome and validate its association with cardiometabolic diseases, we further conducted a long-term follow-up

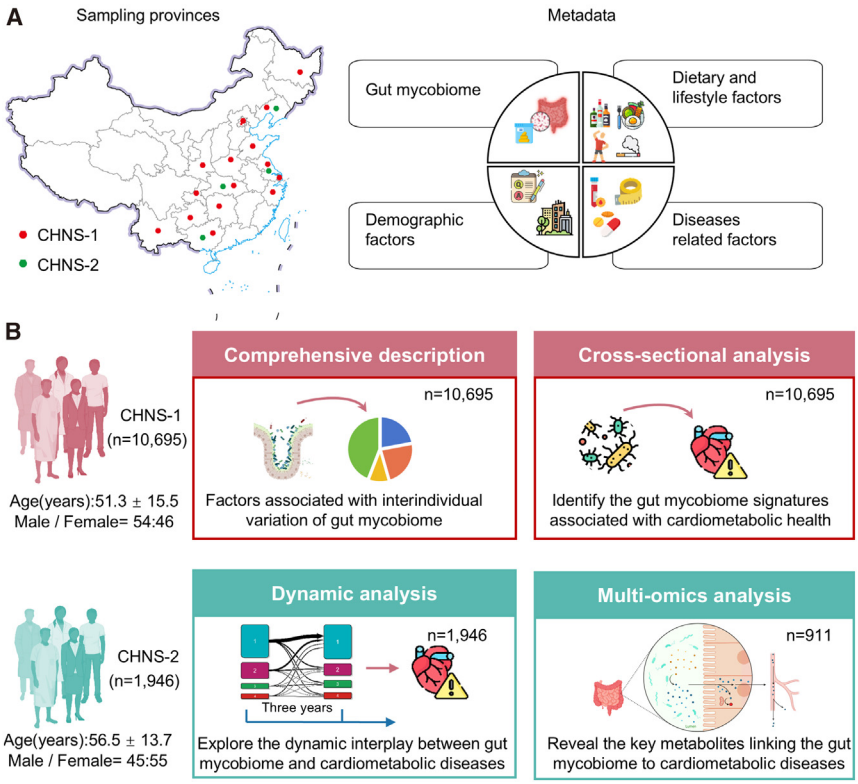


Figure 1. Conceptual outline of the study and analysis workflow

(A) Overview of sampling regions and metadata. This study utilized data from the sub-cohorts of the China Health and Nutrition Survey (CHNS), referred to as CHNS-1 (n = 10,695) and CHNS-2 (n = 1,946). CHNS-1 encompassed participants from 15 provinces or municipalities across China, while CHNS-2 drew from 4 provinces within the country. Notably, there was no overlap in participants between these two cohorts. All participants contributed ITS2 sequencing data and metadata. Moreover, repeated measurements of mycobiome (n = 1,946) and metabolome (n = 996) data are accessible for CHNS-2 participants.

(B) Overview of analysis workflow. In the CHNS-1 cohort, we initially conducted a comprehensive portrayal of the inter-individual variation of the gut mycobiome and identified its associated factors. Subsequently, we delved deeper into investigating the connection between the gut mycobiome and cardiometabolic health. In the CHNS-2 cohort, our analysis focused on exploring the long-term variability of the gut mycobiome and dynamic interplay between the gut mycobiome and cardiometabolic diseases. Furthermore, we uncovered the key metabolites potential linking the gut mycobiome to cardiometabolic diseases through multi-omics integration.

See also Table S1.

analysis in a cohort of 1,946 participants over a three-year period. Furthermore, to gain potential biological insights, we integrated serum metabolome data and used mediation analysis to explore the specific metabolites that may medicate the association of the gut mycobiome with cardiometabolic diseases.

RESULTS

Study design and cohort descriptions

In the present study, we profiled the gut mycobiome within a cohort of 12,641 Chinese adult participants (Figure 1A). These individuals were drawn from two independent sub-cohorts of the China Health and Nutrition Survey (CHNS), identified as CHNS-1 and CHNS-2. Overall characteristics of the study participants are shown in Table S1. The CHNS is a nationally representative longitudinal cohort within China, distinguished by its utilization of a meticulously designed multistage, random cluster approach, and the implementation of standardized experimental protocols.[14]

The CHNS-1 cohort consisted of individuals from 15 prominent provinces or municipalities across China (Figure 1A). Within this cohort, detailed metadata and gut mycobiome profiles were available for a total of 10,695 participants (mean age [standard deviation, SD] 51.3 ± 15.5 years, 54% were women). In the CHNS-1 cohort, we provided a comprehensive depiction of the gut mycobiome's variability and investigated the intricate associations between various host environmental phenotypes and the gut mycobiome's variation (Figure 1B). Next, we further investigated the gut mycobiome signatures of cardiometabolic health. Subsequently, our analysis was extended to the CHNS-2 cohort, which consisted of individuals from four prominent provinces in China (Figure 1A), and featured with repeated measured clinical metadata (n = 1,946 pairs), gut mycobiome data (n = 1,946 pairs), and serum metabolome data (n = 996 pairs) spanning a median period of 3.02 years (interquartile range: 3–3.07). The inclusion of the CHNS-2 cohort facilitated an exploration of the long-term variability of the gut mycobiome and uncovered the key metabolites that may link the gut mycobiome to cardiometabolic diseases (Figure 1B).

Overall gut mycobiome community profile

We profiled the gut mycobiome by sequencing the internal transcribed spacer 2 (ITS2) region. A total of 761 genera (processed via QIIME [Quantitative Insights Into Microbial Ecology] 2[15]) were detected in the CHNS-1 cohort, and 630 were identified in the CHNS-2 cohort. After filtering out genera of low prevalence (<10% prevalence in the respective cohort), 54 genera were retained in the CHNS-1 cohort, and 75 in the CHNS-2 cohort. Among these retained genera, 53 genera were shared across both the CHNS-1 and CHNS-2 cohorts.

In the CHNS-1 cohort, the two most predominant phyla in the gut mycobiome were Ascomycota and Basidiomycota (Figure 2A), consistent with previous data from the US,[16] Europe,[17] and China.[18] The prevalence and relative abundance of the 54 included genera are presented in Table S2. Additionally, the inter-correlation of included fungal genera is depicted in Figure S1. We identified five genera (*Aspergillus*, *Candida*, *Cladosporium*, *Penicillium*, and *Saccharomyces*) present in at least 50% of

CellPress
OPEN ACCESS

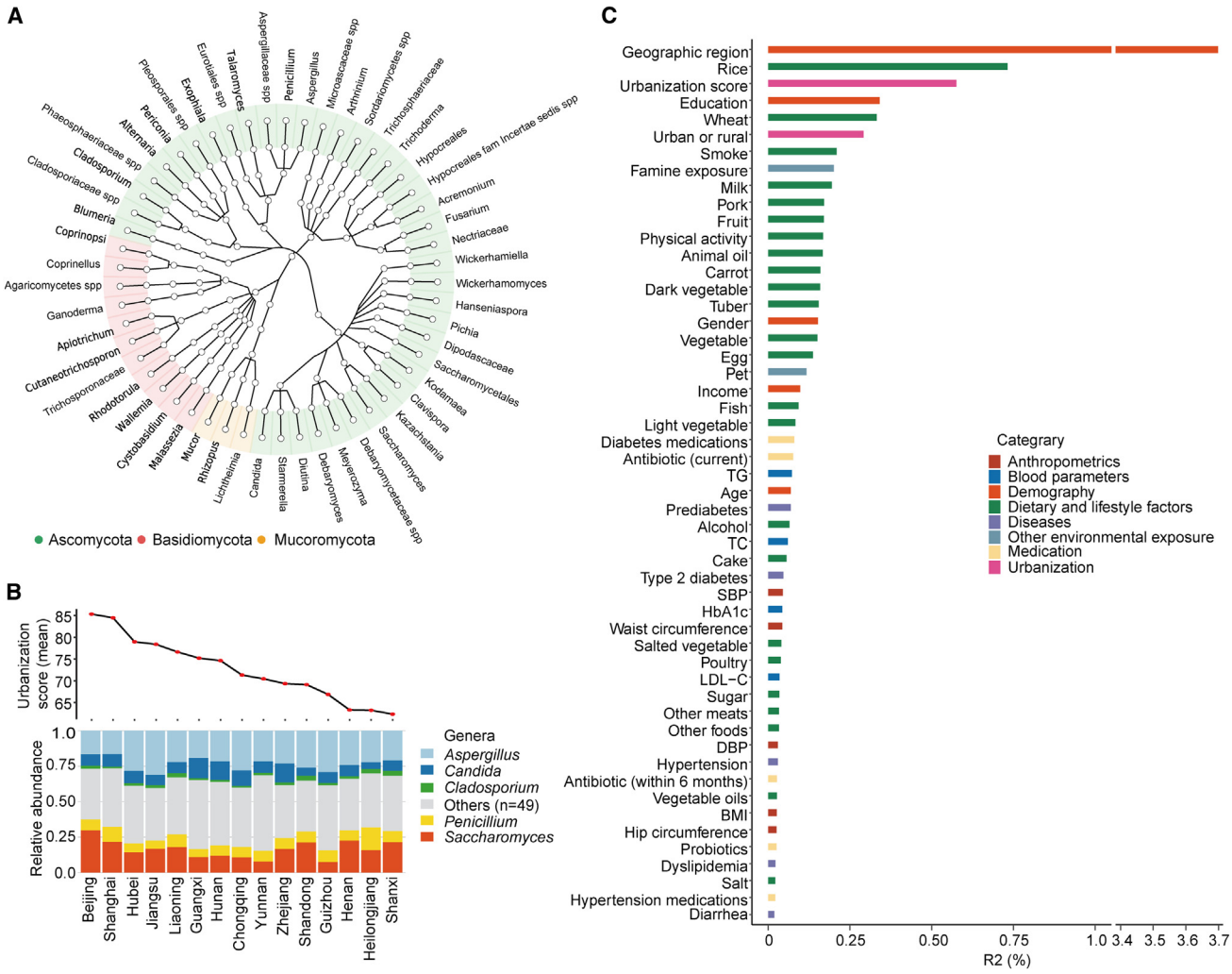


Figure 2. Landscape of the gut mycobiome and key factors explaining the gut mycobiome inter-individual variation

(A) Taxonomic tree illustrating the genera with a prevalence greater than 10% in the CHNS-1 cohort (n = 10,695). Nodes (n = 54) within the inner to outer circles depict fungal taxa from phylum to genus levels, with genus-level nodes colored based on their respective phylum assignments. Reads that could not be confidently assigned to known taxa due to insufficient matching with reference sequences were labeled with the highest taxonomic rank that could be assigned, followed by "spp."

(B) Distribution of fungal taxa that are present in at least 50% of samples at a minimum relative abundance of 0.1%.

(C) Variance in the mycobiome composition explained by phenotype categories was assessed through multivariate PERMANOVA analysis (n = 10,695). In this analysis, 54 genera with a prevalence greater than 10% were retained. The p value was determined through 999 permutations and subsequently adjusted for multiple testing. Notably, only the statistically significant results (FDR<0.05) are depicted in the figure.

See also Figures S1 and S2 and Tables S2 and S3.

samples at a minimum relative abundance of 0.1%. These genera also showed enrichment in other cohorts.[16,17,19] The distribution of these genera across geographic regions and urbanization gradients is depicted in the Figure 2B and S2A, respectively. We observed a noticeable increase in the relative abundance of fungal *Saccharomyces* with an increase in urbanization score (Figure S2A). Linear mixed-effects model revealed that the relative abundance of *Saccharomyces* increased by 0.029 (95% confidence interval [CI], 0.024 to 0.035) with per increase in urbanization quintile. Additionally, *Saccharomyces* exerted the strongest inverse association with fungal diversity (Figure S2B). Co-abundance network analysis found that *Blumeria*, *Debaryomyces*, and *Candida* formed central nodes in fungal networks (Figure S2C and Table S3).

Factors explaining inter-individual gut mycobiome variation

In the CHNS-1 cohort, we used PERMANOVA (permutational multivariate analysis of variance) analysis individually to assess the contributions of each environmental and host factor to the inter-individual variability in fungal composition and found that 52 factors displayed associations with the fungal variation (false

discovery rate [FDR]<0.05, Figure 2C). These factors included demography (5 factors), urbanization (2 factors), dietary and lifestyle factors (24 factors), anthropometric factors (5 factors), blood parameters (4 factors), diseases (5 factors), medications (5 factors), and other environmental factors (2 factors).

The geographic region emerged as the dominant factor associated with variations in the gut mycobiome, explaining 3.7% of the observed variability (Figure 2C). Urbanization, whether defined categorically (urban or rural) or continuously (urbanization score), along with staple foods such as rice and wheat intake, as well as exposure to smoking, ranked prominently among contributors to inter-individual variation in the gut mycobiome (Figure 2C).

Gut mycobiome signatures of cardiometabolic health

To elucidate the distinctive fungal profiles associated with cardiometabolic health, we examined the relationships between 58 fungal features (including 54 genera and 4 alpha diversity indices) and 3 cardiometabolic disorders, along with 8 clinical biomarkers, including measures of glycemic control, lipid levels, and blood pressure in the CHNS-1 cohort. The prevalence of these cardiometabolic conditions in the cohort is as follows: T2D at 11.1%, hypertension at 28%, and dyslipidemia at 32.4%. Notably, 46.6% of the T2D participants also suffer from hypertension, and 51.7% suffer from dyslipidemia. This comprehensive analysis yielded a total of 48 significant associations (FDR<0.05). Higher fungal alpha diversity indices showed protective associations with multiple cardiometabolic diseases (Figure 3A) and related biomarkers (Figure 3B). Furthermore, the abundance of eight specific genera exhibited noteworthy alterations in at least one of the cardiometabolic disease groups (Figure 3C). For instance, *Saccharomyces* was positively associated with T2D risk (OR [odds ration] = 1.14; 95% CI, 1.06 to 1.21). *Saccharomyces* was also positively associated with HbA1c ($p = 0.0013$), although this association did not remain significant after correcting for multiple comparisons (FDR = 0.07). *Saccharomyces* was found to be enriched in participants with type 1 diabetes[20] and multiple sclerosis[21] and has the potential to exacerbate gut inflammation as indicated by an animal study.[22] *Candida* exhibited a positive correlation with dyslipidemia (OR = 1.08; 95% CI, 1.03 to 1.12), concurrently displaying positive association with triglycerides (TGs) and systolic blood pressure (SBP), while manifesting an inverse association with high-density lipoprotein cholesterol (HDL-C) in our study. Importantly, the inverse association of alpha diversity indices with cardiometabolic diseases persisted after further adjusting for sequencing depth (Table S4). Interestingly, 16 fungal taxa exhibited an association with TG (Figure 3B), underscoring the potential role of the mycobiome in lipid metabolism and overall cardiometabolic health.

Our subsequent analyses revealed compelling evidence of a synergistic interaction between the gut *Saccharomyces* and environmental factors on T2D risk ($p_{interaction} = 0.012$). The traditional risk score combines multiple factors, including age, gender, waist circumference, BMI, and SBP, to estimate an individual's overall risk of developing T2D. We divided participants into four groups based on the median levels of T2D traditional risk score and *Saccharomyces* abundance. Comparing with participants with lower levels of both traditional risk score and *Saccharomyces*, the adjusted odds ratio (95% CI) for T2D was 1.30 (1.03–1.65) for those with higher *Saccharomyces* levels but with low traditional risk score (Figure 3D). It further increased to 4.28 (3.48–5.27) for participants exhibiting both higher levels of traditional risk scores and *Saccharomyces* (Figure 3D). Overall, these results suggest *Saccharomyces* may emerge as an effect modifier of traditional risk factors in promoting T2D risk.

Temporal variability of the gut mycobiome over time

To assess the temporal variability of the gut mycobiome, we undertook a dynamic analysis of fungal composition. This analysis focused on genera with prevalence higher than 10% in the participants with repeated mycobiome profiling data across an average period of 3.02 years (interquartile range: 3–3.07) in the CHNS-2 cohort. There was a slight increase in the inter-individual variation of the gut mycobiome over time (Figure 4A). The differences in fungal composition were more pronounced between individuals than within individuals ($p < 0.0001$; Figure 4A). We found that age exhibited inverse association with intra-individual variation in gut fungal composition across the baseline and follow-up (Figure 4B). Importantly, this association persisted even after adjusting for sequencing depth, as indicated by the Spearman partial correlation coefficient (corr = −0.058, $p = 0.01$). When considering individual fungal genera, 45.3% (34 out of 75) showed significant alterations in abundance over time (Figure S3A). Specifically, among core taxa, there was a reduction in the abundance of *Aspergillus*, *Penicillium*, and *Debaryomyces*, while *Candida* exhibited an increased over the follow-up period. Additionally, we detected an average 1.57-fold reduction in observed OTUs (operational taxonomic units) at the amplicon sequence variant (ASV) level during the follow-up compared to the baseline (Figure 4C). This observation was validated by a linear mixed model adjusted for sequencing depth (adjusted beta = −5.96, 95% CI: −8.79 to −2.40). A consistent decrease was also detected for other alpha diversity indices (Figure S3B).

We further discerned four distinct fungal clusters by applying Dirichlet mixture modeling on the fungal composition data at both time points (Figure S4A). We found that cluster 4 (C4) was remarkably different from the remaining clusters, primarily characterized by the dominance of *Saccharomyces* (Figure 4D and S4B). Characteristics of the participants at baseline, stratified by fungal clusters, are presented in Table S5. Through transition trajectory analysis, we investigated the dynamic changes in participants' fungal pattern over time and found that only a small proportion of participants exchanged within the cluster 3 and cluster 4 (Figure 4E). Those individuals who transitioned to cluster 4 during the follow-up period exhibited lower alpha diversity indices compared to those who transitioned to alternative clusters, which was found to be independent of their fungal cluster status at baseline (Figure 4F).

Dynamic interplay between the gut mycobiome, cardiometabolic diseases, and serum metabolites

To investigate the dynamic interplay between the gut mycobiome and human health in a longitudinal context, we fitted a generalized estimating equation model to account for the

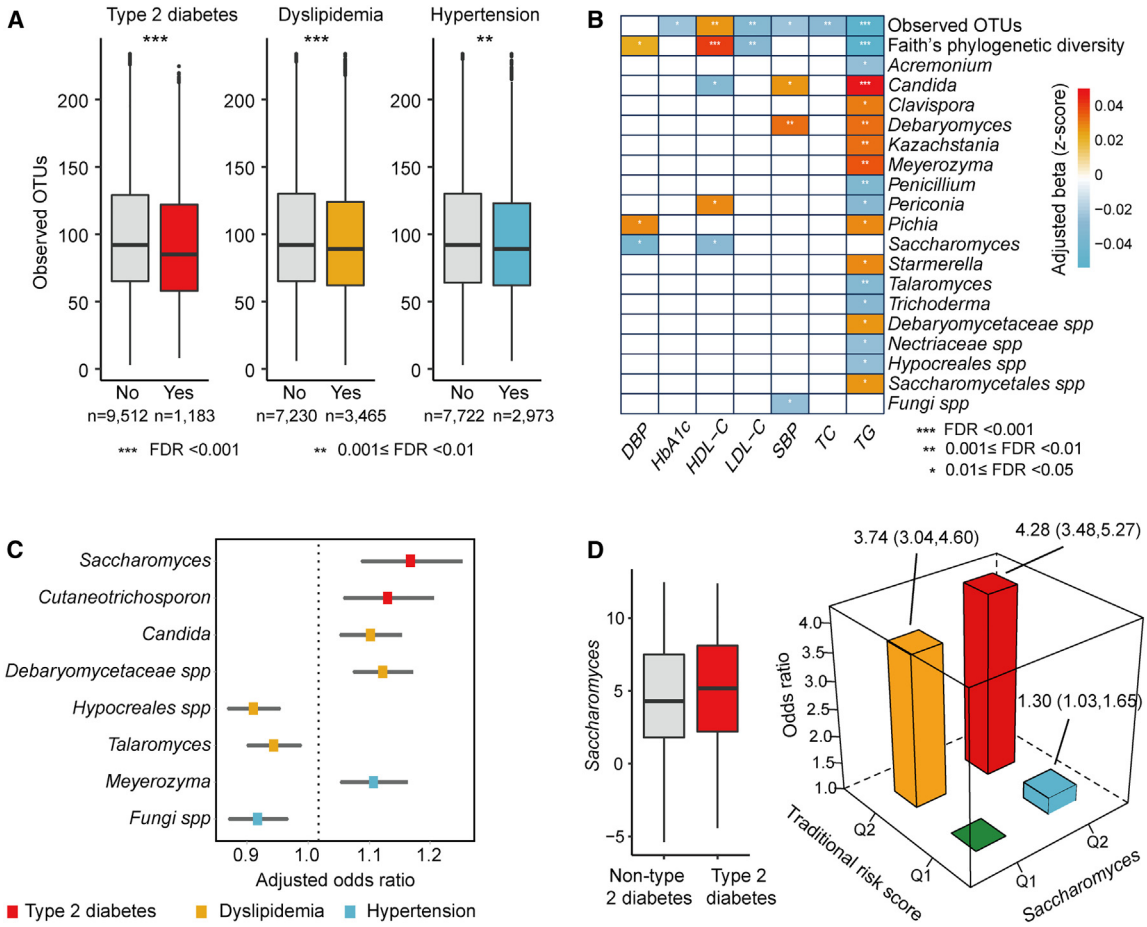


Figure 3. Gut mycobiome signatures associated with cardiometabolic health

(A) Boxplots illustrating observed OTUs in participants with different cardiometabolic diseases compared with those without (n = 10,695). Significance was obtained by logistic mixed-effects regression model, with adjustment for age, gender, and BMI. The logistic mixed-effects model included a random intercept and random coefficient for provinces or megacities, accounting for the variability in gut fungal composition across different regions.

(B) Associations of mycobiome with indicators of cardiometabolic diseases (n = 10,695). The analysis was conducted using linear mixed-effects regression model, with adjustment for age, gender, and BMI. The linear mixed-effects model included a random intercept and random coefficient for provinces or megacities, accounting for the variability in gut fungal composition across different regions. Associations are colored by direction of Z-scored effect (red, positive; blue, negative). Notably, only the statistically significant results (FDR < 0.05) are depicted in the figure.

(C) Associations of fungal genera (per SD unit) with cardiometabolic diseases (n = 10,695). Odds ratios and 95% CI were calculated using logistic mixed-effects regression model described above. Notably, only the statistically significant results (FDR < 0.05) are depicted in the figure.

(D) Left, boxplots illustrating *Saccharomyces* in participants with type 2 diabetes compared with those without. Right, interaction between *Saccharomyces* and traditional risk factors in relation to the risk of type 2 diabetes (n = 10,695). The traditional risk score combines multiple factors, including age, gender, waist circumference, BMI, and SBP, to estimate an individual's overall risk of developing T2D.

See also Table S4.

dynamic nature of the gut mycobiome and status of host cardiometabolic diseases in the CHNS-2 cohort. In comparison to participants assigned to fungal enterotype of cluster 4 (dominance of *Saccharomyces*), those assigned to cluster 3 demonstrated a significantly lower risk of cardiometabolic diseases, with odds ratio of 0.45 (0.29–0.71) for T2D and 0.67 (0.49–0.91) for hypertension (Figure 4G).

To gain a deeper understanding of the distinctive variations in serum metabolites between fungal enterotype of cluster 4 and the other clusters, we utilized the LightGBM (light gradient-boosting machine) algorithm to construct machine learning classifiers. These classifiers aimed to differentiate individuals with the fungal enterotype of cluster 4 from those with each of the other fungal enterotypes, based on their serum metabolome. Receiver operating characteristic curve analysis demonstrated that the classification for cluster 3 and cluster 4 showed the highest performance, with a mean 10-fold cross-validation area under the curve (AUC) of 0.90 (Figure 5A).

Through further analysis using orthogonal projection to latent structures discriminant analysis (OPLS-DA), we pinpointed 212 metabolites that contributed to the differentiation of cluster 3 and cluster 4 (Table S6). These distinguished metabolites were predominantly associated with fatty acid and amino acid metabolism and their derivatives. By reconstructing a classifier using

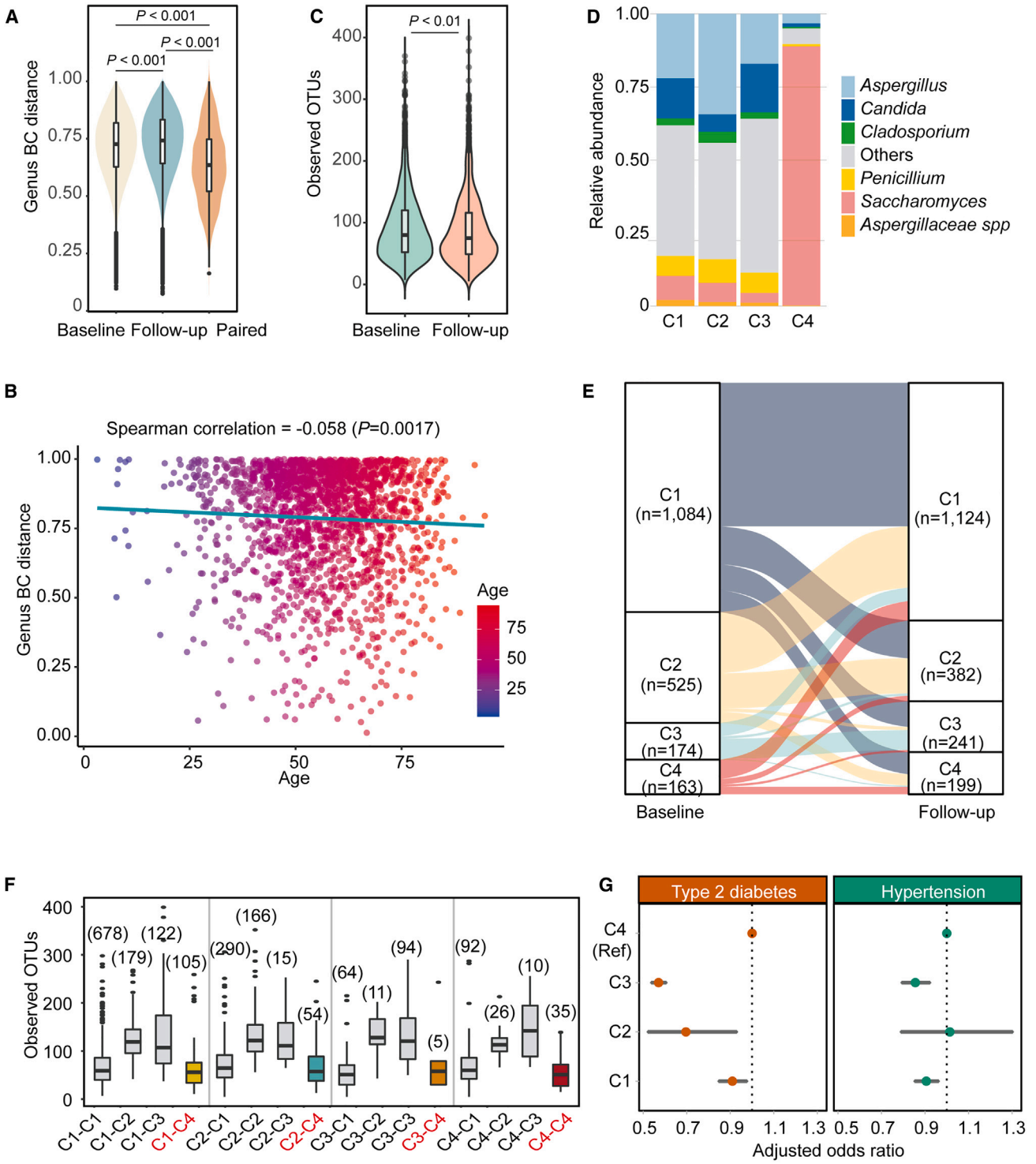


Figure 4. Dynamic interplay between the gut mycobiome and cardiometabolic diseases

(A) Comparison of inter-individual Bray-Curtis distance of fungal composition between baseline and follow-up, and the within-pair Bray-Curtis distances across the two time points (n = 1,946).

(B) Correlation between age and intra-individual variation in gut fungal composition across the baseline and follow-up (n = 1,946).

(C) Comparison of the fungal alpha diversity (observed OTUs) between baseline and follow-up. The y axis refers to the observed OTUs at the ASV level (n = 1,946).

(D) Distribution of fungal taxa that are present in at least 50% of samples at a minimum relative abundance of 0.1% across different fungal clusters (n = 1,946 pairs).

(legend continued on next page)

these OPLS-DA-identified metabolites at the baseline, we achieved a similar classification accuracy for cluster 3 and cluster 4 as that of the primary all-metabolites-based model (AUC = 0.90; Figure 5B). Importantly, this classifier was successfully validated with follow-up participants who were distinct from the baseline participants (AUC = 0.84; Figure 5B). These follow-up participants transitioned from cluster 1 or cluster 2 at baseline to cluster 3 or cluster 4 during the follow-up phase. Additionally, our analysis employing the Wilcoxon rank-sum test highlighted significant differences in the abundance of 132 metabolites between the two groups (Figure 5C).

Metabolites as functional link mediating the association of mycobiome with T2D

In the CHNS-2 cohort, to investigate the intricate interplay among the gut mycobiome, serum metabolites, and cardiometabolic health, we firstly performed a weighted gene co-expression network analysis. This analysis grouped the OPLS-DA-identified metabolites into 15 distinct metabolite modules based on the similarity of metabolism profiles. The distribution of identified metabolites across different metabolism modules is presented in Figure S5A. The inter-correlations among these metabolite modules ranged from weak to moderate (Figure S5B). The majority of these modules, specifically 9 out of the 15, displayed significant differences between two distinct fungal patterns, denoted as cluster 3 and cluster 4 (Figure S5C). This underscores the substantial variations in metabolic profiles associated with these fungal patterns. Additionally, four metabolite modules (M2, M9, M12, and M14) exhibited significant associations with risk of T2D or hypertension (Figure S5D).

We then hypothesized that serum metabolite modules might mediate the impact of the gut mycobiome on cardiometabolic diseases. We found that two metabolite modules (M14 and M9) emerged as mediators of the effects of the gut mycobiome on T2D, with mediated effects of 37% and 16%, respectively (Figure S6A). Importantly, we did not identify any significant inverse effect, where gut fungi might influence metabolites through their impact on the disease. Furthermore, a majority of the metabolites belonging to the M9 and M14 modules exhibited significant differences between the T2D cases and controls (Figure 6A). The inter-correlations among these metabolites ranged from weak to moderate (Figure S6B).

For the individual core fungal taxa and metabolites belonging to the M9 and M14 modules, we uncovered three significant results in the mediation analyses, with forward mediation effects showing an FDR <0.05 and reverse mediation effects showing an FDR >0.05. These results highlight the potential role of two fungi (*Saccharomyces* and *Blumeria*) in the development of T2D via several metabolites (γ-linolenic acid, L-valine, and LPE (lysophosphatidylethanolamine) (Figure 6B). A noteworthy instance was the mediation effect of L-valine, which accounted for 45% of the total effect of *Blumeria* on T2D ($FDR_{mediation}$ = 0.02, Figure 6C). L-valine, one of the branched-chain amino acids (BCAAs), exhibited a positive correlation with *Blumeria* and was found to be elevated in individuals with T2D in our study (Figure 6A). Consistently, L-valine has previously been implicated in causing insulin resistance in animal studies[23,24] and has shown a positive association with T2D in human cohort studies.[25]

Another example was γ-linolenic acid, which displayed a positive correlation with the abundance of *Saccharomyces* and was found to be enriched in T2D patients in our study (Figure 6A). It acted as a mediator of *Saccharomyces* in the progression of T2D ($FDR_{mediation}$ = 0.04), with a mediation effect of 16% (Figure 6D). Taken together, our findings provide potential functional insights underlying the relationships between the gut mycobiome and human cardiometabolic diseases.

DISCUSSION

Here, we systematically characterized the gut mycobiome composition and dynamics in a large nationwide cohort of Chinese adults. Leveraging a rich array of phenotypic measurements, our research builds upon existing knowledge by uncovering robust correlations between gut mycobiome composition and cardiometabolic health. Moreover, we observed temporal variability in both individual and fungal patterns of the gut mycobiome over three years, emphasizing its inherently dynamic nature. Additionally, we found the dynamic interplay between the gut mycobiome and cardiometabolic diseases. Finally, our investigations revealed the potential functional role of metabolites as mediators in the association between the gut mycobiome and T2D. These findings provide deep insights into the intricate interactions between the gut mycobiome and the host, shedding light on its relevance to cardiometabolic health.

Several prior studies have reported associations of the gut mycobiome with geographic locations and dietary and lifestyle factors, including milk, fruit, vegetable, fish, sugar, and alcohol consumption.[13,17] Our substantially expanded sample size and comprehensive phenotype data have enabled us to extend the investigation, thereby enhancing the scope of previous knowledge. Our investigation underscores the considerable impact of dietary and lifestyle factors on fungal variation, highlighting the contribution of staple food consumption (specifically wheat and rice) and physical activity. Both rice and wheat, as fundamental components of the Chinese diet, contribute substantially to caloric intake and represent distinct dietary patterns.[26]

Fungi are common inhabitants of the intestinal barrier surface. Animal studies have shown that the gut mycobiome is influenced

(E) Transition of all participants between the fungal clusters from the baseline to follow-up (n = 1,946).
(F) Distribution of alpha diversity across different fungal pattern transitions. The numbers of participants in each fungal pattern were marked on the top of the boxplot. Key combinations highlighted: C1-C4, C2-C4, C3-C4, and C4-C4. Others shaded in gray.
(G) Dynamic interplay between mycobiome and human cardiometabolic diseases (n = 1,946). Generalized estimating equation model was employed to assess the difference between other fungal clusters and cluster 4. The models were adjusted for age, gender, and BMI and used independent correlation structure as the generalized estimating equation model's working correlation matrix. Additionally, the models included both a random intercept and a random coefficient for provinces or megacities.
See also Figures S3 and S4 and Table S5.

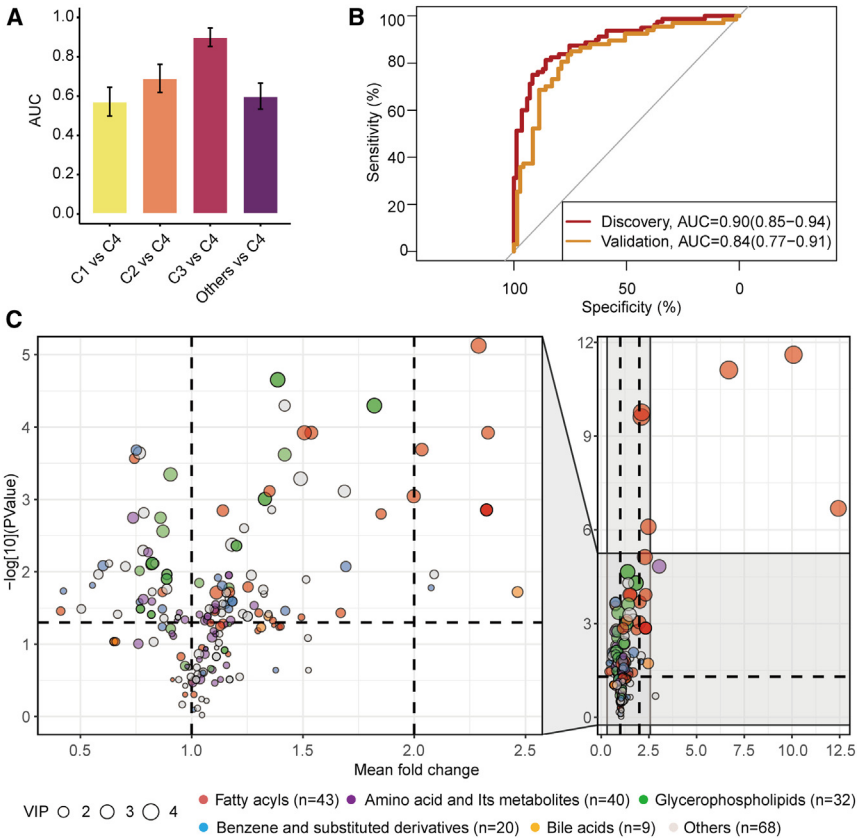


Figure 5. Association of the gut mycobiome with serum metabolites

(A) Comparing the performance of metabolome-based classifiers (LightGBM) in categorizing participants' fungal clusters at the baseline (n = 911). Performance of the classifiers was assessed by area under the receiver operating characteristic curve (AUC) across 10-folds, with cluster 4 as the reference group, and each of other groups and their combinations as the comparison group. Here, the numbers of participants included in the cluster 1 to 4 were 518, 227, 85, and 80, respectively.

(B) Receiver operating characteristic curve for the orthogonal projections to latent structures discriminant analysis (OPLS-DA)-selected metabolites in categorizing participants into either cluster 3 or cluster 4 groups within the discovery (n = 165) and validation cohorts (n = 138). The effectiveness of these selected metabolites in predicting fungal clusters was assessed using the machine learning model (LightGBM).

(C) Fold change of differential metabolites between the fungal cluster 3 and cluster 4 (n = 165). The identified metabolites are visualized as dots, with colors corresponding to their respective categories. The dot size indicates the importance of the metabolite in distinguishing fungal clusters during the OPLS-DA analysis.

See also Figure S5 and Table S6.

by the environment and correlates with metabolic outcomes in response to diet.[27] Dysbiosis in fungal communities can affect mucus production, epithelial function, and immune defense, potentially contributing to the development of metabolic diseases.[6] While higher levels of gut bacterial alpha diversity have consistently been associated with improved health status,[28,29] evidence for fungal diversity remains limited. In the present study, we found that fungal alpha diversity was inversely associated with multiple cardiometabolic diseases and risk factors. A recent study[8] reported that fungal alpha diversity was lower in participants with T2D and hypertension based on public fecal metagenomes. However, the association was not significant, likely due to the small sample size (n < 350) and technical challenges in detecting fungi using metagenomic methods.[3] We also observed that gut *Saccharomyces* was associated with T2D but not with dyslipidemia or hypertension, suggesting that *Saccharomyces* might influence T2D through mechanisms that were independent of those typically associated with dyslipidemia or hypertension. This was consistent with previous research, which also found that no single fungal genus was consistently associated with different cardiometabolic diseases.[8]

There was a mycobiome-dependent association between the traditional risk factors and T2D, where a stronger association was observed in individuals with enrichment of *Saccharomyces*. These findings suggest the importance of considering *Saccharomyces* as a potential effect modifier of traditional risk factors in the prevention of T2D risk. Importantly, the link between *Saccharomyces* and T2D is further supported by results from our independently conducted dynamic analysis. This reinforces the potential role of *Saccharomyces* as a predictive biomarker of T2D. It is crucial to note that these findings warrant further investigation to elucidate the underlying mechanisms and biological pathways by which *Saccharomyces* may influence T2D risk. Additionally, considering the potential role of *Saccharomyces* in modifying the impact of traditional risk factors may open avenues for personalized approaches to diabetes prevention and management.

Our mediation analysis provided evidence that several metabolites may play a mediating role in the relationship between the gut mycobiome and T2D. Of particular interest, we demonstrated that L-valine, a BCAA, may mediate the association between *Blumeria* and T2D. BCAAs are synthesized in bacteria, plant, and fungi and are well-known contributors to insulin resistance.[23,24] Additionally, γ-linolenic, an omega-6 fatty acid, exhibited a positive association with T2D, aligning with previous findings.[30] Furthermore, our observations suggest that γ-linolenic acid might mediate the impact of *Saccharomyces* on T2D. These associations between specific fungi and metabolites with T2D provide valuable insights into the potential mechanism by which gut mycobiota may contribute to the pathogenesis of the complex metabolic disorders. Further investigations into these specific pathways may shed light on therapeutic targets and interventions for T2D and potentially other cardiometabolic diseases. Nevertheless, additional studies are needed to validate and elucidate the underlying mechanism of these intriguing findings.

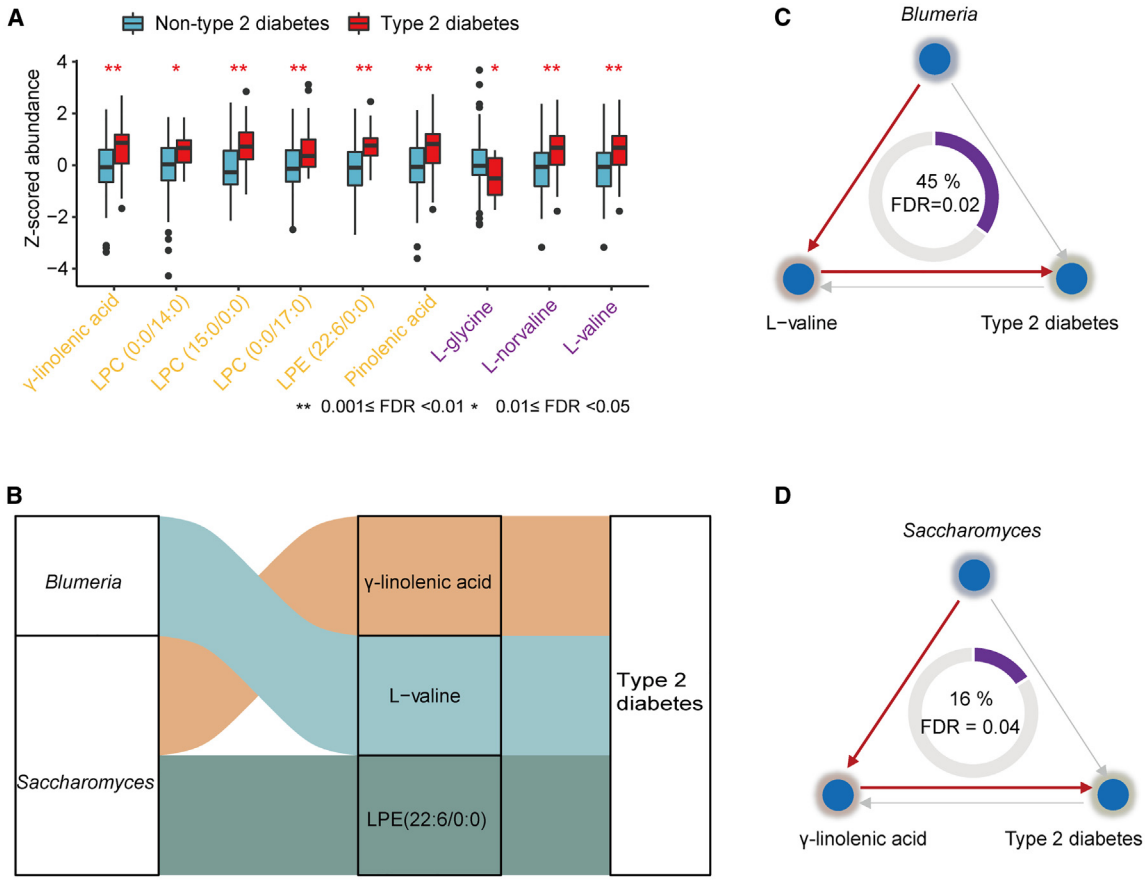


Figure 6. Functional metabolites mediate the effect of the gut mycobiome on type 2 diabetes

(A) Association of serum metabolites with type 2 diabetes (n = 165). Boxplot shows the distribution of mycobiome-related metabolites abundance among the type 2 diabetes cases and healthy controls. Significance was examined by logistic regression, with adjustment for age, gender, and BMI. The names of the metabolites are color-coded based on their belonging to specific metabolite modules: purple represented M9, and yellow corresponded to M14.

(B) Sankey diagram showing the significant mediation linkages (FDR < 0.05) where fungal genera acted as effect factors, metabolites as mediators, and type 2 diabetes as outcome (n = 165).

(C) Mediation linkages among the *Blumeria*, L-Valine, and type 2 diabetes (n = 165). The proportion of mediation effect is displayed at the center of the ring chart.

(D) As in (C), but for *Saccharomyces*, γ-linolenic acid, and type 2 diabetes (n = 165).

See also Figure S6.

In summary, our study offers a comprehensive characterization of the gut mycobiome composition and dynamics in a nationwide Chinese cohort. By extensively mapping the gut mycobiome to environmental factors, we have created a valuable resource for unraveling the factors associated with inter-individual variation of the gut mycobiome. We reveal the inter-connections between the gut mycobiome and cardiometabolic health, underscoring gut *Saccharomyces* as a potential predictive biomarker of T2D risk. This study advances our understanding of the roles of the gut mycobiome in cardiometabolic health and provides potential targets for intervention and therapeutics aimed at modulating the gut mycobiome.

Limitations of the study

Our study has several limitations. Firstly, our findings regarding the associations of the gut mycobiome with cardiometabolic health are observational, without the establishment of causation. Future research is needed to confirm the potential causal relationships. Secondly, while mediation analysis has been utilized in observational microbiome studies,[31–33] further animal studies are necessary to elucidate the underlying mechanisms and potential causal direction. Thirdly, current fungal databases are still immature, particularly for less-studied and non-culturable fungi, which results in many reads being classified as "unknown/unclassified." Additionally, primer design can introduce biases affecting amplification efficiency.[3] Future research should prioritize developing validated methodologies, including refining DNA extraction techniques, improving the quality of fungal databases, and creating more effective primers. Finally, all participants included in the present study are Chinese, so caution should be exercised in extrapolating our findings to other ethnic groups.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to the lead contact, Prof. Ju-Sheng Zheng (zhengjusheng@westlake.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The raw ITS sequencing data have been deposited in the Genome Sequence Archive (GSA) (https://ngdc.cncb.ac.cn/gsa/) at the accession number CRA013806.
- Due to informed consent regulations, phenotypic data for the cohort are available upon request from the CHNS (https://www.cpc.unc.edu/projects/china/).
- The computer codes for main analyses in this study are deposited in GitHub under https://github.com/wenutrition/CHNS-fungi-cardiometabolic-diseases. The DOI at Zenodo is https://doi.org/10.5281/zenodo.13353503. The software and packages used for this analysis are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this work is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank the High-Performance Computing Center and High-Throughput Core Facility at Westlake University for assistance in computing and data generation. This study was funded by the "Pioneer" and "Leading goose" R&D Program of Zhejiang (2024SSYS0032), the National Key R&D Program of China (2022YFA1303900), the National Natural Science Foundation of China (82073529, 82404242, 82103826, and U21A20427), the Research Program of Westlake Laboratory of Life Sciences and Biomedicine (202208012), and China Postdoctoral Science Foundation (2023M733177, 2022M722833). The China Health and Nutrition Survey received funding from the National Institutes of Health (NIH) (R01HD30880, R01AG065357, P30DK056350, and R01HD38700) from 1989 to 2019 and was supported by the National Institutes of Health and National Institute of Diabetes and Digestive and Kidney Diseases (R01DK104371) and the Carolina Population Center (P2CHD050924 and P30AG066615). The funders had no role in collecting data, study design, interpretation of data, or the decision to submit the manuscript for publication.

AUTHOR CONTRIBUTIONS

J.-S.Z., B.Z., and H.W. conceived the study. B.Z. and H.W. established the human cohorts. J.-S.Z. and W.G. designed and performed the statistical analyses. H.W., J.Z., W.D., X.J., and C.S. collected the data. W.G., C.S., W.D., J.Z., K.Z., X.W., Z.M., C.G., and Y.F. contributed to data curation. W.G. and J.-S.Z. wrote the first draft of the manuscript. W.G., H.W., C.S., and Y.F. contributed equally to the work. All authors contributed to the discussion and critical revision of the manuscript and approved the final draft.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2024.101775.

Received: June 14, 2024
Revised: July 30, 2024
Accepted: September 13, 2024
Published: October 4, 2024

REFERENCES

1. Peay, K.G., Kennedy, P.G., and Talbot, J.M. (2016). Dimensions of biodiversity in the Earth mycobiome. Nat. Rev. Microbiol. *14*, 434–447. https://doi.org/10.1038/nrmicro.2016.59.
2. Iliev, I.D., and Leonardi, I. (2017). Fungal dysbiosis: immunity and interactions at mucosal barriers. Nat. Rev. Immunol. *17*, 635–646. https://doi.org/10.1038/nri.2017.55.
3. Richard, M.L., and Sokol, H. (2019). The gut mycobiota: insights into analysis, environmental interactions and role in gastrointestinal diseases. Nat. Rev. Gastroenterol. Hepatol. *16*, 331–345. https://doi.org/10.1038/s41575-019-0121-2.
4. Saftien, A., Puschhof, J., and Elinav, E. (2023). Fungi and cancer. Gut *72*, 1410–1425. https://doi.org/10.1136/gutjnl-2022-327952.
5. Stefan, N., and Schulze, M.B. (2023). Metabolic health and cardiometabolic risk clusters: implications for prediction, prevention, and treatment. Lancet Diabetes Endocrinol. *11*, 426–440. https://doi.org/10.1016/S2213-8587(23)00086-4.
6. Wang, L., Zhang, K., Zeng, Y., Luo, Y., Peng, J., Zhang, J., Kuang, T., and Fan, G. (2023). Gut mycobiome and metabolic diseases: The known, the unknown, and the future. Pharmacol. Res. *193*, 106807. https://doi.org/10.1016/j.phrs.2023.106807.
7. Wheeler, M.L., Limon, J.J., Bar, A.S., Leal, C.A., Gargus, M., Tang, J., Brown, J., Funari, V.A., Wang, H.L., Crother, T.R., et al. (2016). Immunological Consequences of Intestinal Fungal Dysbiosis. Cell Host Microbe *19*, 865–873. https://doi.org/10.1016/j.chom.2016.05.003.
8. Yan, Q., Li, S., Yan, Q., Huo, X., Wang, C., Wang, X., Sun, Y., Zhao, W., Yu, Z., Zhang, Y., et al. (2024). A genomic compendium of cultivated human gut fungi characterizes the gut mycobiome and its relevance to common diseases. Cell *187*, 2969–2989.e24. https://doi.org/10.1016/j.cell.2024.04.043.
9. Van Syoc, E., Nixon, M.P., Silverman, J.D., Luo, Y., Gonzalez, F.J., Elbere, I., Klovins, J., Patterson, A.D., Rogers, C.J., and Ganda, E. (2024). Changes in the type 2 diabetes gut mycobiome associate with metformin treatment across populations. mBio *15*, e0016924. https://doi.org/10.1128/mbio.00169-24.
10. Jayasudha, R., Das, T., Kalyana Chakravarthy, S., Sai Prashanthi, G., Bhargava, A., Tyagi, M., Rani, P.K., Pappuru, R.R., and Shivaji, S. (2020). Gut mycobiomes are altered in people with type 2 Diabetes Mellitus and Diabetic Retinopathy. PLoS One *15*, e0243077. https://doi.org/10.1371/journal.pone.0243077.
11. Qiu, J., Zhao, L., Cheng, Y., Chen, Q., Xu, Y., Lu, Y., Gao, J., Lei, W., Yan, C., Ling, Z., and Wu, S. (2023). Exploring the gut mycobiome: differential composition and clinical associations in hypertension, chronic kidney disease, and their comorbidity. Front. Immunol. *14*, 1317809. https://doi.org/10.3389/fimmu.2023.1317809.

12. Zou, Y., Ge, A., Lydia, B., Huang, C., Wang, Q., and Yu, Y. (2022). Gut mycobiome dysbiosis contributes to the development of hypertension and its response to immunoglobulin light chains. Front. Immunol. *13*, 1089295. https://doi.org/10.3389/fimmu.2022.1089295.

13. Sun, Y., Zuo, T., Cheung, C.P., Gu, W., Wan, Y., Zhang, F., Chen, N., Zhan, H., Yeoh, Y.K., Niu, J., et al. (2021). Population-Level Configurations of Gut Mycobiome Across 6 Ethnicities in Urban and Rural China. Gastroenterology *160*, 272–286.e11. https://doi.org/10.1053/j.gastro.2020.09.014.

14. Popkin, B.M., Du, S., Zhai, F., and Zhang, B. (2010). Cohort Profile: The China Health and Nutrition Survey–monitoring and understanding socio-economic and health change in China, 1989-2011. Int. J. Epidemiol. *39*, 1435–1440. https://doi.org/10.1093/ije/dyp322.

15. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. *37*, 852–857. https://doi.org/10.1038/s41587-019-0209-9.

16. Nash, A.K., Auchtung, T.A., Wong, M.C., Smith, D.P., Gesell, J.R., Ross, M.C., Stewart, C.J., Metcalf, G.A., Muzny, D.M., Gibbs, R.A., et al. (2017). The gut mycobiome of the Human Microbiome Project healthy cohort. Microbiome *5*, 153. https://doi.org/10.1186/s40168-017-0373-4.

17. Szóstak, N., Handschuh, L., Samelak-Czajka, A., Tomela, K., Schmidt, M., Pruss, Ł., Milanowska-Zabel, K., Kozlowski, P., and Philips, A. (2023). Host Factors Associated with Gut Mycobiome Structure. mSystems *8*, e0098622. https://doi.org/10.1128/msystems.00986-22.

18. Shuai, M., Fu, Y., Zhong, H.-L., Gou, W., Jiang, Z., Liang, Y., Miao, Z., Xu, J.J., Huynh, T., Wahlqvist, M.L., et al. (2022). Mapping the human gut mycobiome in middle-aged and elderly adults: multiomics insights and implications for host metabolic health. Gut *71*, 1812–1820. https://doi.org/10.1136/gutjnl-2021-326298.

19. Ahmad, H.F., Castro Mejia, J.L., Krych, L., Khakimov, B., Kot, W., Bechshøft, R.L., Reitelseder, S., Højfeldt, G.W., Engelsen, S.B., Holm, L., and Nielsen, D.S. (2020). Gut Mycobiome dysbiosis is linked to hypertriglyceridemia among home Dwelling elderly Danes. Preprint at bioRxiv. https://doi.org/10.1101/2020.04.16.044693.

20. Kowalewska, B., Zorena, K., Szmigiero-Kawko, M., Wąż, P., and Myśliwiec, M. (2016). Higher diversity in fungal species discriminates children with type 1 diabetes mellitus from healthy control. Patient Prefer. Adherence *10*, 591–599. https://doi.org/10.2147/PPA.S97852.

21. Shah, S., Locca, A., Dorsett, Y., Cantoni, C., Ghezzi, L., Lin, Q., Bokoliya, S., Panier, H., Suther, C., Gormley, M., et al. (2021). Alterations of the gut mycobiome in patients with MS. EBioMedicine *71*, 103557. https://doi.org/10.1016/j.ebiom.2021.103557.

22. Chiaro, T.R., Soto, R., Zac Stephens, W., Kubinak, J.L., Petersen, C., Gogokhia, L., Bell, R., Delgado, J.C., Cox, J., Voth, W., et al. (2017). A member of the gut mycobiota modulates host purine metabolism exacerbating colitis in mice. Sci. Transl. Med. *9*, eaaf9044. https://doi.org/10.1126/scitranslmed.aaf9044.

23. Pedersen, H.K., Gudmundsdottir, V., Nielsen, H.B., Hyotylainen, T., Nielsen, T., Jensen, B.A.H., Forslund, K., Hildebrand, F., Prifti, E., Falony, G., et al. (2016). Human gut microbes impact host serum metabolome and insulin sensitivity. Nature *535*, 376–381. https://doi.org/10.1038/nature18646.

24. Neinast, M., Murashige, D., and Arany, Z. (2019). Branched Chain Amino Acids. Annu. Rev. Physiol. *81*, 139–164. https://doi.org/10.1146/annurev-physiol-020518-114455.

25. Morze, J., Wittenbecher, C., Schwingshackl, L., Danielewicz, A., Rynkiewicz, A., Hu, F.B., and Guasch-Ferré, M. (2022). Metabolomics and Type 2 Diabetes Risk: An Updated Systematic Review and Meta-analysis of Prospective Cohort Studies. Diabetes Care *45*, 1013–1024. https://doi.org/10.2337/dc21-1705.

26. Xu, K., Zhang, B., Liu, Y., Mi, B., Wang, Y., Shen, Y., Shi, G., Dang, S., Liu, X., and Yan, H. (2022). Staple Food Preference and Obesity Phenotypes: The Regional Ethnic Cohort Study in Northwest China. Nutrients *14*, 5243. https://doi.org/10.3390/nu14245243.

27. Mims, T.S., Abdallah, Q.A., Stewart, J.D., Watts, S.P., White, C.T., Rousselle, T.V., Gosain, A., Bajwa, A., Han, J.C., Willis, K.A., and Pierre, J.F. (2021). The gut mycobiome of healthy mice is shaped by the environment and correlates with metabolic outcomes in response to diet. Commun. Biol. *4*, 281. https://doi.org/10.1038/s42003-021-01820-z.

28. Chen, Z., Radjabzadeh, D., Chen, L., Kurilshikov, A., Kavousi, M., Ahmadizar, F., Ikram, M.A., Uitterlinden, A.G., Zhernakova, A., Fu, J., et al. (2021). Association of Insulin Resistance and Type 2 Diabetes With Gut Microbial Diversity: A Microbiome-Wide Analysis From Population Studies. JAMA Netw. Open *4*, e2118811. https://doi.org/10.1001/jamanetworkopen.2021.18811.

29. Aasmets, O., Krigul, K.L., Lüll, K., Metspalu, A., and Org, E. (2022). Gut metagenome associations with extensive digital health data in a volunteer-based Estonian microbiome cohort. Nat. Commun. *13*, 869. https://doi.org/10.1038/s41467-022-28464-9.

30. Miao, Z., Lin, J.S., Mao, Y., Chen, G.D., Zeng, F.F., Dong, H.L., Jiang, Z., Wang, J., Xiao, C., Shuai, M., et al. (2020). Erythrocyte n-6 polyunsaturated fatty acids, gut microbiota and incident type 2 diabetes: a prospective cohort study. Diabetes Care *43*, 2435–2443. https://doi.org/10.1101/2020.03.29.20039693.

31. Lin, L., Yi, X., Liu, H., Meng, R., Li, S., Liu, X., Yang, J., Xu, Y., Li, C., Wang, Y., et al. (2023). The airway microbiome mediates the interaction between environmental exposure and respiratory health in humans. Nat. Med. *29*, 1750–1759. https://doi.org/10.1038/s41591-023-02424-2.

32. Chen, L., Wang, D., Garmaeva, S., Kurilshikov, A., Vich Vila, A., Gacesa, R., Sinha, T., Lifelines Cohort Study; Segal, E., Weersma, R.K., et al. (2021). The long-term genetic stability and individual specificity of the human gut microbiome. Cell *184*, 2302–2315.e12. https://doi.org/10.1016/j.cell.2021.03.024.

33. Chen, L., Zhernakova, D.V., Kurilshikov, A., Andreu-Sánchez, S., Wang, D., Augustijn, H.E., Vich Vila, A., Lifelines Cohort Study; Weersma, R.K., Medema, M.H., et al. (2022). Influence of the microbiome, diet and genetics on inter-individual variation in the human plasma metabolome. Nat. Med. *28*, 2333–2343. https://doi.org/10.1038/s41591-022-02014-8.

34. Op De Beeck, M., Lievens, B., Busschaert, P., Declerck, S., Vangronsveld, J., and Colpaert, J.V. (2014). Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. PLoS One *9*, e97629. https://doi.org/10.1371/journal.pone.0097629.

35. Peschel, S., Müller, C.L., von Mutius, E., Boulesteix, A.L., and Depner, M. (2021). NetCoMi: network construction and comparison for microbiome data in R. Brief. Bioinform. *22*, bbaa290. https://doi.org/10.1093/bib/bbaa290.

36. Zhang, X., Zhang, J., Du, W., Su, C., Ouyang, Y., Huang, F., Jia, X., Li, L., Bai, J., Zhang, B., et al. (2021). Multi-Trajectories of Macronutrient Intake and Their Associations with Obesity among Chinese Adults from 1991 to 2018: A Prospective Study. Nutrients *14*, 13. https://doi.org/10.3390/nu14010013.

37. Ng, S.W., Howard, A.-G., Wang, H.J., Su, C., and Zhang, B. (2014). The physical activity transition among adults in China: 1991-2011. Obes. Rev. *15*, 27–36. https://doi.org/10.1111/obr.12127.

38. Jones-Smith, J.C., and Popkin, B.M. (2010). Understanding community context and adult health changes in China: development of an urbanicity scale. Soc. Sci. Med. *71*, 1436–1446. https://doi.org/10.1016/j.socscimed.2010.07.027.

39. American Diabetes Association (2013). Diagnosis and Classification of Diabetes Mellitus. Diabetes Care *36*, S67–S74. https://doi.org/10.2337/dc13-S067.

40. 2016 Chinese guidelines for the management of dyslipidemia in adults (2018). Journal of geriatric cardiology *15*, 1–29. https://doi.org/10.11909/j.issn.1671-5411.2018.01.011.

41. Mills, K.T., Stefanescu, A., and He, J. (2020). The global epidemiology of hypertension. Nat. Rev. Nephrol. *16*, 223–237. https://doi.org/10.1038/s41581-019-0244-2.

42. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583. https://doi.org/10.1038/nmeth.3869.

43. Palarea-Albaladejo, J., and Martín-Fernández, J.A. (2015). zCompositions—R package for multivariate imputation of left-censored data under a compositional approach. Chemometr. Intell. Lab. Syst. *143*, 85–96. https://doi.org/10.1016/j.chemolab.2015.02.019.

44. Zhou, X., Qiao, Q., Ji, L., Ning, F., Yang, W., Weng, J., Shan, Z., Tian, H., Ji, Q., Lin, L., et al. (2013). Nonlaboratory-based risk assessment algorithm for undiagnosed type 2 diabetes developed on a nation-wide diabetes survey. Diabetes Care *36*, 3944–3952. https://doi.org/10.2337/dc13-0593.

45. Holmes, I., Harris, K., and Quince, C. (2012). Dirichlet multinomial mixtures: generative models for microbial metagenomics. PLoS One *7*, e30126. https://doi.org/10.1371/journal.pone.0030126.

46. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree (NIPS). https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

47. Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinf. *9*, 559. https://doi.org/10.1186/1471-2105-9-559.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Human feces	This paper	N/A
Critical commercial assays		
E.Z.N.A.® Soil DNA Kit	Omega Bio-tek, Norcross, GA, U.S.	Cat# D5625-01
Deposited data		
ITS2 rRNA sequencing data	This paper	Genome Sequence Archive (GSA): CRA013806
Metadata	This paper	https://www.cpc.unc.edu/projects/china/
Oligonucleotides		
ITS2 rRNA Forward Primer GCATCGATGAAGAACGCAGC	Op De et al.[34]	ITS3F
ITS2 rRNA Reverse Primer TCCTCCGCTTATTGATATGC	Op De et al.[34]	ITS4R
Software and algorithms		
Stata package v15.0	StataCorp LLC, TX, USA	https://www.stata.com
NetCoMi R package v.1.0.2	Peschel et al.[35]	https://github.com/stefpeschel/NetCoMi?tab=readme-ov-file
vegan R package v.2.5.7	CRAN	https://cran.r-project.org/web/packages/vegan/index.html
microbiome R package v.1.19.1	bioconductor	https://www.bioconductor.org/packages/release/bioc/html/microbiome.html
lightgbm Python package v.3.2.1	sklearn	https://lightgbm.readthedocs.io/en/stable/
ropls R package v.1.26.4	bioconductor	https://bioconductor.org/packages/release/bioc/html/ropls.html
WGCNA R package v.1.70.3	CRAN	https://cran.r-project.org/web/packages/WGCNA/index.html
mediation R package v.4.5.0	CRAN	https://cran.r-project.org/web/packages/mediation/index.html
QIIME 2	Bolyen et al.[15]	https://qiime2.org/
The computer code for main analyses	This paper	https://github.com/wenutrition/CHNS-fungi-cardiometabolic-diseases Zenodo: https://doi.org/10.5281/zenodo.13353503

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study cohorts

The present study was based on data from the China Health and Nutrition Survey (CHNS). CHNS is nationwide longitudinal cohort within China, distinguished by its utilization of a multistage, random cluster design.[14] This intricate design ensures the comprehensive coverage of a diverse range of critical public health risk factors, health outcomes, dietary and lifestyle components, as well as demographic and economic factors. The CHNS rounds have been completed in 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011, 2015 and 2018. Notably, during the 2015 and 2018 rounds of the CHNS, stool samples were collected for the measurement of the gut mycobiome. In this study, we included two sub-cohorts of the CHNS, identified as CHNS-1 and CHNS-2. In the CHNS-1 cohort, a total of 10,865 participants from 12 provinces (Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, Shandong, Shanxi, Yunnan, Zhejiang) and 3 municipalities (Beijing, Chongqing, Shanghai) provided stool samples for the measurement of ITS rRNA. After excluding participants with low depth of sequencing reads (<20,000, $n = 32$) and those without detailed metadata ($n = 138$), 10,695 participants (mean age 51.3 ± 15.5 years, 54% were women) were included in the CHNS-1 cohort. They were utilized to comprehensively profile the variations in the mycobiome and establish associations between the gut mycobiome and cardiometabolic health.

After adopting the same inclusion and exclusion criteria as used in the CHNS-1 cohort, we included an additional sub-cohort of independent participants (CHNS-2) from four prominent provinces in China (Liaoning, Guangxi, Jiangsu, Hubei), with repeatedly collected metadata (n = 1,946 pairs, mean age 56.5 ± 13.7 years, 55% were women), mycobiome data (1,946 pairs), and metabolome data (n = 996 pairs) over a period of 3 years. This sub-cohort was utilized to explore the inherent long-term variability of the gut mycobiome and to provide a deeper biological insight into the intricate interplay between the gut mycobiome, metabolome and cardiometabolic diseases.

Ethical approval

The CHNS protocol was approved by the Institutional Review Boards of the Chinese Center for Disease Control and Prevention (No. 201524), the Westlake University (No. 20220429ZJS002), the University of North Carolina at Chapel Hill, USA, and the US National Institute for Nutrition and Health (No. 07–1963). Informed consent was obtained from all participants.

METHOD DETAILS

Metadata collection

We collected phenotypes across a diverse range of categories, including demographic factors, dietary and lifestyle components, anthropometric factors, chronic diseases, medications, blood parameters, urbanization-related indices, and various other environmental factors. Demographic, medication, lifestyle, and dietary data were collected through standard questionnaires during home visits over three consecutive days. Habitual dietary intakes were evaluated using three successive 24-h dietary recalls, including two on weekdays and one on a weekend day. Trained investigators guided participants in reporting the specific types and quantities of all foods consumed within the preceding 24 h.[36] Physical activity was assessed as a total metabolic equivalent (MET) for task hours per week.[37] Anthropometric factors were measured on-site by trained staff.

We utilized the urbanization index as a comprehensive measure to assess the degree of urbanization within each community. This index comprises 12 community indicators, encompassing factors such as population density, economic activity, presence of traditional and modern markets, transportation and health infrastructure, sanitation facilities, communication access, availability of social services, cultural diversity, and housing conditions. Each of these 12 components was computed based on the presence of relevant infrastructure, the percentage of households utilizing these services within the community, and assigned a maximum score of 10 (ranging from 0 to 10). The methodology for scoring algorithms, threshold values are elaborated elsewhere.[38]

Following an overnight fasting, blood samples were collected via venipuncture. All samples were analyzed in a national central lab in Beijing (medical laboratory accreditation certificate ISO 15189:2007) with strict quality control. Blood glucose levels were measured using a glucose oxidase phenol 4-aminoantipyrine peroxidase kit (Randox, Crumlin, UK) and a Hitachi 7,600 Analyzer (Hitachi, Tokyo, Japan). Glycated hemoglobin (HbA1c) was measured via high-performance liquid chromatography system (model HLC-723 G7; Tosoh Corporation, Tokyo, Japan). All blood lipid measures were on the Hitachi 7,600 automated analyzer (Hitachi Inc., Tokyo, Japan).

Assessments of cardiometabolic diseases

Type 2 diabetes (T2D) cases were ascertained based on fasting blood glucose $\geq$7.0 mmol/L or HbA1c $\geq$ 47.5 mmol/mol (6.5%), or being currently under medical treatment for T2D during the collection of stool samples, according to the American Diabetes Association criteria for the diagnosis of diabetes.[39] Dyslipidemia cases were ascertained based on TC $\geq$ 240 mg/dL (6.2 mmol/L) or TG $\geq$ 200 mg/dL (2.3 mmol/L) or LDL-C $\geq$ 160 mg/dL (4.1 mmol/L) or HDL-C as <40 mg/dL (1.0 mmol/L), according to the guidelines for the prevention and treatment of dyslipidemia in Chinese adults.[40] Hypertension was defined as systolic blood pressure over 140 mmHg, diastolic blood pressure over 90 mmHg, or taking antihypertensive medicine.[41]

Fecal sample collection and gut mycobiome analyses

Stool samples were collected by the participants themselves, who received instructions for the collection and storage process, and were immediately frozen at −20°C refrigerators after collection. All stool samples were transported through a cold chain to the central laboratory within 1–2 days and stored at −40°C until processing.

Total fungal genomic DNA was extracted using the E.Z.N.A. soil DNA Kit (Omega Bio-tek, Norcross, GA, U.S.) according to manufacturer's instructions. The quality and concentration of extracted DNA were determined using 1.0% agarose gel electrophoresis and a NanoDrop ND-2000 spectrophotometer (Thermo Scientific Inc., USA). The hypervariable region of ITS2 was amplified using the primer pairs ITS3F (GCATCGATGAAGAACGCAGC) and ITS4R (TCCTCCGCTTATTGATATGC)[34] by an ABI GeneAmp 9700 PCR thermocycler (ABI, CA, USA). The PCR reaction mixture including 4 μL of 5 × Fast Pfu buffer, 2 μL of 2.5 mM dNTPs, 0.8 μL of forward and reverse primer (5 μM), 0.4 μL of Fast Pfu polymerase, and 10 ng of template DNA. The volume was then brought up to 20 μL with ddH2O. The PCR was carried out under the following conditions: initial denaturation for 3 min at 95°C, followed by 35 cycles of denaturing for 30 s at 95°C, annealing for 30 s at 55°C, and extension for 45 s at 72°C, and single extension at 72°C for 10 min, and finally hold at 10°C. The PCR product was extracted from a 2% agarose gel and purified using the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union City, CA, USA) according to manufacturer's instructions, and quantified using Quantus Fluorometer (Promega, USA).

The gut mycobiome was profiled by sequencing of Internal Transcribed Spacer (ITS). The hypervariable region of ITS2 was amplified using the primer pairs ITS3F (GCATCGATGAAGAACGCAGC) and ITS4R (TCCTCCGCTTATTGATATGC) by an ABI GeneAmp 9700 PCR thermocycler (ABI, CA, USA). Purified amplicons were pooled in equimolar amounts and paired-end sequenced on NovaSeq 6000 platform (Illumina, San Diego,USA) according to the standard protocols by Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China). We obtained an average of 145,675 ± 65,187 reads/sample (mean ± SD) in the CHNS-1 cohort and 226,503 ± 65,272 reads in the CHNS-2 cohort.

The raw sequencing reads were further processed with the Quantitative Insights Into Microbial Ecology 2 platform (QIIME 2).[15] In summary, DADA2[42] was used to filter sequencing reads with quality score Q < 25 and to denoise reads into amplicon sequence variants (ASVs), resulting in feature tables and representative sequences. The ASV features that were presented in only one sample were excluded before conducting the taxonomy analysis. Taxonomy was assigned to ASVs based on the UNITE (version 8.2, 99%) database using the VSEARCH tool wrapped in QIIME2. Reads that could not be confidently assigned to known taxa due to insufficient matching with reference sequences were labeled with the highest taxonomic rank that could be assigned, followed by 'spp'. ASVs that were not filtered were rarefied at 20,000 reads to calculate alpha diversity. The measures of alpha diversity at the ASV level include observed OTUs, Shannon's diversity index, Faith's phylogenetic diversity, and Pielou's evenness.

Serum metabolome profiling

The serum samples were stored at −80°C in a refrigerator and thawed on ice before being vortexed for 10 s. For metabolomic analysis, 50 μL of the sample and 300 μL of the extraction solution (ACN:Methanol = 1:4, V/V) containing internal standards were added to a 2 mL microcentrifuge tube. The sample was vortexed for 3 min and then centrifuged at 12,000 rpm for 10 min at 4°C. Subsequently, 200 μL of the supernatant was collected and placed at −20°C for 30 min, followed by another centrifugation at 12,000 rpm for 3 min at 4°C. An aliquot of 180 μL of the supernatant was transferred for LC-ESI-MS/MS analysis.

The sample extracts were analyzed using an LC-ESI-MS/MS system (UPLC, ExionLC AD, https://sciex.com.cn/; MS, QTRAP System, https://sciex.com/) following standard protocols. The triple quadrupole-linear ion trap mass spectrometer (QTRAP) was used to perform LIT and triple quadrupole (QQQ) scans, operated and controlled by Analyst 1.6.3 software (Sciex) with standard parameters. The source temperature 500°C; ion spray voltage (IS) was 5500 V (positive) and −4500 V (negative); ion source gas I (GSI), gas II (GSII), and curtain gas (CUR) were set at 55, 60, and 25.0 psi, respectively; the collision gas (CAD) was set to high. Instrument tuning and mass calibration were performed with10 and 100 μmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period.

QUANTIFICATION AND STATISTICAL ANALYSIS

The gut mycobiome taxa data were analyzed at the genus level. To address the issue of sparse fungal data in subsequent analyses, taxa present in less than 10% of the population were excluded. For the mycobiome-phenotype association analysis, the taxa data were transformed using the centered log-ratio (CLR) method to address the compositional nature of the mycobiome data. Prior to performing the CLR transformation, multiplicative imputation was employed to handle zero values.[43] As for blood metabolome data, a natural logarithmic (loge) transformation was applied. All statistical tests used were two sided. To account for potential false positive discoveries, multiple-testing correction was implemented using the Benjamini–Hochberg method. An adjusted p value of less than 0.05 was considered indicative of statistical significance. All regression analyses were performed using Stata (v15.0).

Core taxa identification

Core fungal taxa were identified as those appearing in at least 50% of samples with a minimum relative abundance of 0.1%, or forming central nodes in fungal network. NetCoMi package [43] was used to perform fungal co-abundance networks analyses. We used SparCC analysis to estimate the fungal correlation matrix. Correlations with FDR adjusted p values <0.05 and with a magnitude above 0.2 were selected for further visualization and network analysis. Here, we used eigenvector centrality to define the central node in the co-abundance network (node with a centrality value above the empirical 95% quantile in the network).

Estimation of the effect of environmental factors on fungal variability

We used the *vegdist* function in the R package *vegan* (version 2.5.7) to calculate the gut fungal Bray-Curtis dissimilarity matrix. The variations of Bray-Curtis dissimilarity matrix explained by each of the environmental variables was determined by PERMANOVA analysis using the function adonis2 in *vegan*. The p value was determined through 999 permutations.

Correlation between fungal genera and observed OTUs

Spearman correlation analysis was used to examine the correlation between measured fungal genera and observed OTUs.

Association of the gut mycobiome with human cardiometabolic diseases

We used logistic mixed-effects regression and linear mixed-effects regression, respectively, to independently evaluate the associations of each fungal features (per SD) with human cardiometabolic diseases and their indicative factors (fasting glucose, HbA1c,

HDL, LDL, TC, TG, SBP and DBP). The models were adjusted for age, gender, and BMI and included a random intercept and random coefficient for provinces or megacities, accounting for the variability in gut fungal composition across different regions.

For the fungus that was significantly associated with T2D, we further evaluated the interaction between the fungus and T2D traditional risk score for the risk of T2D using a logistic mixed-effects regression model. The model simultaneously included the two main effects, as well as the product of the two main effects, along with random intercepts and random coefficients to account for the variability across provinces or megacities. The risk score utilized in the analysis was developed based on the standards of medical care for T2D in China. It encompassed factors such as age, gender, waist circumference, BMI, and systolic blood pressure.[44] Furthermore, for the fungal taxa that showed a significant interaction with risk score for T2D, we performed a subgroup analysis stratified by the median level of the corresponding factor.

Temporal variability of gut mycobiome over time

We conducted dynamic analyses on participants with paired mycobiome data from the 2015 and 2018 surveys. To evaluate differences in the gut mycobiome between the two time points, we performed a principal component analysis (PCA) based on between-samples Aitchison distance, obtained from CLR-transformed abundance data. We performed paired two-sided Wilcoxon-signed rank tests on the first two axes of principal components to assess the variation between the two time points. To compare differences in overall fungal composition between and within individuals, we employed the Wilcoxon rank-sum test on the fungal Bray-Curtis dissimilarity. Furthermore, we utilized paired two-sided Wilcoxon-signed rank test to evaluate temporal changes in both alpha diversity and the relative abundance of gut fungi. For the clustering of samples across the two time points, we adopted the R package microbiome (version 1.19.1) to perform Dirichlet Multinomial Mixture (DMM) analysis.[45] The optimal number of clusters was identified based on the local minimum of the Laplace approximation score. To ensure the robustness of this cluster count, we used additional alternative criteria, such as the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC). We quantified the contribution of each fungal taxa to the fungal clusters, considering taxa with contribution values exceeding the 99th percentile of the overall contribution values as the primary contributor to the fungal cluster. We then analyzed the transition of all participants between the DMM clusters from the baseline to follow-up. Subsequently, we analyzed the transitions of all participants between the DMM clusters from baseline to follow-up.

Dynamic interplay between gut mycobiome and human cardiometabolic diseases

We fitted a generalized estimating equation (GEE) model to examine the association of dynamic changes in gut mycobiome clusters over time with cardiometabolic diseases. The models were adjusted for age, gender, and BMI and used independent correlation structure as the GEE model's working correlation matrix. Additionally, the models included both a random intercept and a random coefficient for provinces or megacities.

Associations of gut mycobiome with human serum metabolites

Among the participants with measured mycobiome and metabolome data in the 2015 survey, we constructed metabolome-based machine learning classifiers (LightGBM)[46] to classify the fungal clusters. Performance of the classifiers was assessed by area under the Receiver-Operating Characteristic curve (AUC) across 10-folds. The reference group was cluster 4, and each of other groups and their combinations as the comparison group.

We performed orthogonal projections to latent structures discriminant analysis (OPLS-DA) to investigate the differential metabolites between fungal cluster 3 and cluster 4. This analysis was performed in the R language using the 'ropls (version 1.26.4). Metabolites with a variable importance in projection (VIP) score exceeding 1 were selected for further investigation. We evaluated the chosen metabolites using the machine learning model (LightGBM). This analysis was conducted using the Python package lightgbm (version 3.2.1). The model was trained and initially validated using 10-fold cross-validation with the 2015 dataset. Subsequently, the model trained on the 2015 data (discovery model) was directly applied and validated on the 2018 dataset, which did not include any participants from the training dataset. Additionally, we employed the Wilcoxon rank-sum test to calculate the fold change of selected metabolites between the two groups.

Mediation analysis

We utilized weighted correlation network analysis (WGCNA, version 1.70.3)[47] to detect co-expressed metabolite modules based on the metabolites selected from OPLS-DA. To assess the associations between fungal clusters and diseases, as well as between fungal clusters (cluster 3 and cluster 4) and metabolite modules, logistic mixed-effects and linear mixed-effects regression analyses were conducted, respectively. Furthermore, logistic mixed-effects regression was used to evaluate the association between metabolite modules and diseases. All models were adjusted for age, gender, and BMI and included a random intercept and random coefficient for provinces or megacities, accounting for the variability in gut fungal composition across different regions. For metabolite modules that exhibited significant associations with both fungal partitions and diseases, we conducted mediation analysis using the R package mediation (v4.5.0) to investigate the potential mediation effect of metabolites on the link between mycobiome and cardiometabolic diseases. In addition, we also performed mediation analysis to explore whether disease status could mediate the association between mycobiome and the metabolites.

Cell Reports Medicine, Volume *5*

Supplemental information

The temporal dynamics of the gut mycobiome and its association with cardiometabolic health in a nationwide cohort of 12,641 Chinese adults

Wanglong Gou, Huijun Wang, Chang Su, Yuanqing Fu, Xinyu Wang, Chang Gao, Menglei Shuai, Zelei Miao, Jiguo Zhang, Xiaofang Jia, Wenwen Du, Ke Zhang, Bing Zhang, and Ju-Sheng Zheng

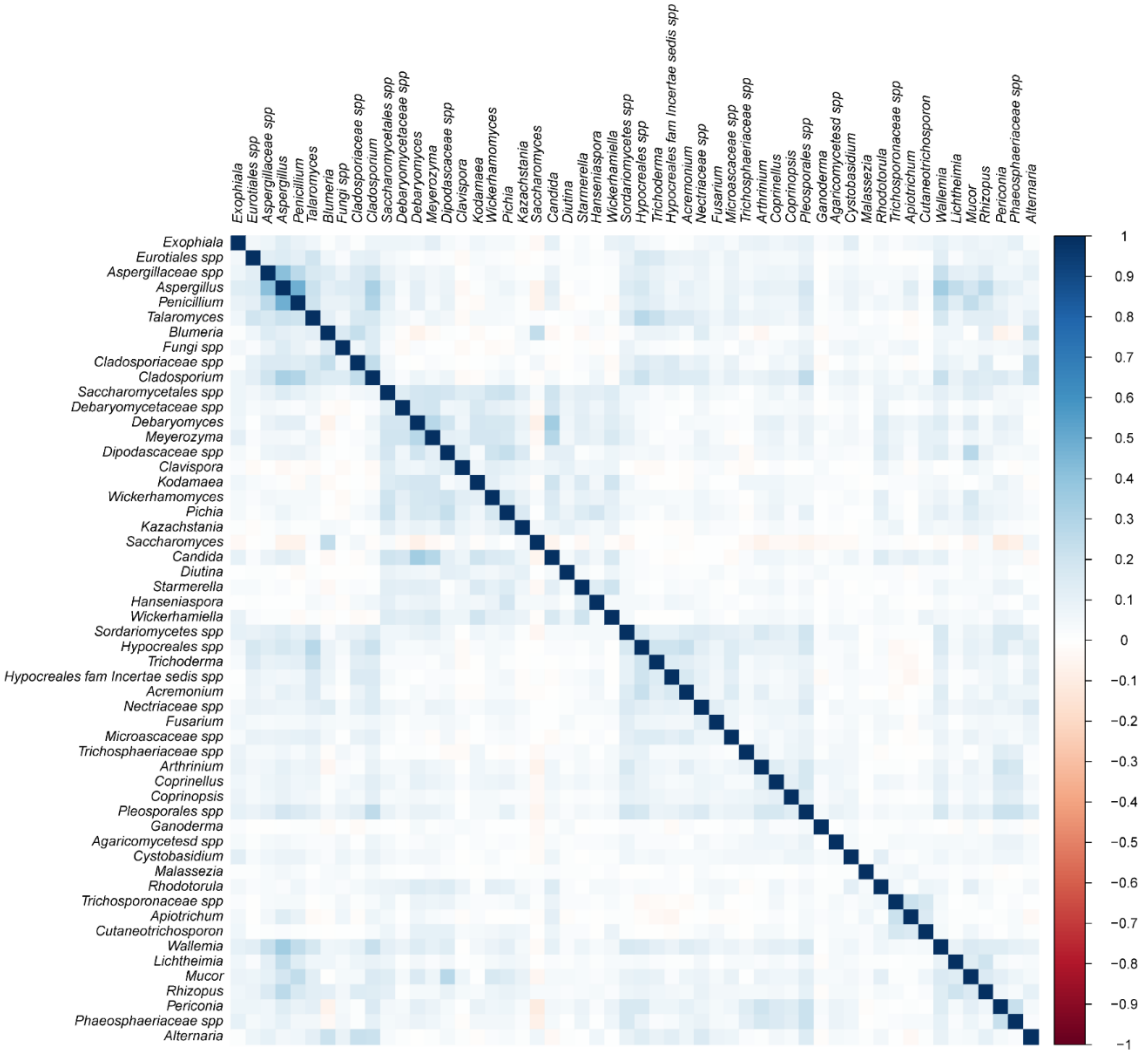


Figure S1. The inter-correlation of fungal genera, related to Figure 2. The correlation between each genus was calculated by Spearman correlation.

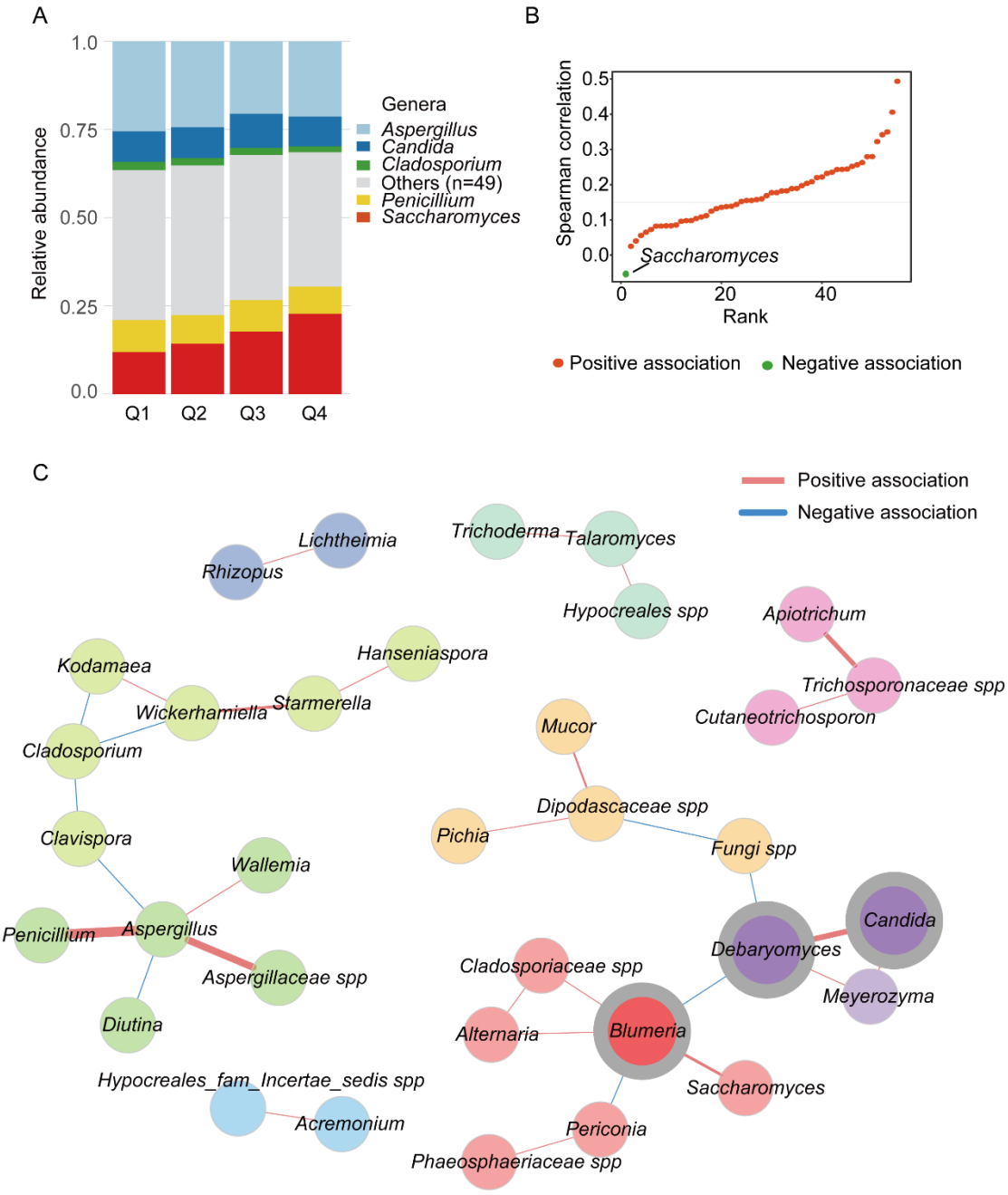


Figure S2. Distribution of fungal genera across urbanization score and correlation analyses, related to Figure 2. A, Distribution of fungal taxa that present in at least 50% of samples at a minimum relative abundance of 0.1%, across urbanization score. Q1 to Q4 represent the quantiles of the urbanization score. **B,** Correlations between the gut fungal genera and observed OUT. The values on the x-axis and y-axis represent the rank and Spearman's coefficient of each genus's correlation with observed OTUs, respectively. **C,** Co-abundance network construction and keystone taxa identification. Co-abundance network was constructed by SparCC analysis. Correlations with FDR adjusted P values < 0.05 and with a magnitude above 0.2 were selected for further visualization and network analysis. Here, we used eigenvector centrality to define the central node in the co-abundance network (node with a centrality value above the empirical 95% quantile in the network). The size of each node was proportional to its eigenvector centrality value. Nodes were clustered and marked with different colors to visually distinguish between clusters.

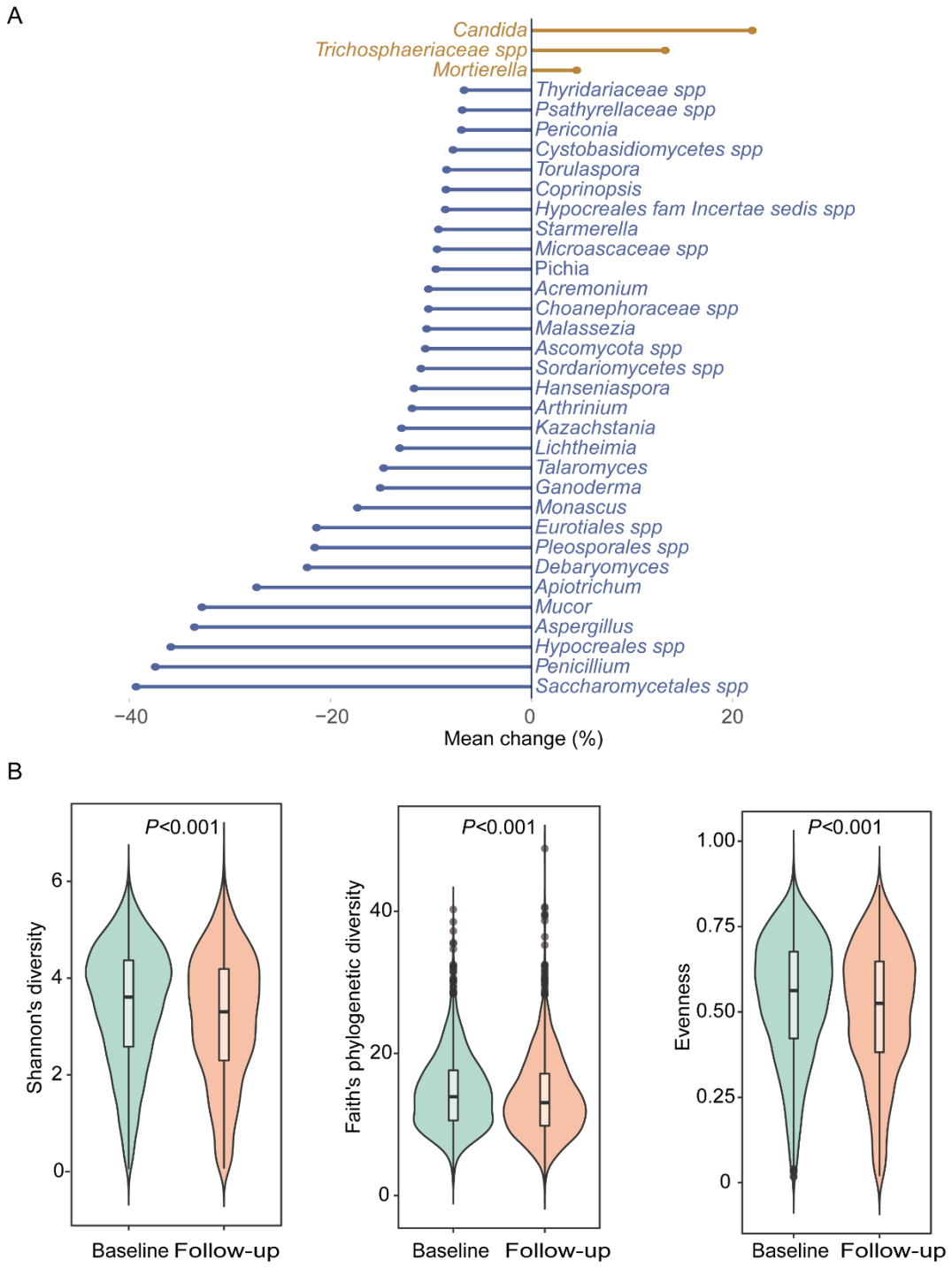


Figure S3. Temporal variability of gut mycobiome over time, related to Figure 4. A, Percentage changes of fungal genera with significant increase (brown-orange) or decrease (blue) in abundance over three years. Temporal changes in the relative abundance of the gut mycobiome were evaluated using a paired two-sided Wilcoxon signed-rank test. Only statistically significant results (FDR < 0.05) are depicted in the figure. **B,** Comparison of the fungal alpha diversity indices between baseline and follow-up.

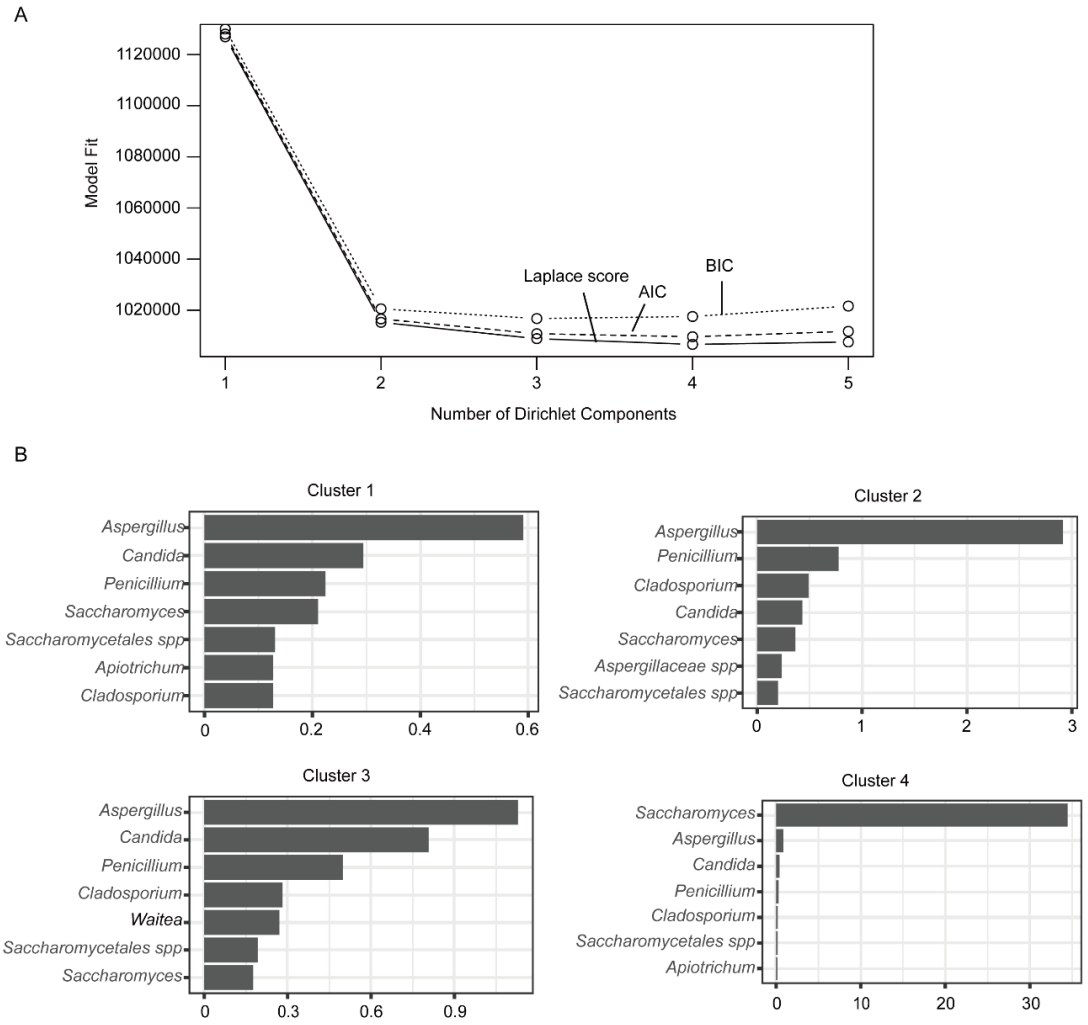


Figure S4. DMM clustering of ITS gene sequencing data, related to Figure 4. A, Model fitting performance for different numbers of clusters. The x-axis represents the number of clusters, while the y-axis showcases the corresponding values of the performance indices. AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion. **B,** Box plot showing the importance for dominant fungal taxa per DMM cluster.

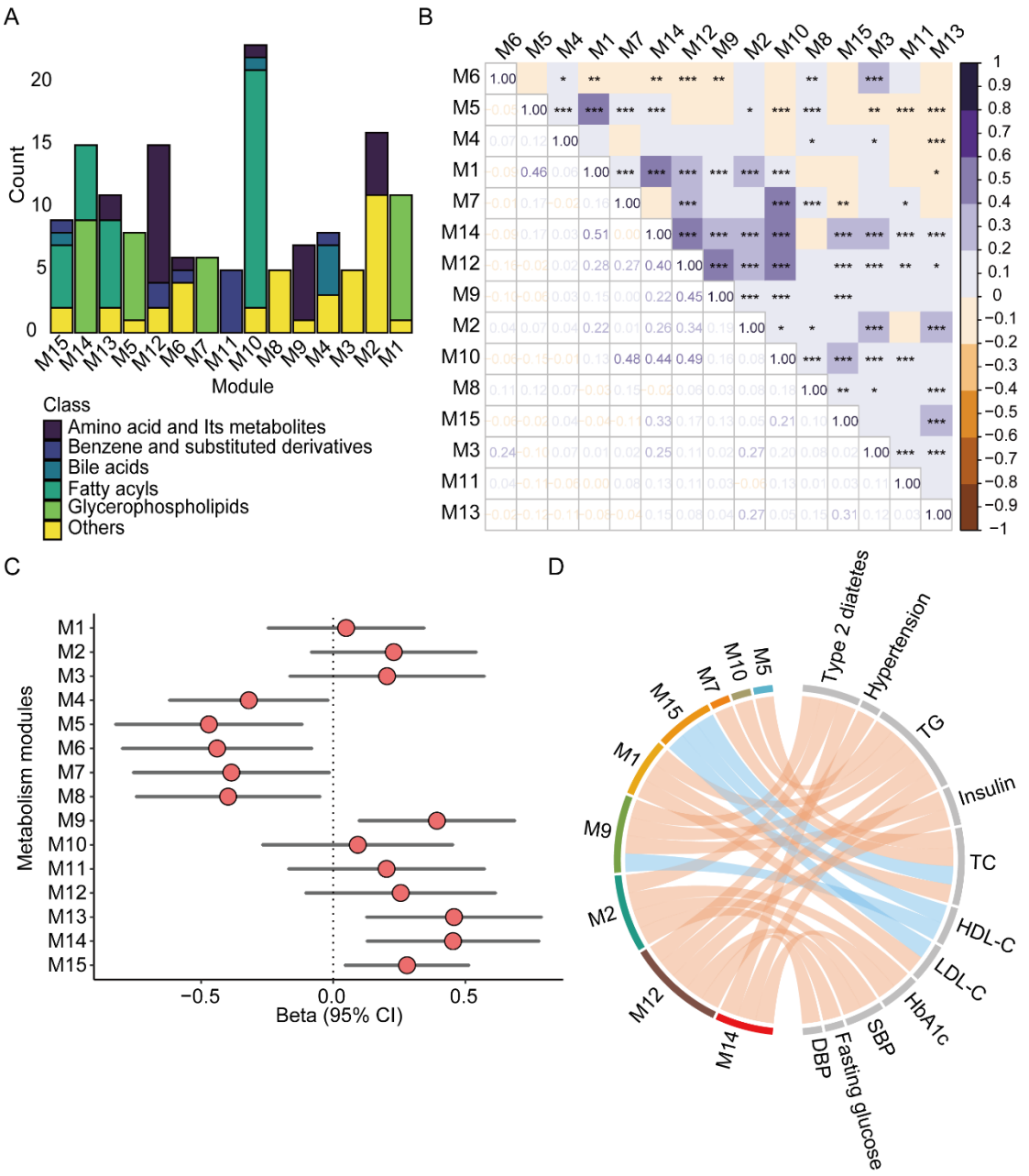


Figure S5. Dynamic interplay between gut mycobiome, metabolism modules and cardiometabolic health, related to Figure 5. A, The distribution of metabolites across different metabolism modules. **B,** Paired association of metabolism modules. The correlation between each metabolism module was calculated by Spearman correlation coefficient. Significance levels are indicated as follows: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. **C,** Association of gut mycobiome with metabolism modules. The analysis was conducted using linear mixed-effects regression model, with adjustment for the age, gender and BMI. The linear mixed-effects model included a random intercept and random coefficient for provinces or megacities, accounting for the variability in gut fungal composition across different regions. The figure presents Z-scored beta coefficients along with their corresponding 95% confidence intervals. **D,** Association of metabolism modules with cardiometabolic diseases, as well as host metabolism indicators. Weighted Gene Co-expression Network Analysis was employed to identify metabolism modules of differential metabolites. Only significant associations (orange for positive, blue for negative) were showed on the chord diagram.

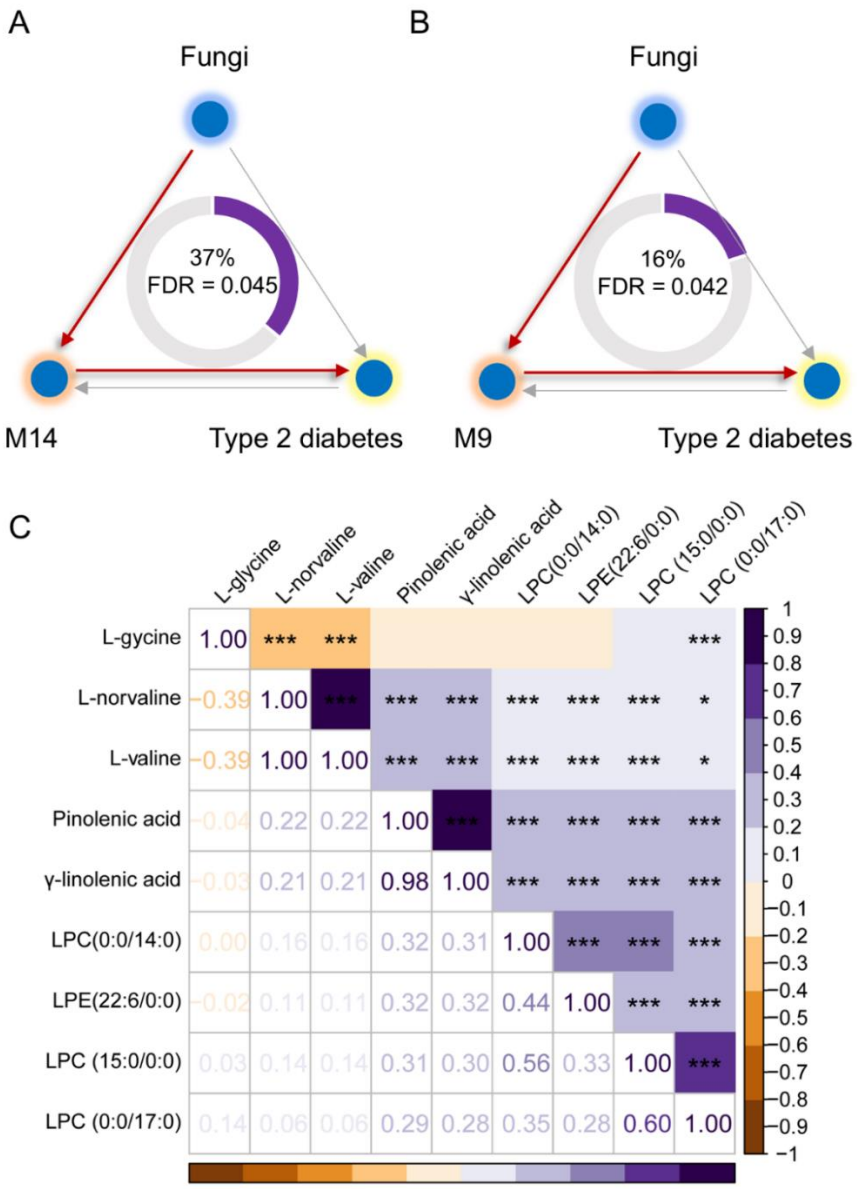


Figure S6. Metabolism modules mediate the effect of gut mycobiome on type 2 diabetes, related to Figure 6. A, Mediation linkages among the gut mycobiome, metabolism modules (M14 and M9) and type 2 diabetes. The proportion of mediation effect is displayed at the center of the ring chart. **B,** As in A, but for M9 module. **C,** Paired associations of metabolites belong to M14 and M9. The correlation between each metabolite was calculated by Spearman correlation coefficient. Significance levels are indicated as follows: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Table S1. Characteristics of the participants included in the study, related to Figure 1.

Factor	CHNS-1	CHNS-2 (baseline)	CHNS-2 (follow-up)
Number of participants	10695	1946	1946
Age, years	51.30 (15.48)	56.51 (13.73)	59.54 (13.73)
Women, n (%)	5727 (53.6%)	1064 (54.7%)	1064 (54.7%)
Urban, n (%)	4281 (40.0%)	584 (30.0%)	584 (30.0%)
Urbanization score	71.95 (17.64)	75.05 (16.68)	75.05 (16.68)
Pet owership, n (%)	1973 (19.0%)	496 (26.0%)	515 (26.7%)
Education, n (%)			
Middle school or lower	6822 (63.8%)	1459 (75.0%)	1459 (75.0%)
High school or professional college	2296 (21.5%)	353 (18.1%)	353 (18.1%)
University	1577 (14.7%)	134 (6.9%)	134 (6.9%)
Married, n (%)	10216 (95.5%)	1908 (98.0%)	1908 (98.0%)
Income, yuan / year per household	68145.23 (92725.05)	77164.02 (2.3e+05)	83565.78 (1.4e+05)
BMI, kg/m2	24.15 (4.10)	24.19 (4.23)	24.24 (3.85)
Waist circumference, cm	83.79 (12.72)	84.59 (12.05)	84.69 (11.12)
Hip circumference, cm	94.68 (10.34)	94.89 (9.56)	94.60 (9.17)
SBP, mmHg	126.91 (19.04)	131.16 (18.80)	133.75 (20.12)
DBP, mmHg	80.62 (11.17)	82.51 (10.57)	83.47 (11.59)
HbA1c, %	5.72 (0.97)	5.73 (0.92)	5.69 (0.97)
Fasting glucose, mmol/l	5.49 (1.53)	5.44 (1.40)	5.87 (1.62)
Insulin, mmol/l	7.62 (7.87)	7.07 (7.48)	10.44 (19.35)
HDL-C, mmol/l	1.28 (0.33)	1.25 (0.34)	1.46 (0.43)
LDL-C, mmol/l	3.12 (0.90)	3.12 (0.94)	3.15 (0.92)
TC, mmol/l	4.91 (1.10)	5.02 (1.04)	5.01 (0.99)
TG, mmol/l	1.50 (1.12)	1.53 (1.29)	1.73 (1.62)
Type 2 diabetes, n (%)	1183 (11.1%)	235 (12.1%)	285 (14.6%)
Prediabetes, n (%)	2900 (31.1%)	624 (32.1%)	514 (26.4%)
Hypertension, n (%)	2973 (28.0%)	721 (37.1%)	815 (41.9%)
Dyslipidemia, n (%)	3465 (32.4%)	698 (40.0%)	640 (34.7%)
Intestinal disease, n (%)	245 (2.4%)	25 (1.3%)	48 (2.5%)
Diarrhea, n (%)	101 (1.0%)	23 (1.2%)	18 (0.9%)
Myocardial infarction, n (%)	87 (0.8%)	16 (0.8%)	19 (1.0%)
Stroke, n (%)	135 (1.3%)	32 (1.6%)	53 (2.7%)
Cancer, n (%)	115 (1.1%)	19 (1.0%)	26 (1.3%)
Hypertension medications, n (%)	1446 (13.5%)	332 (17.1%)	414 (21.3%)
Diabetes medications, n (%)	437 (4.1.0%)	80 (4.1%)	104 (5.3%)
Antibiotic (current), n (%)	205 (2.0%)	29 (1.5%)	42 (2.2%)
Antibiotic (within 6 months), n (%)	1000 (9.7%)	153 (8.0%)	188 (9.8%)
Probiotics, n (%)	335 (3.2%)	43 (2.3%)	86 (4.5%)
Anti-inflammatory medications, n (%)	259 (2.5%)	30 (1.6%)	51 (2.6%)
Antiacid medications, n (%)	104 (1.0%)	18 (0.9%)	22 (1.1%)
Proton pump inhibitor, n (%)	154 (1.5%)	27 (1.4%)	33 (1.7%)
Current smoking, n (%)	2785 (26.3%)	490 (25.4%)	462 (23.8%)
Current alcohol consumption, n (%)	2905 (27.5%)	539 (27.9%)	503 (26.0%)
Physical activity, MET	143.48 (164.77)	173.93 (175.42)	171.24 (193.38)
Wheat intake, g/day	137.27 (136.18)	92.00 (109.05)	100.89 (108.53)
Rice intake, g/day	220.75 (172.69)	241.95 (132.27)	233.65 (142.74)
Dark vegetable intake, g/day	65.28 (75.69)	79.02 (93.01)	61.01 (81.76)
Light vegetable intake, g/day	200.56 (132.86)	214.45 (133.23)	214.30 (136.44)
Vegetable intake, g/day	265.83 (153.97)	293.47 (150.02)	275.31 (153.13)
Salted vegetable intake, g/day	3.58 (12.58)	5.34 (14.39)	5.13 (18.65)
Fruit intake, g/day	41.77 (73.87)	52.31 (81.80)	52.82 (90.17)
Nuts intake, g/day	4.47 (14.53)	4.68 (14.23)	4.76 (14.98)
Pork intake, g/day	68.54 (70.36)	65.58 (65.53)	63.87 (66.97)
Poultry intake, g/day	16.02 (36.24)	19.12 (39.95)	17.93 (37.95)
Milk intake, g/day	27.48 (75.69)	20.45 (61.36)	20.73 (62.03)
Egg intake, g/day	27.01 (31.03)	29.70 (33.36)	28.76 (35.16)
Fish intake, g/day	26.45 (46.70)	35.89 (54.12)	35.66 (54.20)
Carrot intake, g/day	27.52 (61.04)	28.28 (53.46)	32.48 (67.69)
Tuber intake, g/day	36.40 (54.73)	33.60 (59.24)	34.02 (66.63)
Pastes intake, g/day	1.74 (10.77)	1.15 (4.76)	1.16 (6.21)
Other meat intake, g/day	10.18 (27.12)	7.73 (23.57)	8.83 (26.44)
Cake intake, g/day	14.43 (38.36)	10.82 (29.48)	15.24 (39.97)
Sugar intake, g/day	2.26 (9.00)	2.35 (5.91)	1.93 (6.77)
Vegetable oils intake, g/day	35.94 (34.95)	39.14 (34.79)	35.55 (27.08)
Animal oil intake, g/day	3.47 (13.03)	0.45 (3.45)	1.03 (7.22)
Salt intake, g/day	8.53 (11.19)	9.99 (30.19)	7.70 (5.67)
Sauce intake, g/day	9.44 (27.57)	9.71 (13.70)	10.08 (13.04)
Other foods intake, g/day	18.77 (30.00)	16.19 (28.53)	14.09 (26.77)

Table S2. The prevalence and relative abundance of included genera in the CHNS-1 cohort, related to Figure 2.

Phylum	Class	Order	Family	Genus	Relative abundance (%)	Prevalence (%)
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	22.60	99.21
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	16.12	82.56
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	8.80	89.76
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	8.35	91.93
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	3.31	67.18
Ascomycota	Saccharomycetes	Saccharomycetales	Unassigned	Unassigned	3.04	64.82
Ascomycota	Saccharomycetes	Saccharomycetales	Dipodascaceae	Unassigned	2.88	57.28
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	2.00	71.86
Ascomycota	Leotiomycetes	Erysiphales	Erysiphaceae	Blumeria	1.20	46.11
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Unassigned	1.03	50.30
Ascomycota	Dothideomycetes	Pleosporales	Unassigned	Unassigned	0.99	48.21
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	0.88	26.42
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	0.82	26.63
Ascomycota	Sordariomycetes	Hypocreales	Unassigned	Unassigned	0.71	35.41
Ascomycota	Dothideomycetes	Pleosporales	Periconiaceae	Periconia	0.54	24.09
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Meyerozyma	0.48	26.45
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	0.47	19.41
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	0.45	15.95
Ascomycota	Eurotiomycetes	Eurotiales	Trichocomaceae	Talaromyces	0.39	28.85
Ascomycota	Eurotiomycetes	Eurotiales	Unassigned	Unassigned	0.38	22.69
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	0.34	22.48
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Clavispora	0.33	12.36
Ascomycota	Eurotiomycetes	Chaetothyriales	Herpotrichiellaceae	Exophiala	0.29	21.43
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	0.27	22.64
Ascomycota	Sordariomycetes	Unassigned	Unassigned	Unassigned	0.26	23.22
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Wickerhamiella	0.26	13.28
Ascomycota	Sordariomycetes	Trichosphaeriales	Trichosphaeriaceae	Unassigned	0.25	15.08
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Diutina	0.25	10.08
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Unassigned	0.24	20.34
Ascomycota	Sordariomycetes	Hypocreales	Hypocreaceae	Trichoderma	0.21	14.58
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Unassigned	0.21	20.15
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Kodamaea	0.18	11.88

Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Unassigned	0.17	12.63
Ascomycota	Sordariomycetes	Hypocreales	Hypocreales_fam_Incertae_sedis	Acremonium	0.14	10.04
Ascomycota	Sordariomycetes	Microascales	Microascaceae	Unassigned	0.13	11.01
Ascomycota	Sordariomycetes	Xylariales	Apiosporaceae	Arthrinium	0.13	12.03
Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Unassigned	0.12	11.82
Ascomycota	Sordariomycetes	Hypocreales	Hypocreales_fam_Incertae_sedis	Unassigned	0.11	10.29
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	0.11	12.59
Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	1.97	62.46
Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	0.85	39.34
Basidiomycota	Agaricomycetes	Unassigned	Unassigned	Unassigned	0.76	20.95
Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	0.76	43.52
Basidiomycota	Agaricomycetes	Polyporales	Ganodermataceae	Ganoderma	0.43	12.12
Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Cutaneotrichosporon	0.38	19.06
Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Unassigned	0.35	23.00
Basidiomycota	Cystobasidiomycetes	Cystobasidiales	Cystobasidiaceae	Cystobasidium	0.28	23.38
Basidiomycota	Agaricomycetes	Agaricales	Psathyrellaceae	Coprinopsis	0.20	16.14
Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	0.12	11.81
Basidiomycota	Agaricomycetes	Agaricales	Psathyrellaceae	Coprinellus	0.09	10.83
Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	3.49	61.93
Mucoromycota	Mucoromycetes	Mucorales	Rhizopodaceae	Rhizopus	0.83	41.25
Mucoromycota	Mucoromycetes	Mucorales	Lichtheimiaceae	Lichtheimia	0.11	15.63
Unassigned	Unassigned	Unassigned	Unassigned	Unassigned	9.95	89.71

Table S3. Normalized eigenvector centrality and degree of genera in the network for the CHNS-1 cohort, related to Figure 2.

Genus	Normalized eigenvector centrality	Degree
Blumeria	1	5
Debaryomyces	0.92	4
Candida	0.57	2
Cladosporiaceae spp	0.56	2
Alternaria	0.55	2
Meyerozyma	0.55	2
Fungi spp	0.44	2
Periconia	0.40	2
Saccharomyces	0.38	1
Dipodascaceae spp	0.21	3
Phaeosphaeriaceae spp	0.14	1
Mucor	0.079	1
Pichia	0.074	1
Lichtheimia	0	1
Aspergillus	0	5
Wickerhamiella	0	3
Cladosporium	0	3
Talaromyces	0	2
Trichosporonaceae spp	0	2
Clavispora	0	2
Kodamaea	0	2
Starmerella	0	2
Acremonium	0	1
Apiotrichum	0	1
Aspergillaceae spp	0	1
Cutaneotrichosporon	0	1
Diutina	0	1
Hanseniaspora	0	1
Hypocreales spp	0	1
Penicillium	0	1
Rhizopus	0	1
Trichoderma	0	1
Wallemia	0	1
Hypocreales_fam_Incertae_sedis spp	0	1
Agaricomycetes spp	0	0
Arthrinium	0	0
Coprinellus	0	0
Coprinopsis	0	0
Cystobasidium	0	0
Debaryomycetaceae spp	0	0
Eurotiales spp	0	0
Exophiala	0	0
Fusarium	0	0
Ganoderma	0	0

Kazachstania	0	0
Malassezia	0	0
Microascaceae spp	0	0
Nectriaceae spp	0	0
Pleosporales spp	0	0
Rhodotorula	0	0
Saccharomycetales spp	0	0
Sordariomycetes spp	0	0
Trichosphaeriaceae spp	0	0
Wickerhamomyces	0	0

Table S4. Association of fungal alpha diversity with cardiometabolic diseases, adjusted for sequencing depth, related to Figure 3.

Fungal alpha diversity	Cardiometabolic diseases	Adjusted odds ratio	Lower confidence interval	Higher confidence interval	P
Faith_pd	Dyslipidemia	0.90	0.86	0.94	0.000017
Observed OTUs	Type 2 diabetes	0.89	0.83	0.96	0.002
Observed OTUs	Dyslipidemia	0.93	0.89	0.98	0.003
Shannon	Type 2 diabetes	0.92	0.86	0.98	0.010
Observed OTUs	Hypertension	0.94	0.90	0.99	0.020
Faith_pd	Type 2 diabetes	0.92	0.86	0.99	0.029
Evenness	Type 2 diabetes	0.94	0.88	1.00	0.043
Shannon	Dyslipidemia	0.97	0.93	1.01	0.187
Shannon	Hypertension	0.97	0.93	1.02	0.269
Evenness	Dyslipidemia	0.99	0.94	1.03	0.493
Evenness	Hypertension	0.99	0.94	1.04	0.668
Faith_pd	Hypertension	0.99	0.94	1.04	0.711

Table S5. Characteristics of the participants at baseline, stratified by fungal clusters, related to Figure 4.

Factor	C1	C2	C3	C4
Number of participants	1084	525	174	163
Age, years	55.96 (13.79)	58.15 (12.77)	56.63 (15.25)	54.80 (14.29)
Women, n (%)	582 (53.7%)	298 (56.8%)	93 (53.4%)	91 (55.8%)
Urban, n (%)	398 (36.7%)	86 (16.4%)	33 (19.0%)	67 (41.1%)
Urbanization score	76.29 (17.33)	69.01 (17.47)	65.85 (15.31)	78.28 (16.50)
Education, n (%)				
Middle school or lower	767 (70.8%)	425 (81.0%)	149 (85.6%)	100 (61.3%)
High school or professional college	224 (20.7%)	71 (13.5%)	20 (11.5%)	45 (27.6%)
University	93 (8.6%)	29 (5.5%)	5 (2.9%)	18 (11.0%)
Married, n (%)	1063 (98.1%)	517 (98.5%)	172 (98.9%)	159 (97.5%)
Income, yuan / year per household	84922.58 (1.3e+05)	82833.71 (1.6e+05)	67058.78 (1.2e+05)	94251.86 (1.0e+05)
BMI, kg/m2	24.28 (3.72)	24.34 (3.66)	22.85 (7.65)	24.58 (3.83)
Waist circumference, cm	84.90 (12.09)	85.47 (10.75)	78.89 (13.29)	85.78 (12.81)
Hip circumference, cm	95.31 (9.76)	95.62 (7.97)	88.96 (11.60)	96.08 (8.38)
SBP, mmHg	131.32 (18.08)	131.70 (19.65)	129.59 (20.26)	130.03 (19.14)
DBP, mmHg	82.40 (10.15)	83.24 (11.04)	80.50 (10.46)	83.06 (11.62)
HbA1c, %	5.74 (0.96)	5.73 (0.83)	5.57 (0.93)	5.78 (0.87)
Fasting glucose, mmol/l	5.44 (1.43)	5.39 (1.36)	5.63 (1.23)	5.49 (1.45)
Insulin, mmol/l	7.22 (7.75)	6.58 (7.03)	6.36 (4.68)	8.35 (9.09)
HDL-C, mmol/l	1.25 (0.34)	1.24 (0.34)	1.37 (0.37)	1.21 (0.31)
LDL-C, mmol/l	3.15 (0.93)	3.00 (0.87)	3.36 (0.92)	3.13 (1.20)
TC, mmol/l	5.03 (1.02)	4.92 (0.98)	5.29 (1.06)	5.04 (1.32)
TG, mmol/l	1.54 (1.10)	1.53 (1.34)	1.30 (0.90)	1.69 (2.21)
Type 2 diabetes, n (%)	136 (12.6%)	61 (11.6%)	10 (5.8%)	28 (17.2%)
Hypertension, n (%)	394 (36.5%)	203 (38.7%)	61 (35.1%)	63 (38.7%)
Dyslipidemia, n (%)	388 (40.3%)	190 (39.0%)	55 (38.2%)	65 (42.5%)
Current smoking, n (%)	279 (26.1%)	129 (24.7%)	39 (22.4%)	43 (26.5%)
Current alcohol consumption, n (%)	317 (29.6%)	128 (24.5%)	52 (29.9%)	42 (25.9%)
Physical activity, MET	158.65 (162.54)	209.92 (201.01)	181.86 (172.98)	149.75 (152.88)
Wheat intake, g/day	87.39 (102.33)	118.22 (124.52)	36.87 (64.77)	96.27 (110.14)
Rice intake, g/day	237.89 (126.15)	228.43 (124.33)	322.50 (171.69)	226.44 (115.96)
Dark vegetable intake, g/day	76.67 (86.58)	64.78 (82.91)	147.41 (140.84)	68.05 (67.90)
Light vegetable intake, g/day	215.59 (132.90)	219.21 (133.08)	183.38 (130.83)	223.84 (132.24)
Vegetable intake, g/day	292.35 (149.76)	283.99 (149.98)	330.79 (151.52)	291.89 (145.70)
Salted vegetable intake, g/day	5.23 (13.55)	4.49 (12.97)	7.71 (16.79)	5.65 (12.35)
Fruit intake, g/day	57.04 (86.59)	48.07 (76.01)	30.65 (53.86)	56.43 (80.83)
Nuts intake, g/day	5.27 (14.88)	4.18 (11.91)	2.89 (10.90)	3.69 (11.88)
Pork intake, g/day	68.02 (65.82)	54.30 (62.77)	89.06 (70.08)	60.75 (58.90)
Poultry intake, g/day	18.53 (38.16)	14.90 (33.84)	32.84 (48.51)	20.50 (39.56)
Milk intake, g/day	24.83 (66.79)	10.39 (39.92)	5.36 (34.58)	39.33 (86.33)
Egg intake, g/day	31.04 (32.64)	30.83 (37.49)	17.92 (26.49)	29.64 (27.21)
Fish intake, g/day	37.71 (55.77)	32.96 (50.67)	22.18 (36.28)	47.52 (64.52)
Carrot intake, g/day	26.43 (48.29)	33.73 (58.67)	29.48 (58.80)	20.49 (47.99)
Tuber intake, g/day	36.64 (63.52)	35.10 (56.16)	8.22 (30.23)	35.27 (54.78)
Pastes intake, g/day	1.12 (4.68)	1.48 (5.54)	0.44 (2.57)	1.00 (3.62)

Other meat intake, g/day	8.38 (22.82)	5.42 (20.54)	3.98 (13.58)	14.09 (33.12)
Cake intake, g/day	9.65 (26.35)	13.30 (34.57)	5.46 (19.01)	15.92 (36.38)
Sugar intake, g/day	2.35 (5.47)	2.32 (6.36)	1.53 (5.69)	3.25 (6.96)
Vegetable oils intake, g/day	37.38 (32.60)	42.22 (35.59)	38.37 (30.04)	40.91 (41.60)
Animal oil intake, g/day	0.38 (3.04)	0.60 (3.74)	0.36 (2.31)	0.49 (4.34)
Salt intake, g/day	9.14 (14.32)	10.03 (21.09)	16.71 (83.58)	8.05 (6.98)
Sauce intake, g/day	9.40 (13.00)	10.01 (15.25)	10.68 (13.39)	9.65 (12.11)
Other foods intake, g/day	17.94 (31.64)	14.39 (24.96)	9.65 (17.61)	17.19 (24.59)