OPEN ACCESS

Original research

Landscape of the gut mycobiome dynamics during pregnancy and its relationship with host metabolism and pregnancy health

Yuanqing Fu ,[1,2,3] Wanglong Gou,[1,2,3] Ping Wu,[4] Yuwei Lai,[4] Xinxiu Liang,[2,3] Ke Zhang,[2,3] Menglei Shuai,[2,3] Jun Tang,[2,3] Zelei Miao,[2,3] Jieteng Chen,[2,3] Jiaying Yuan,[5] Bin Zhao,[6] Yunhaonan Yang,[7] Xiaojuan Liu,[8] Yayi Hu,[9] An Pan,[4] Xiong-Fei Pan,[7,10,11] Ju-Sheng Zheng [1,2,3,12,13]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/gutjnl-2024-332260).

For numbered affiliations see end of article.

Correspondence to
Professor Ju-Sheng Zheng, School of Medicine, Westlake University, Hangzhou 310030, Zhejiang, China; zhengjusheng@westlake.edu.cn, Professor Xiong-Fei Pan, West China Second University Hospital, Sichuan University, Chengdu 610041, China; pxiongfei@scu.edu.cn and Professor An Pan, Huazhong University of Science and Technology, Wuhan 430030, China; panan@hust.edu.cn

YF, WG, PW and YL contributed equally.

YF, WG, PW and YL are joint first authors.

Received 21 February 2024
Accepted 23 April 2024

To cite: Fu Y, Gou W, Wu P, *et al*. *Gut* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/gutjnl-2024-332260

ABSTRACT

Objective The remodelling of gut mycobiome (ie, fungi) during pregnancy and its potential influence on host metabolism and pregnancy health remains largely unexplored. Here, we aim to examine the characteristics of gut fungi in pregnant women, and reveal the associations between gut mycobiome, host metabolome and pregnancy health.

Design Based on a prospective birth cohort in central China (2017 to 2020): Tongji-Huaxi-Shuangliu Birth Cohort, we included 4800 participants who had available ITS2 sequencing data, dietary information and clinical records during their pregnancy. Additionally, we established a subcohort of 1059 participants, which included 514 women who gave birth to preterm, low birthweight or macrosomia infants, as well as 545 randomly selected controls. In this subcohort, a total of 750, 748 and 709 participants had ITS2 sequencing data, 16S sequencing data and serum metabolome data available, respectively, across all trimesters.

Results The composition of gut fungi changes dramatically from early to late pregnancy, exhibiting a greater degree of variability and individuality compared with changes observed in gut bacteria. The multiomics data provide a landscape of the networks among gut mycobiome, biological functionality, serum metabolites and pregnancy health, pinpointing the link between *Mucor* and adverse pregnancy outcomes. The prepregnancy overweight status is a key factor influencing both gut mycobiome compositional alteration and the pattern of metabolic remodelling during pregnancy.

Conclusion This study provides a landscape of gut mycobiome dynamics during pregnancy and its relationship with host metabolism and pregnancy health, which lays the foundation of the future gut mycobiome investigation for healthy pregnancy.

WHAT IS ALREADY KNOWN ON THIS SUBJECT

⇒ The human intestine is home to a diverse range of bacterial and fungal species, forming the ecological community that contributes to normal physiology.
⇒ The human gut ecological community changes during pregnancy and plays a role in gestational dysmetabolic conditions.
⇒ The remodelling of gut fungi during pregnancy and its potential influence on host metabolism and pregnancy health remains largely unexplored.

WHAT THIS STUDY ADDS

⇒ The compositional changes of gut fungi from early to late pregnancy exhibit a greater degree of variability and individuality compared with changes observed in gut bacteria.
⇒ The prepregnancy overweight status is a key factor influencing both gut mycobiome compositional alteration as well as the pattern of metabolic remodelling during pregnancy.
⇒ Gut fungal *Mucor* during early pregnancy is positively associated with the risk of gestational diabetes mellitus and macrosomia.
⇒ The multiomics data provide a landscape of the networks among gut fungi, biological functionality, serum metabolites and pregnancy health.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our findings indicate the dynamic nature of the gut mycobiome throughout each trimester of pregnancy and its impacts on host metabolism as well as pregnancy health.
⇒ The manipulation of the gut fungi, a crucial constituent of the gut ecological community, holds great potential to serve as a novel approach to promoting a healthy pregnancy.

INTRODUCTION

During normal pregnancy, the maternal body undergoes dramatic physiological changes including immunological, hormonal and metabolic changes.[1 2] Gut microbiota is considered as a virtual organ, and pregnancy status is associated with a profound alteration of the gut microbiota.[3 4] It is hypothesised that the host can manipulate the gut microbiota to promote metabolic changes during pregnancy, ultimately supporting the growth and development of the fetus.[4]

It is important to note that the gastrointestinal tract is populated not only by bacteria, but also by fungi, known as gut mycobiome.[5–7] However, fungal communities are still far less studied, compared with the extensive research conducted on the gut bacteria. As a major component of the gut microbiota, gut fungi are believed to play a crucial role in intestinal ecology, which is essential for host health.[8 9] Recently, several studies have investigated the gut mycobiome in pregnant women, particularly among those who are obese or have been diagnosed with gestational diabetes mellitus (GDM).[10–12] Although the sample sizes of these studies are limited, they provide crucial indications that the gut mycobiome may interact with host metabolism during pregnancy and influence the development of GDM. Therefore, the gut mycobiome has the potential to be an intervention target for promoting a healthy pregnancy.

Regarding the remodelling of gut mycobiome during pregnancy, however, the dynamics of gut mycobiome and its interactions with gut microbial functionality, host metabolism, pregnancy complications and adverse birth outcomes have not been well studied. Here, we aim to comprehensively examine the underlying determinants for gut mycobiome based on a large-scale cohort of pregnant women (n=4800) and profile the dynamics of gut mycobiome based on a subcohort of deeply phenotyped participants (n=750). Moreover, leveraging the repeat measurements of multiomics and deep phenotypes in the established subcohort, we aim to provide a landscape of the networks among the gut mycobiome, gut microbial functionality, host metabolism and pregnancy health.

METHODS

Methods are available as online supplemental file 1.

RESULTS

Participant characteristics and gut mycobiome composition

This study was based on a prospective birth cohort study in central China: Tongji-Huaxi-Shuangliu Birth Cohort (THSBC). The THSBC recruited pregnant women who initiated prenatal care in a local maternal and child health hospital during their early pregnancy. Exclusion criteria were (1) Receiving infertility treatment (eg, in vitro fertilisation or intrauterine insemination); (2) Reporting severe chronic or infectious diseases (eg, cancer, HIV infection or tuberculosis); or (3) Were unable to or refused to sign the informed consent. In the present analysis, we included 4800 participants who had available ITS2 sequencing data, dietary information and clinical records during their pregnancy. This data set enables us to comprehensively profile the gut mycobiome among pregnant women and investigate potential determinants contributing to the variations of the gut mycobiome. To examine how pregnancy impacts the gut mycobiome over time and investigate their potential associations with host metabolism, we established a subcohort of 1059 participants, which included 514 women who gave birth to preterm (n=240), low birthweight (n=137) or macrosomia (n=216) infants, as well as 545 randomly selected participants who did not experience the above three adverse pregnancy outcomes. ITS2 sequencing was performed for the whole cohort involving 4800 participants, while the shotgun metagenomics sequencing was performed for the first trimester (T1) samples within the established subcohort (n=1059). Additionally, within the subcohort, 750 and 748 participants had ITS2 and 16S sequencing data available, respectively, across all trimesters. Figure 1 provides an overview of the study workflow.

The age of the included participants (n=4800) ranged from 18 years to 40 years (mean age, 26.4; SD, 3.6; table 1). More than half (57.5%) of these women were primiparous, while a majority of the remaining participants (41.3%) were multiparous. Prior to pregnancy, 947 women in the study were underweight and 666 women were overweight or obese, while most of the women had normal body weights (n=3187). Age and parity are most important prepregnancy anthropometric factors contributing to the interindividual variation of gut mycobiome composition. Antibiotics use and dietary factors including steamed bread, egg, fruits, meat and tea consumption are also identified as significant contributors (figure 2A, $p<0.05$).

We observed three fungal enterotypes, including *Saccharomyces*-dominated enterotype (prevalence, 26.5%), *Candida*-dominated enterotype (18.8%) and *Aspergillus*-dominated enterotype (54.7%, figure 2B,C). Parity and dietary factors including steamed bread and tea consumption are most important factors influencing the fungal enterotypes (online supplemental table S1). Specifically, women with a *Saccharomyces*-dominated enterotype during early pregnancy are more likely to be primiparous and have a dietary preference for consuming steamed bread within the past year. On the other hand, women with a *Candida*-dominated enterotype are more likely to have a lifestyle characterised by tea consumption within the past year. Of note, *Saccharomyces*, *Candida* and *Aspergillus* are the top three prevalent fungi, while the overall prevalence of gut mycobiome is very sparse (figure 2D,E). Specifically, we identify a total of 626 fungal genera at T1 (n=4800), and more than 96% (606 out of 626) of the identified genera have a prevalence lower than 40%. We consider those present among more than 40% of participants as core fungal genera (n=20) in subsequent analyses.

Compositional dynamics of the gut mycobiome during pregnancy in the longitudinal subcohort

We examined the compositional changes during pregnancy among the 750 participants who had ITS2 sequencing data available for all trimesters. The results showed a global shift in microbial community composition from T1 to the second trimester (T2), but not T2 to the third trimester (T3) (online supplemental figure 1A,B). Inconsistent with prior knowledge that the gut mycobiome was relatively stable,[9-10] our findings showed that as many as 68.5% of the 750 participants experienced shifts of the fungal enterotype during pregnancy (figure 3A). Specifically, the proportion of *Saccharomyces*-dominated enterotype increased (T1, 27.7%; T2, 31.7% and T3, 34.4%) and the *Aspergillus*-dominated enterotype decreased (T1, 53.1%; T2, 49.3% and T3, 48.4%) from T1 to T3. The proportion of *Candida*-dominated enterotype was relatively stable throughout the pregnancy (ranging from 17.2% to 19.1%).

The within-sample α diversity, including phylogenetic diversity and richness, was substantially reduced from T1 to T3, while the Shannon Index was not that dynamic (figure 3B–D). Moreover, the alterations in the within-sample α diversity during pregnancy were very similar between women who gave birth to healthy infants and those who had preterm, low birthweight or macrosomia infants. Although dietary factors contributed to the gut mycobiome compositional variation across participants, we did not observe significant associations between changes in the consumption of eight main food groups from T1 to T3 and decreased richness, after multiple testing correction (false discovery rate (FDR)>0.05, online supplemental table S2).

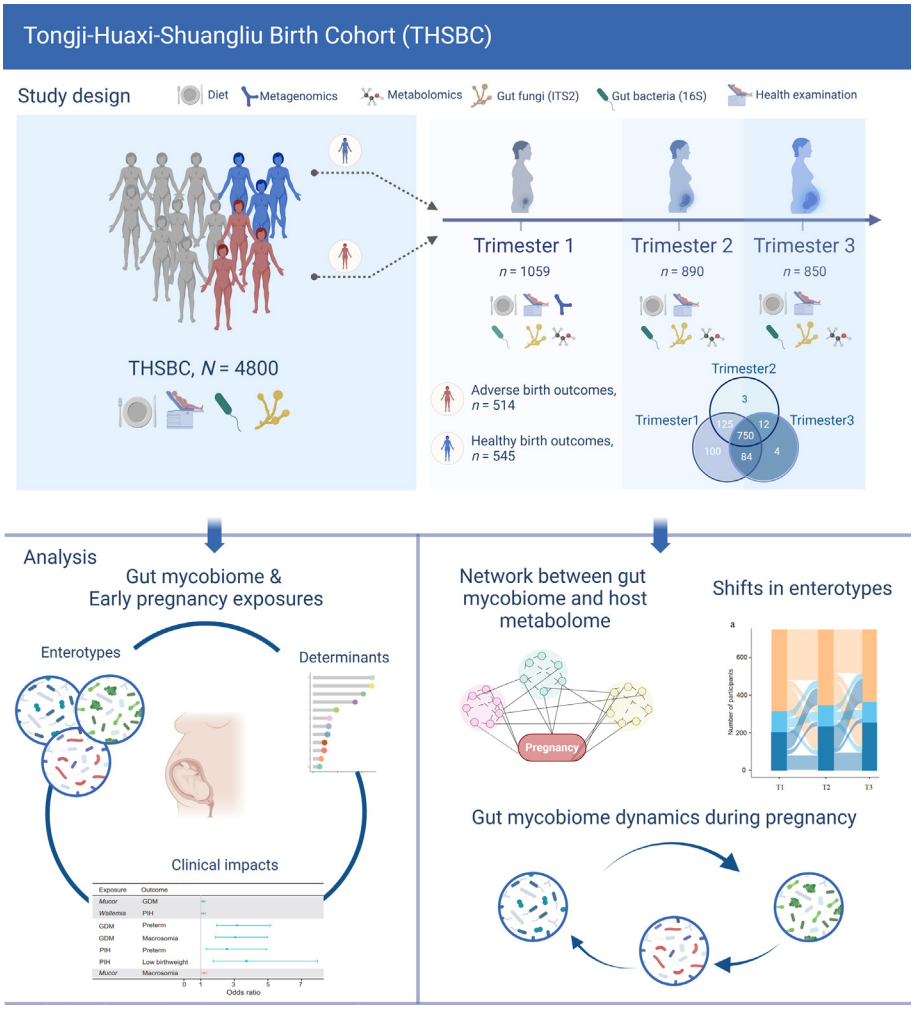


Figure 1 Study workflow for profiling the gut fungi during pregnancy and exploring its relationship with host metabolism and health. To comprehensively profile the gut mycobiome-host interaction among pregnant women, we investigated potential determinants contributing to the variations of the gut mycobiome and explored the impact of gut mycobiome during early pregnancy on later pregnancy complications as well as birth outcomes in a large cohort involving 4800 pregnant women. To examine how pregnancy impacts the gut mycobiome over time and investigate potential associations between the gut mycobiome and host metabolism, we established a subcohort of 1059 participants, which included 514 women who gave birth to preterm (n=240), low birth weight (n=137) or macrosomia (n=216) infants, as well as 545 randomly selected healthy controls. Within this subcohort, ITS2 sequencing was performed on 1059 stool samples collected during the first trimester of pregnancy, 890 during the second trimester of pregnancy and 850 during the third trimester. A total of 750 participants in this subcohort had ITS2 sequencing data available for all trimesters.

Therefore, the decreased richness was not likely to be driven by the change of diet during pregnancy.

We further profiled the extent of intraindividual shift in the composition in the context of interindividual variation. There was a wide range of intraindividual Bray-Curtis distance from T1 to T2 or T1 to T3, most of which were even larger than the mean interindividual distance at a single time point (figure 4A, online supplemental figure 1C). By contrast, the gut bacteria composition was much more conserved during pregnancy. We observed consistency between compositional alterations and the transition of enterotypes, as participants who exhibited a changed fungal enterotypes displayed considerably greater alterations in gut mycobiome composition (online supplemental figure 1D). Moreover, we found that women who were overweight or obese prior to pregnancy experienced much more variation of gut fungi from T1 to T3 compared with those who were underweight before pregnancy (figure 4B).

Dynamics of the individual gut fungal genus during pregnancy in the longitudinal subcohort

Among the 750 participants with gut mycobiome data available across all trimesters, we identified 410 genera in T1 samples. We then assessed the instability for each genus by calculating the loss rate from T1 to T3. For the 390 less prevalent fungi genera, the mean loss rate was as high as 97.6%, indicating extreme instability. By contrast, the 20 core fungi exhibited a mean loss rate of 55.7%, with the *Aspergillus*, *Candida* and *Saccharomyces* being the most stable fungal genera (figure 4C).

To examine the remodelling of the gut fungi from early to late pregnancy, we found that 4 out of the 20 core genera were substantially altered (paired t test on centered log-ratio (CLR)-transformed data between T1 and T3, FDR<0.05). Specifically, these genera were *Aspergillus*, *Cladosporium*, *Penicillium* and *Candida*, and all these fungal genera were depleted during late pregnancy compared with early pregnancy, which occurred

Table 1 Characteristics of the study population

	Total (n=4800)	Subcohort (n=1059)	Without adverse birth outcomes* (n=545)	With adverse birth outcomes (n=514)
Age (SD)	26.4 (3.6)	26.5 (3.7)	26.2 (3.6)	26.7 (3.9)
Prepregnancy BMI (SD)	20.9 (2.9)	21.2 (3.1)	20.9 (3.0)	21.5 (3.2)
Weight gain (SD)	14.5 (5.0)	14.5 (5.2)	14.5 (4.9)	14.4 (5.4)
Gestation duration (SD)	39.2 (1.3)	38.6 (1.9)	39.5 (0.9)	37.8 (2.3)
Delivery mode				
Vaginal	54.9%	49.5%	59.1%	39.4%
Caesarean	45.1%	50.5%	40.9%	60.6%
Pregnancy complications				
Gestational diabetes mellitus	3.0%	8.3%	7.6%	9.1%
Gestation-induced hypertension	1.9%	2.1%	0.6%	3.8%
Anaemia in pregnancy	45.0%	39.2%	39.6%	38.8%
Adverse birth outcomes				
Preterm	5.5%	22.3%	0	46.7%
Low birth weight	3.1%	12.9	0	26.7%
Macrosomia	4.9%	20.4%	0	42.0%
Gravidity				
1	35.0%	36.6%	40.9%	32.2%
2	28.4%	28.5%	27.2%	29.8%
3	19.0%	17.3%	16.3%	18.3%
4	11.2%	10.2%	9.4%	10.9%
>4	6.4%	7.5%	6.3%	8.8%
Parity				
0	57.5%	60.7%	61.4%	60.0%
1	41.3%	38.5%	38.3%	38.8%
≥2	1.2%	0.76%	0.4%	1.2%

*Adverse birth outcomes of interest in this study are preterm delivery, low birth weight and macrosomia.
BMI, body mass index.

in 40.4%–57.6% of women. Moreover, we used a machine learning algorithm (Light GBM) to discriminate T1 and T3 samples based on the gut mycobiome composition (10-fold cross validation, area under the curve (AUC)=0.728, figure 4D). Using the interpretable Shapley Additive exPlanations (SHAP) value, we found that the core genera accounted for all the top 10 discriminate fungal genera. Then we only used the core fungal genera for discrimination, and the performance was comparable (AUC=0.725, figure 4D). Most of these top discriminatory genera were over-represented in T1 and belonged mostly to the *Saccharomycetales, Eurotiales* or *Capnodiales* order of the *Ascomycota* (n=8). The relative abundance and prevalence of the top 10 discriminate genera during T1 and T3 were shown in figure 4E. Additionally, we conducted a classification analysis to distinguish between T1 and T2 samples, and between T2 and T3 samples. The AUCs for these comparisons were 0.78 and 0.61, respectively (online supplemental figure 2A,B). Moreover, 8 out of the top 10 fungal genera that contributed significantly to the discrimination between T1 and T2 samples, were also top 10 discriminant genera between T1 and T3 samples. These findings suggest that the gut mycobiome features of T2 and T3 samples are quite similar.

Key microbial functional pathways and host serum metabolites correlated with gut mycobiome

To examine the potential interaction between gut fungi and bacteria in the context of gut microbial functionality, we annotated biological pathways mostly specific to bacteria based on the paired shotgun metagenomics data for T1 samples (n=1039) within the established subcohort. As the gut mycobiome clusters to well-defined gut fungal enterotypes, we then investigated whether the matrix of pathways recapitulates this structure. The fungal enterotypes significantly contributed to the variations of the pathway matrix ($p<0.05$, $R^2=0.69\%$, online supplemental figure 2C), which may indicate an interaction between the gut mycobiome composition and the gut bacterial functionality. As expected, the bacterial enterotypes, identified following the same procedure as that of the fungal enterotypes, explained much more variations of the pathway matrix ($R^2=5.15\%$, online supplemental figure 2D).

To further characterise the relationship between fungal enterotypes and individual gut microbial functional pathways, we identified eight pathways whose distributions varied across enterotypes (FDR<0.05 with Kruskal-Wallis test). Specifically, most of these pathways (seven out of eight) including thiamin salvage IV (yeast), sucrose degradation IV, L-lysine biosynthesis I, superpathway of L-phenylalanine biosynthesis, peptidoglycan maturation, C4 photosynthetic carbon assimilation cycle and seleno-amino acid biosynthesis, were relatively abundant in the *Saccharomyces*-dominated enterotype, but less abundant in the *Aspergillus*-dominated enterotype. At the same time, sucrose degradation IV and seleno-amino acid biosynthesis were also abundant in the *Candida*-dominated enterotype, while preQ0 biosynthesis was abundant in both *Candida*-dominated enterotype and *Aspergillus*-dominated enterotype (figure 5A). Moreover, further regression analysis identified 95 significant associations between the identified pathways and 27 serum metabolites (FDR<0.05, figure 5A). More than two-thirds of the significant metabolites were bile acids (eg, chenodeoxycholic acid, isochodeoxycholic acid, deoxycholic acid, hyodeoxycholic acid and glycolithocholic acid), amino acid metabolites (eg, 4-hydroxyhippurate and phenylacetyl-L-glutamine) and organic acid derivatives. (eg, 3-indolepropionic acid and indoleacrylic acid; figure 5A).

As we repeatedly measured the serum metabolome within the established subcohort, we profiled the metabolic alterations and explored its covarying relationship with gut fungi. We quantified 794 identified metabolites, which belonged to diverse biochemical classes, such as amino acids, lipids, nucleotides and carbohydrates. Similar to the gut mycobiome, the serum metabolome also altered dramatically across trimesters, and the alteration between T1 and T2 was much larger than that between T2 and T3 (figure 5B). The proportion of explained variance by trimester ranged from 1.17% to 11.89% for different classes of metabolites, with hormones and its related metabolites ranking at the top (figure 5C). Correlation analyses between individual core fungal genera and individual metabolites showed 30 covarying relationships involving six genera and 27 serum metabolites (FDR<0.05, figure 5D). Four out of the six identified fungi were annotated to the genus level resolution, including *Cladosporium*, *Aspergillus*, *Rhizopus* and *Mucor*. Specifically, *Aspergillus*, *Rhizopus* and *Mucor* covaried positively with three, eight and one metabolite, respectively, while *Cladosporium* covaried negatively with seven metabolites. Most of the *Rhizopus*-related metabolites belong to fatty acyl and the *Cladosporium*-related metabolites are mainly amino acids or its metabolites.

Similar to the finding that the intraindividual variation of gut mycobiome composition was different across underweight, normal-weight and overweight women, we found that being overweight prior to pregnancy significantly impacted the pattern of metabolic alterations from T1 to T3 (figure 6A). Subsequently, we performed an analysis to identify metabolites that exhibited significant changes dependent on the prepregnancy overweight

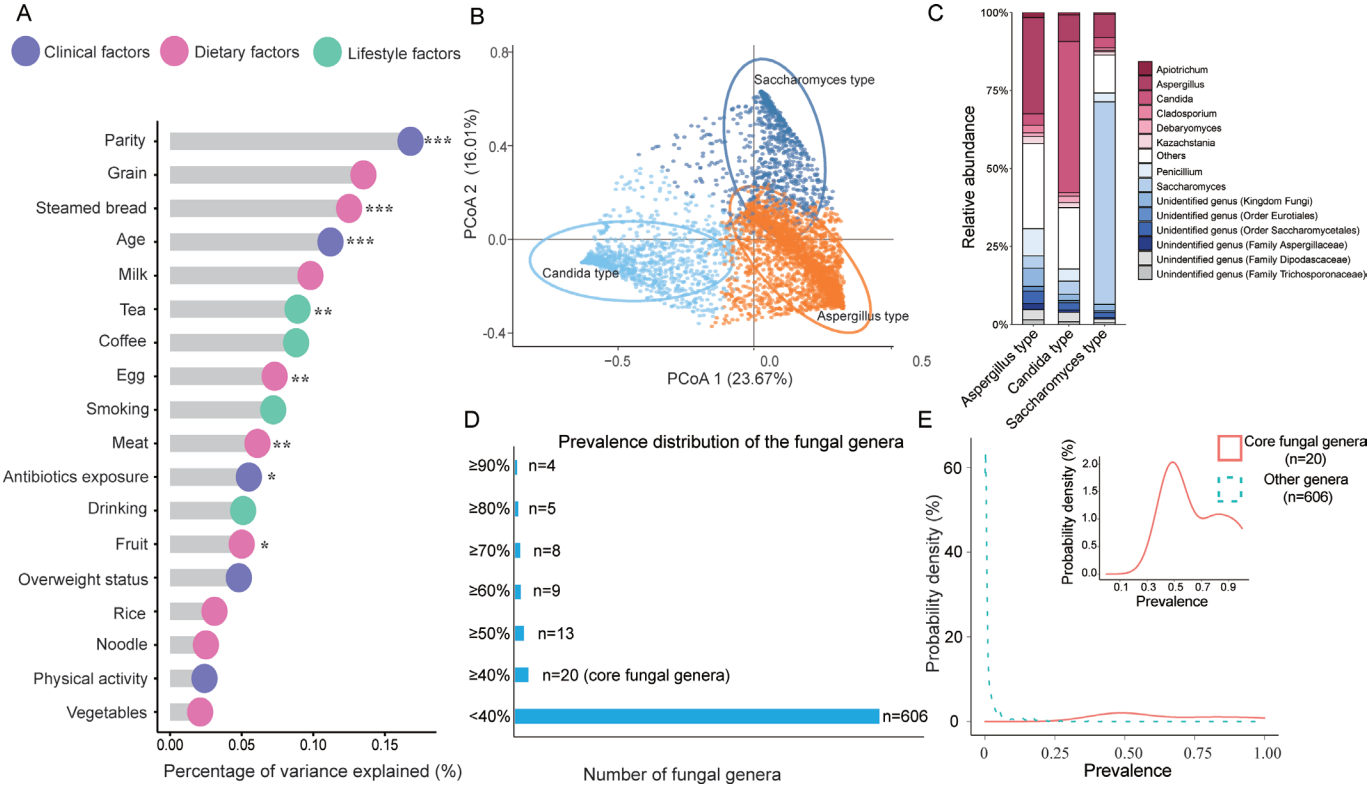


Figure 2 Profiling of the gut mycobiome composition and enterotypes among pregnant women. (A) Variance in the mycobiome composition explained by potential determinants was assessed through permutational multivariate analysis of variance (PERMANOVA) analysis. This analysis was performed based on 4800 independent samples collected during the first trimester. The value of p was determined through 999 permutations. Significance levels are indicated as follows: *, p<0.05; **, p<0.01; ***, p<0.001.(B) Clustering results of fungal enterotypes were visualised by principal coordinate analysis (PCoA). This visualisation was applied for all samples collected in the whole cohort. (C) The most abundant genera within each enterotype were shown. This analysis was based on all samples collected in the whole cohort. (D) The number of gut fungal genera that survived prevalence-based filtering at various cut-off thresholds was shown. This analysis was performed based on 4800 independent samples collected during the first trimester. (E) The distribution of prevalence of gut fungal genera was demonstrated. This analysis was performed based on 4800 independent samples collected during the first trimester.

status. Our findings revealed that 23 specific metabolites (eg, L-Glycine, L-Arginine, hexadecanedioic acid and carnitine) mainly belonging to amino acids and fatty acyl displayed significant alterations exclusively among underweight women during pregnancy (FDR<0.05, figure 6B). In contrast, 24 metabolites (eg, L-Valine, L-Glutamine, 2-Hydroxycinnamic acid and 4-Hydroxybenzyl alcohol) mainly belonging to amino acids and benzene derivatives were found to be significantly altered solely among overweight women (FDR<0.05, figure 6B). Thus, these metabolites likely contributed to the distinct pattern of metabolic changes observed in underweight or overweight women.

Associations of gut mycobiome with pregnancy outcomes

By examining the prospective associations of each core fungal genera during early pregnancy with pregnancy complications in the whole cohort, we found significantly positive associations between *Mucor* and incident GDM (OR: 1.15, 95% CI 1.06 to 1.26; FDR<0.05), and positive associations of *Wallemia* with pregnancy-induced hypertension (PIH; OR: 1.16, 95% CI 1.04 to 1.29; p=0.006, FDR=0.12, figure 6C), after adjustment for potential confounders. Moreover, we validated that both GDM and PIH were risk factors for adverse birth outcomes, including macrosomia and preterm delivery. Specifically, GDM was associated with higher risk of macrosomia (OR: 3.08, 95% CI 1.88 to 5.03; FDR<0.05) and preterm birth (OR: 3.15, 95% CI 1.93 to 5.12; p<0.001, figure 6c). PIH was associated with higher risk of low birth weight (OR: 3.73, 95% CI 1.75 to 7.96; FDR<0.05) and preterm birth (OR: 2.57, 95% CI 1.34 to 4.94; FDR<0.05, figure 6C).

Our analysis of the direct relationship between each core fungal genus and adverse birth outcomes in the whole cohort showed that the relative abundance of *Mucor* was positively associated with the risk of macrosomia (OR: 1.20, 95% CI 1.07 to 1.35; FDR<0.05, figure 6C). We therefore performed a mediation analysis to test whether the effects of *Mucor* on the fetal overgrowth were mediated by GDM. We found that the gut fungal *Mucor* during early pregnancy was associated with macrosomia risk independently of GDM, which suggested that the *Mucor* and GDM might influence the risk of macrosomia through different pathways (figure 6D). We also explored the associations between the extent of gut mycobiome compositional alterations and adverse birth outcomes within the subcohort, which yielded no significant results. This finding may indicate that the shift of overall gut mycobiome composition was a widely shared phenomenon driven by pregnancy, regardless of birth outcomes. However, this did not preclude that the trajectory of some individual gut fungus might be associated with pregnancy health, as we found significant dissimilarity in the trajectory of *Mucor* between pregnant women delivering preterm and non-preterm infants (FDR<0.05, figure 6E).

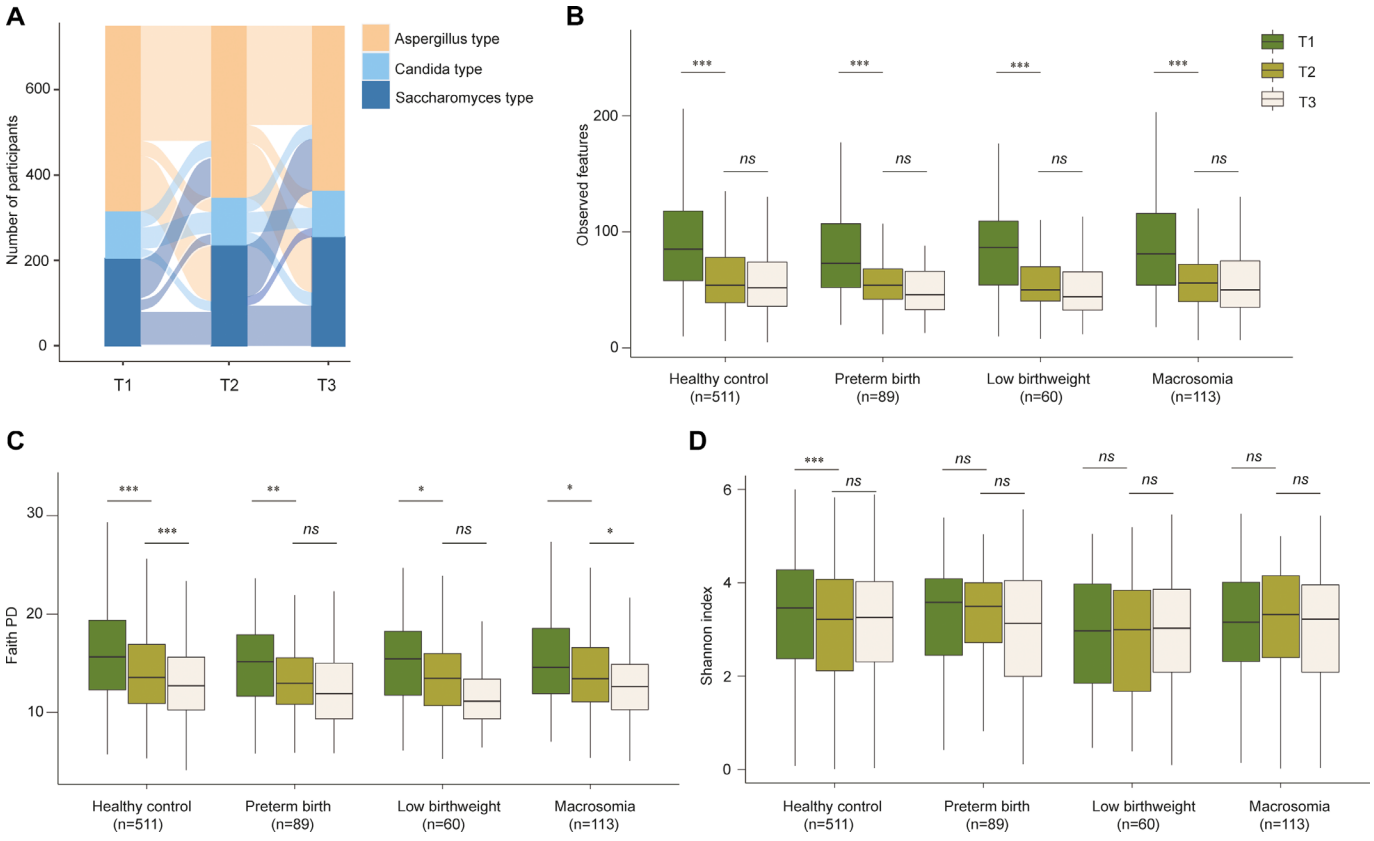


Figure 3 The shifts of gut fungal enterotypes and the dynamics of gut fungal α diversity during pregnancy. (A) Sankey diagram illustrating the shifts of gut fungal enterotypes from early pregnancy to late pregnancy. (B–D) The comparison of gut fungal α diversity across different trimesters, demonstrating the dynamics of gut fungal α diversity from early pregnancy to late pregnancy. Box plot centres show medians of the α diversity metrics with boxes indicating their IQRs, upper and lower whiskers indicating 1.5 times the IQR from above the upper quartile and below the lower quartile, respectively. Paired t test was performed to determine the significance of difference. Significance levels are indicated as follows: *ns*, p>0.05; *, p<0.05; **, p<0.01; ***, p<0.001. T1, the first trimester of pregnancy; T2, the second trimester of pregnancy; T3, the third trimester of pregnancy.

DISCUSSION

To the best of our knowledge, this is the first large-scale prospective cohort study characterising the determinants for gut mycobiome and profiling the gut mycobiome dynamics among pregnant women. We validate three of the previously reported gut fungal enterotypes, and the shifts of gut fungal enterotypes are common during pregnancy. Our quantification analysis of compositional alterations throughout the pregnancy supports a much higher dynamic characteristic of gut fungi compared with the gut bacteria, and highlights that prepregnancy overweight status is a significant contributor to the extent of alterations in gut mycobiome composition. Moreover, we perform network analysis among gut mycobiome, biological functionality, serum metabolites and pregnancy health, identifying that *Mucor* is prospectively associated with both GDM and macrosomia risk.

Our previous work investigated the determinants and stability of gut mycobiome among middle-aged and elderly adults.[9] Gut mycobiome composition was temporally stable while modulated by age, long-term habitual diet and host physiological states. A recent study also reported that age could significantly explain the interindividual variation of the human gut mycobiome and strongly affected the fungal enterotypes in several independent cohorts.[13–15] In the present study, we found similar significant determinants of gut mycobiome, such as age and diet, but the mycobiome composition was not that stable during pregnancy. The most significant determinants for gut mycobiome among pregnant women included age, parity and dietary intakes of steamed bread, egg, fruits, meat and tea consumption. Moreover, the consumption of steamed bread was related to the *Saccharomyces*-dominated enterotype while drinking tea was associated with the *Candida*-dominated enterotype. Interestingly, steamed bread is made by fermenting with *Saccharomyces* and the fungi are also involved in tea fermentation in China, thus the consumption of steamed bread or fermented tea may directly affect culture-independent gut mycobiome composition.

The present study provides evidence that gut mycobiome composition and structure are unstable and the well-defined gut fungal enterotypes could be altered over the course of the pregnancy. Taking the gut bacteria dynamics during pregnancy as a comparison, the longitudinal intraindividual distance for gut fungi is highly individualised and the mean intraindividual distance is much larger than that for gut bacteria composition. This finding was also true for stool samples collected among Human Microbiome Project volunteers, which showed more similar faecal bacterial community structure than faecal fungal community structure over time.[16] The longitudinal samples of one individual's faecal fungal mycobiome are even less similar to each other than those of another individual. Prior studies had shown differences in maternal gut bacterial composition by prepregnancy weight, indicating considerable effects of prepregnancy body mass index on gut microbiota composition during pregnancy.[17–19] In the present study, we report that the prepregnancy overweight status has a substantial influence on the alterations in gut mycobiome composition during pregnancy.

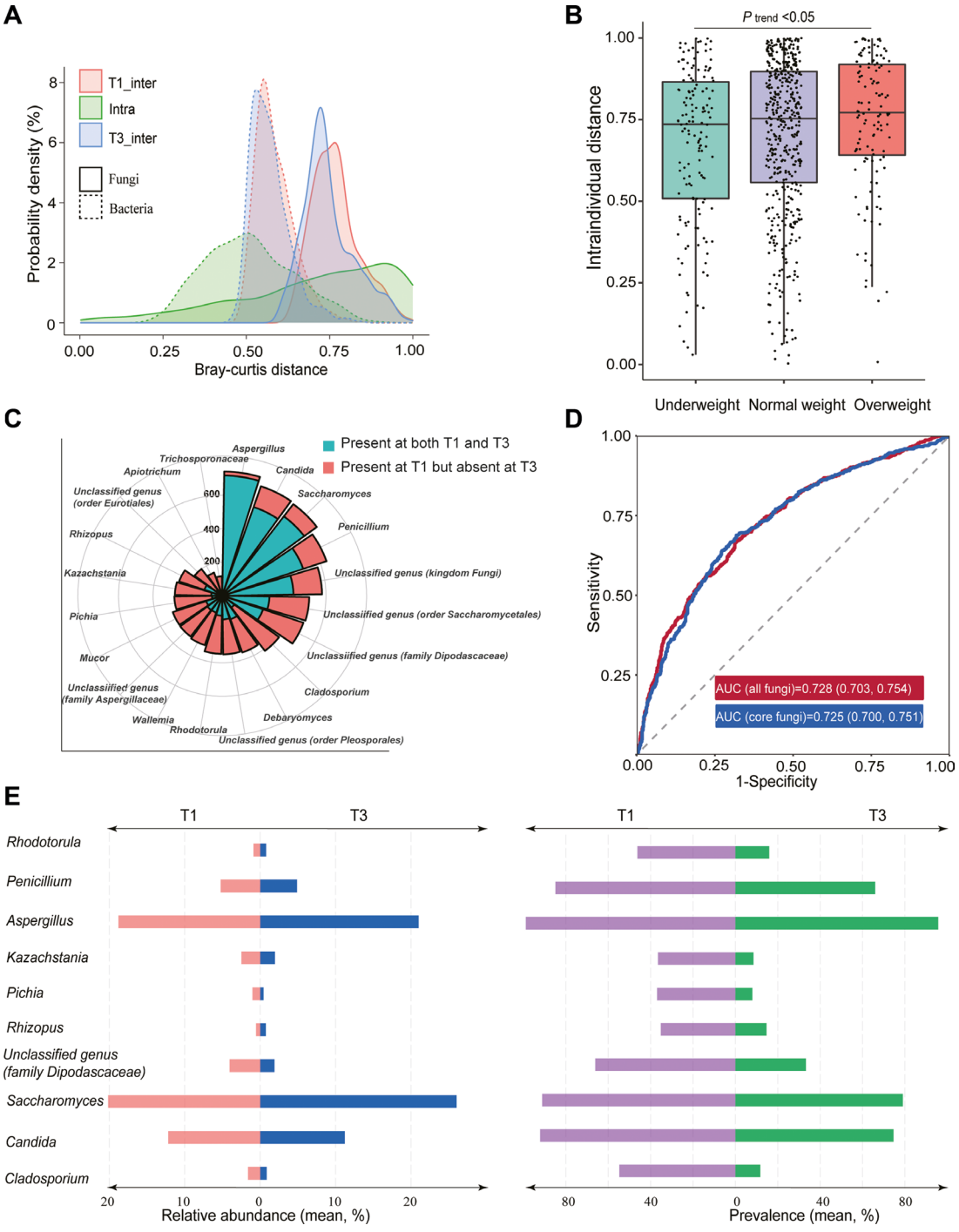


Figure 4 Discriminative gut fungal genera between early and late pregnancy. **(A)** The distributions of variation in gut fungal and bacterial composition over time (from T1 to T3) within individuals, as well as the differences between individuals at T1 or T3. (B) Comparison of the extent of gut mycobiome compositional alteration within individuals over time (from T1 to T3) stratified by prepregnancy overweight status. (C) Nightingale rose diagram visualising the proportion of participants whose core gut fungal genera were lost during later pregnancy in comparison to early pregnancy. (D) A machine learning framework, specifically LightGBM, was employed to train the trimester classifier on the gut mycobiome composition at T1 and T3. Subsequently, this trained classifier was used to predict the trimester to which the samples belong, employing a 10-fold cross-validation strategy and the corresponding area under the curve (AUC) values were presented. (E) The figure displays the relative abundance (left) and prevalence (right) of the top 10 gut fungal genera that contributed to the trimester classifier for T1 and T3.

These results together highlight the importance of body weight management before pregnancy for maintaining a stable gut microbial ecosystem during pregnancy.

We postulate that the considerable alterations in gut mycobiome composition during pregnancy may be attributed to the extremely high loss rate of the less prevalent fungal genera. The less prevalent (prevalence <40%) fungal genera accounted more than 96% of the identified genera during early pregnancy, and these genera showed an average loss rate of 97.6% from T1 to T3. On the other hand, this finding may also support that most of the identified gut fungi are passengers other than residents during pregnancy. Nevertheless, we identified a group of core

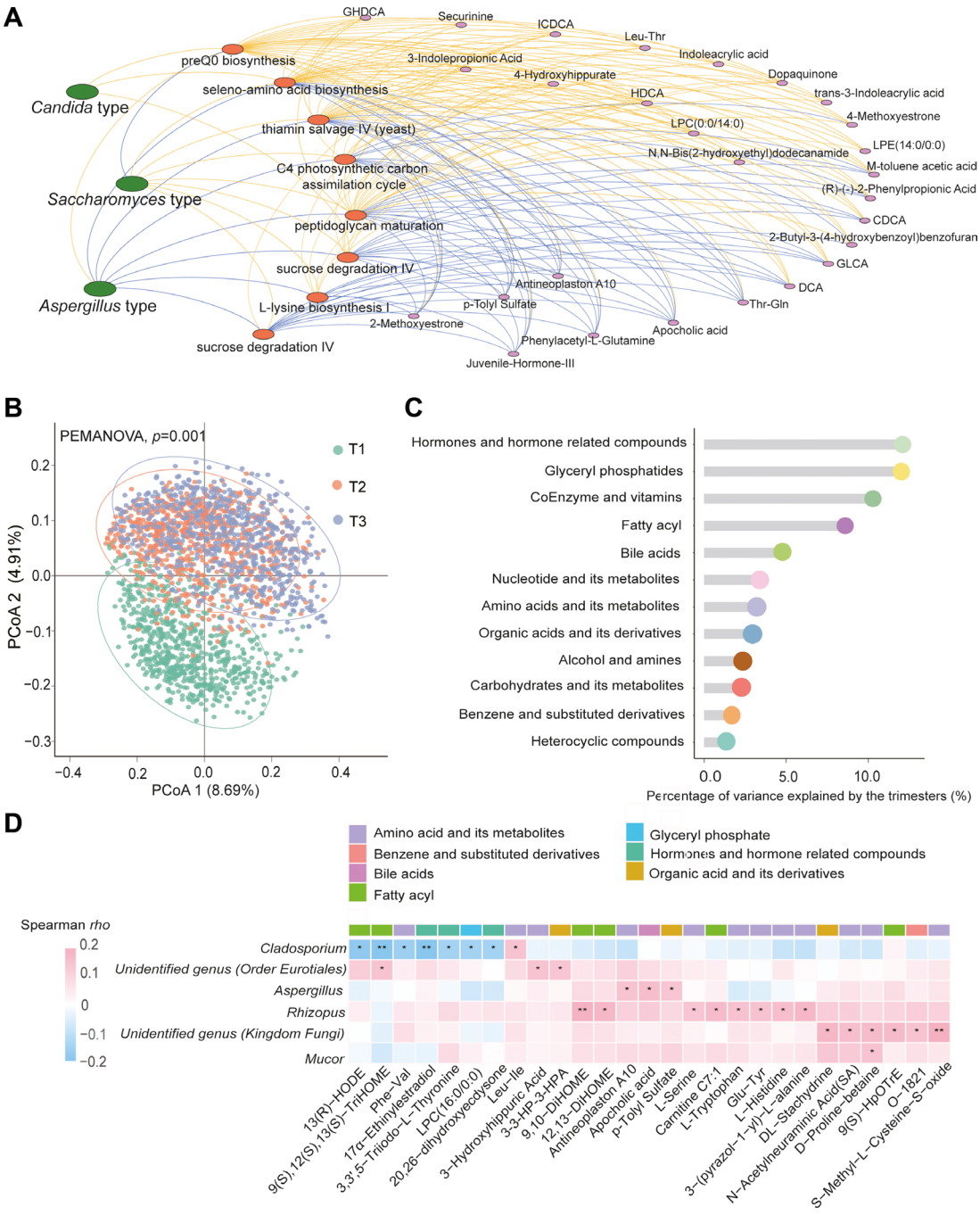


Figure 5 Interactions between gut mycobiome and host metabolism during pregnancy. (A) Network analysis among gut fungal enterotype, microbial functionality and host metabolome. The yellow and blue lines between gut fungal enterotype and functional pathways indicate enrichment and depletion of the pathways, respectively. The yellow and blue lines between functional pathways and serum metabolites indicate positive and negative associations, respectively. (B) Comparison on the overall metabolic pattern of serum samples collected during different trimesters of pregnancy. Principal coordinate analysis (PCoA) was employed, using Canberra dissimilarity, to examine the dissimilarities between all samples. Multivariate PERMANOVA analysis was performed to evaluate the extent to which trimester accounted for the variance in the overall metabolic pattern. The value of p was determined based on 999 permutations. (C) The extent of Canberra dissimilarity-based metabolic alterations during pregnancy for different classes of metabolites over time. To quantify these alterations, we assessed the explained variance of each class of metabolites by trimester using multivariate PERMANOVA. (E) Heatmap of covarying relationship between individual core fungal genera and individual serum metabolites from the first trimester to the third trimester. Spearman correlation analysis was performed between the changes in core gut fungi (CLR-transformed) and changes in metabolites. Significance levels are indicated as follows: *, FDR<0.05; **, FDR<0.01. GHDCA, glycohyodeoxycholic acid; ICDCA, isochodeoxycholic acid; HDCA, hyodeoxycholic acid; CDCA, chenodeoxycholic acid; GLCA, glycolithocholic acid, DCA, deoxycholic acid.

taxa for pregnant women in this study and about half of these core genera were also classified as core fungi for middle-aged and elderly individuals in our previous report and were consistently detected in the Human Microbiome Project and Danish cohorts.[9 16] Therefore, the identified core fungi are more likely to be resident commensals in the human gastrointestinal tract. This provides a rationale for future research focusing on these prevalent core fungi.

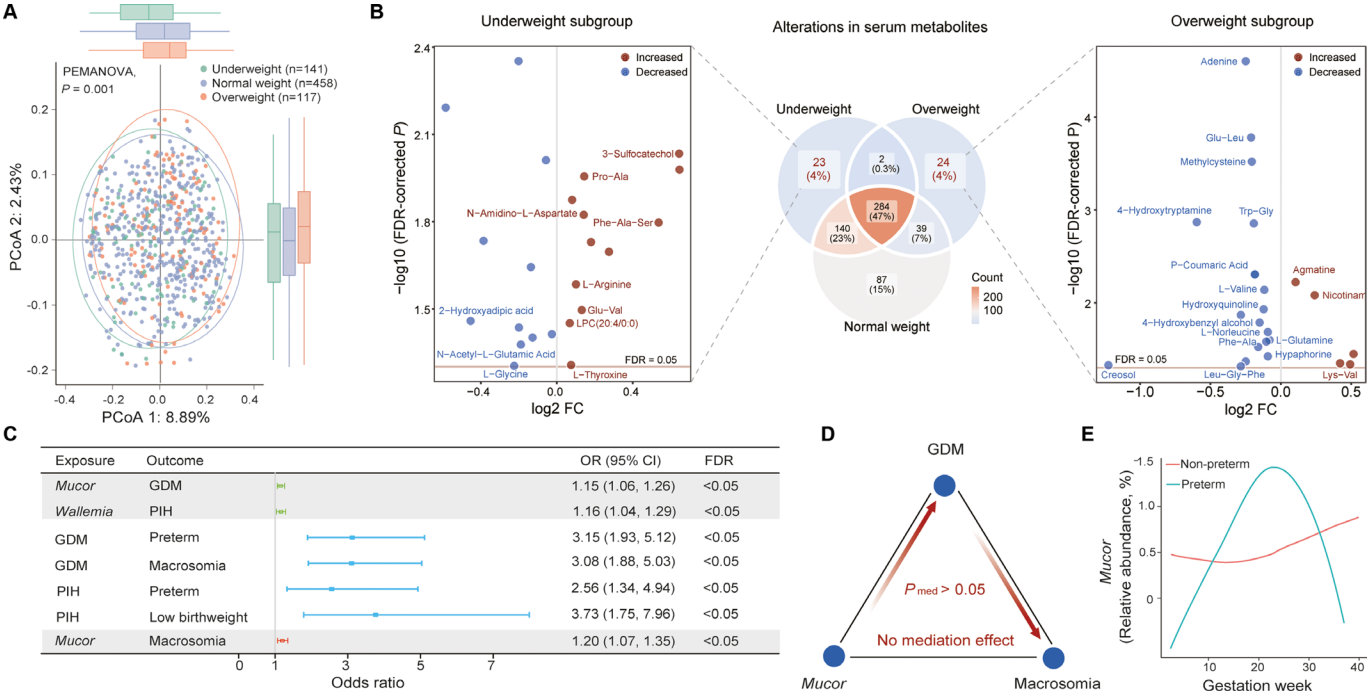


Figure 6 Distinctive metabolic changes stratified by prepregnancy overweight status and the clinical significance of the gut mycobiome during pregnancy. (A) Comparison on the overall metabolic dynamics between subgroups. Principal coordinate analysis (PCoA) was employed, using Canberra dissimilarity, to examine the dissimilarities between pregnant women with different prepregnancy overweight status. The value of p was determined based on 999 permutations. (B) Venn plot showing the number of distinctive and common metabolites that changed significantly from the first trimester to the third trimester. In each subgroup stratified by the prepregnancy overweight status, paired t-tests were conducted for each serum metabolite measured at T1 and T3. An FDR<0.05 was considered statistically significant. For those metabolites which significantly changed solely among pregnant women who were underweight or overweight prior to pregnancy, we showed the fold-change and value of p for each metabolite in the volcano plot. The x-axis shows the log2-transforemd fold-changes, and the y axis indicates the -log (base 10) of the FDR values. Red solid lines indicate the threshold of FDR=0.05. Red dots indicate those metabolites significantly increased from T1 to T3, while blue dots indicate metabolites that significantly decreased. Only the metabolites that were characterised with the highest accuracy and could be matched with internal standards were labelled with compound names in this plot. (C) The relationship between core fungal genera and pregnancy complications as well as adverse birth outcomes. Only the statistically significant associations (FDR<0.05) were illustrated in the forest plot. (D) Mediation analysis among the gut fungal genus *Mucor*, GDM and macrosomia. (E) Curves show LOESS fit for the relative abundance of the identified taxa based on preterm delivery (green) or not (red). The y-axis indicates the relative abundance. 13(R)-HODE, 13R-hydroxy-9Z,11E-octadecadienoic acid; TriHOME, trihydroxyoctadec-10-enoic acid; 3–3-HP-3-HPA, 3-(3-Hydroxyphenyl)−3-hydroxypropanoic acid; DiHOME, dihydroxyoctadec-12-enoic acid; 9(S)-HpOTrE, 9S-hydroperoxy-10E,12Z,15Z-octadecatrienoic acid; O-1821, 7-(3-Hydroxy-2-(3-hydroxy-5-phenylpent-1-enyl)−5-oxocyclopentyl)hept-5-enoic acid. GDM, gestational diabetes mellitus; LOESS, locally weighted regression; PIH, pregnancy-induced hypertension.

Prior studies have indicated host remodelling of the gut microbiome and metabolic changes during pregnancy, which may potentially impact maternal and infant health.[4 20–24] The majority of these studies have primarily focused on gut or vaginal bacteria and metabolites, while very few studies investigated the gut fungi during pregnancy based on a prospective cohort study. Here we present evidence that the interaction between the *Saccharomyces*-dominant enterotype, gut microbial functionality and host metabolism may be particularly significant. The network analysis demonstrated that the *Saccharomyces*-dominant enterotype was associated with more gut microbial functionalities compared with other fungal enterotypes. A previous study reported that the enterotype dominated by *Candida* conferred an increased risk of multiple diseases,[15] but we could not correlate the *Candida*-dominant enterotypes with pregnancy complications or adverse birth outcomes among pregnant women. Nevertheless, we found some specific core fungal genera, which were risk factors for pregnancy complications. The abundances of *Mucor* and *Wallemia* during early pregnancy were positively associated with the risk of GDM and PIH, respectively. We also validated the reported adverse effects of both GDM and PIH on birth outcomes, such as macrosomia, preterm birth and low birth weight.[25–27] Moreover, the abundance of *Mucor* during early pregnancy was also directly associated with the risk of macrosomia, which was independent of GDM. These findings collectively highlight that *Mucor* and *Wallemia* may serve as examples that support the vital role of gut fungi in impacting pregnancy health.

Although the relationship between *Mucor* and blood glucose homoeostasis has not been well studied, prior research reported that the *Mucor* was associated with higher blood glucose levels and inflammatory activity among patients with non-alcoholic fatty liver.[28] On the other hand, high blood glucose level is one of the important factors facilitating the growth of *Mucor*.[29] Our findings strengthen the relationship between *Mucor* and glucose metabolism disorders during pregnancy. Moreover, in a murine model, *Mucor* administration increased intestinal permeability in epithelial cell monolayers, which might be indicative of unstable intestinal barriers.[30] Of note, the increase in the permeability of the gut barrier is thought to contribute to systemic inflammation and diabetes development, and worsen the microvascular complications of existed diabetes.[31 32] A prior study investigating

the gut mycobiota of patients with GDM from middle to late pregnancy also supports that patients with GDM host a predominance of fungal taxa with potential inflammatory effects.[10] Regarding the potential impact of *Mucor* on fetal overgrowth, there are several species belonging to genus *Mucor* (eg, *Mucor circinelloides*) which can provide important alternative sources of bioactive lipids, due to its high efficiency in synthesising and accumulating lipids.[33] These findings may help propose potential mechanisms underlying the detrimental effects of *Mucor* on pregnancy health. Nevertheless, the pathogenic effects of fungi could be species-dependent or even strain-dependent.[11 34] Further studies are needed to validate the underlying mechanisms.

Wallemia has been reported to be a member of fungal microbiota in the human gut,[35] but it is less well known compared with *Candida* or *Saccharomyces*. *Wallemia* represents one of the most xerophilic fungal taxa, including the most xerophilic, osmophilic, and even halophilic and chaophilic microorganisms described to date.[36 37] Several species from the genus *Wallemia* were reported to produce several bioactive metabolites or toxins, which exhibited antiproliferative and antimicrobial activities.[38 39] Previous animal experiments reported that altered Schaedler flora mice colonised with *Wallemia mellicola* experienced enhanced severity of allergic airways disease compared with fungus-free control mice.[40] In the present study, *Wallemia* was positively associated the risk of PIH, while the underlying mechanisms are yet to be investigated. Nevertheless, *Wallemia* are found in various osmotically challenged environments, such as dry, salted or highly sugared foods, so we could not rule out that the *Wallemia* in the present study is just a biomarker of highly salted foods, which may confound our findings.

Our study has several strengths. First, the large sample size enables us to comprehensively profile the gut fungal characteristics and investigate determinants of variations in gut mycobiome composition among pregnant women. Second, the well-established subcohort with longitudinally repeated sample collections and ITS2 sequencing, facilitate the extensive profiling of gut fungal dynamics during pregnancy. Finally, the multiomics data including ITS2 and 16S sequencing, short-gun metagenomics sequencing and serum metabolomics data provide a landscape of the networks among gut fungi, biological functionality and host metabolites. Our study also has several limitations. The pathogenic or probiotic effects of gut fungi could be species-dependent or even strain-dependent, therefore the resolution at the genus level is not high enough in the present study. Future research may improve the accuracy of reference databases for fungal taxon alignment based on metagenomics data. Second, the quantification of gut fungi and bacteria relies on relative abundance measurements, which may introduce unexpected bias into the statistical analysis. Even when data transformation techniques or appropriate methods addressing the compositional nature of the data are employed, the potential bias may still exist. Third, although we perform correlation analysis to gain functional insights of the gut mycobiome, we could not directly annotate the functions specific to gut fungi due to the limited reference databases. In the future, it will be crucial to enhance the function annotation of the human gut mycobiome. Lastly, this study only includes Chinese women, which may inevitably limit the generalisability of our findings.

In summary, we provide evidence for the dynamic nature of gut fungi in comparison to gut bacteria during pregnancy, revealing that prepregnancy overweight status may be a key determinant for this alteration. This study presents a landscape of the networks among the gut mycobiome, biological functionality, serum metabolites and pregnancy health, pinpointing the link between fungal genus *Mucor* and adverse pregnancy outcomes. This study also provides a reference database and resource for future investigation on the functional role of gut mycobiome in healthy pregnancy.

Author affiliations
[1]Affiliated Hangzhou First People's Hospital, School of Medicine, Westlake University, Hangzhou, China
[2]Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, China
[3]Research Center for Industries of the Future, School of Life Sciences, Westlake University, Hangzhou, China
[4]Department of Epidemiology and Biostatistics, Ministry of Education Key Laboratory of Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[5]Department of Science and Education, Shuangliu Maternal and Child Health Hospital, Chengdu, China
[6]Antenatal Care Clinics, Shuangliu Maternal and Child Health Hospital, Chengdu, China
[7]Section of Epidemiology and Population Health & Department of Obstetrics and Gynecology, Ministry of Education Key Laboratory of Birth Defects and Related Diseases of Women and Children & National Medical Products Administration Key Laboratory for Technical Research on Drug Products In Vitro and In Vivo Correlation, West China Second University Hospital, Sichuan University, Chengdu, China
[8]Department of Laboratory Medicine, Ministry of Education Key Laboratory of Birth Defects and Related Diseases of Women and Children, West China Second University Hospital, Sichuan University, Chengdu, China
[9]Department of Obstetrics and Gynecology, Ministry of Education Key Laboratory of Birth Defects and Related Diseases of Women and Children, West China Second University Hospital, Sichuan University, Chengdu, China
[10]Shuangliu Institute of Women's and Children's Health, Shuangliu Maternal and Child Health Hospital, Chengdu, China
[11]West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu, China
[12]Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, China
[13]Zhejiang Key Laboratory of Multi-Omics in Infection and Immunity, Center for Infectious Disease Research, School of Medicine, Westlake University, Hangzhou, China

X Ju-Sheng Zheng @zheng_jusheng

Acknowledgements The authors thank the High-Performance Computing Center and High-Throughput Core Facility at Westlake University for assistance in computing and data generation.

Contributors J-SZ, AP, X-FP and YF conceived the study concept and design. YF, WG, PW and YL analysed the data. PW, YL, JY, BZ, YY, XL, YH contributed to the field work, data collection and data curation. YF, KZ, XL, MS, JT, ZM and JC contributed to visualisation of the data. YF and J-SZ wrote the first draft of the manuscript. AP and X-FP contributed to the critical revision of the manuscript. YF, WG, PW and YL contributed equally to the work. All authors approved the final version of the manuscript for publication. J-SZ, AP and X-FP are the guarantors of this work and, as such, had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding This study was funded by the National Key R&D Program of China (2022YFA1303900, 2023YFC3606300 and 2022YFC3600600), the National Natural Science Foundation of China (82103826,82073529, U21A20427, 82325043, 81930124, 82192902, 82021005 and 92374112), 'Pioneer' and 'Leading goose' R&D Program of Zhejiang (2022C03102), Zhejiang Provincial Natural Science Foundation of China (LQ21H260002), the Research Program of Westlake Laboratory of Life Sciences and Biomedicine (202208012), the Fundamental Research Funds for the Central Universities (YJ202346), China Postdoctoral Science Foundation (2023M733177, 2022M722833). The funders had no role in collecting data, study design, interpretation of data or the decision to submit the manuscript for publication.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Ethics Committee of the Tongji Medical College, Huazhong University of Science and Technology (2017) (S225)-1. Participants gave informed consent to participate in the study before taking part.

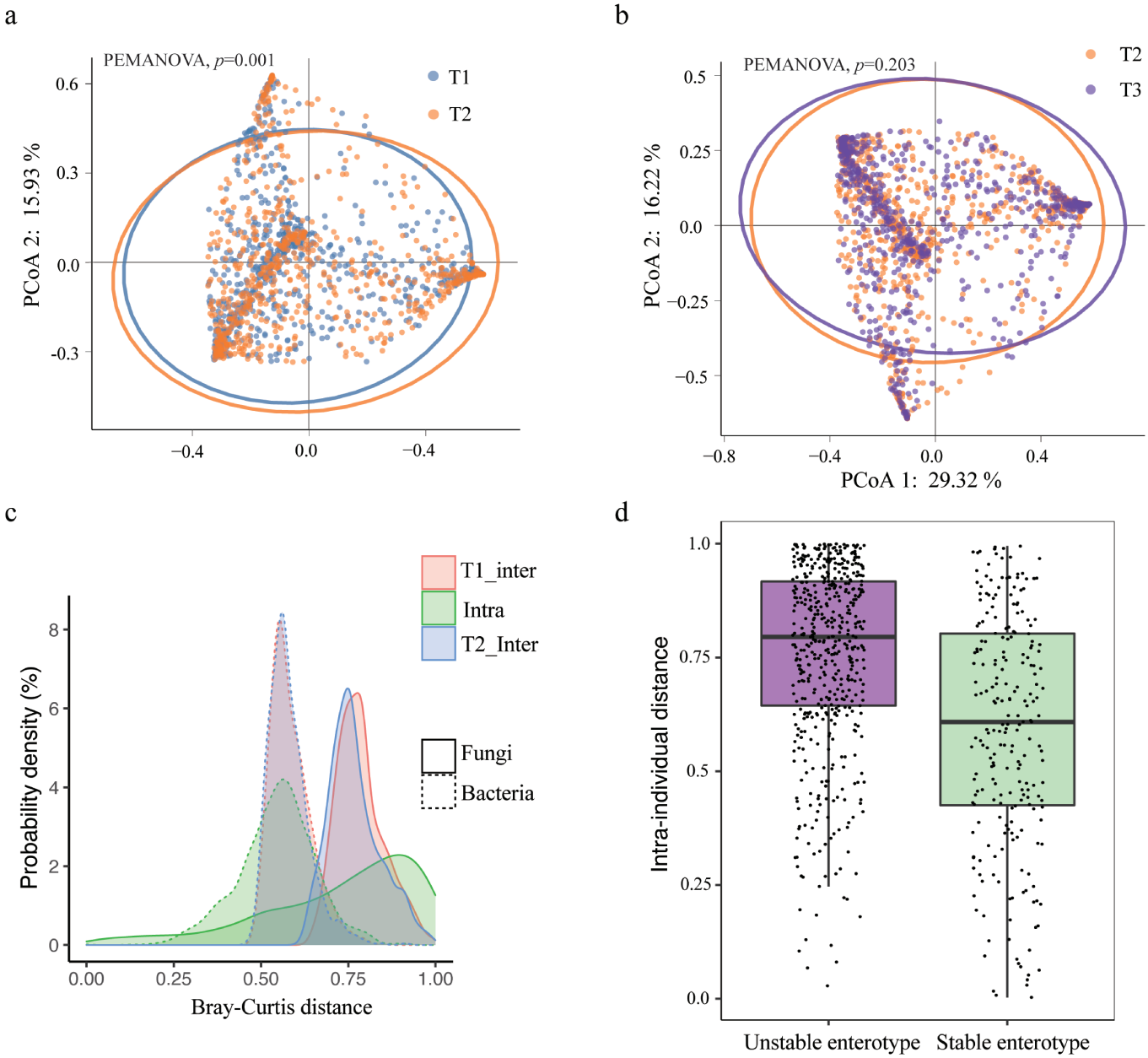
Provenance and peer review Not commissioned; externally peer reviewed.

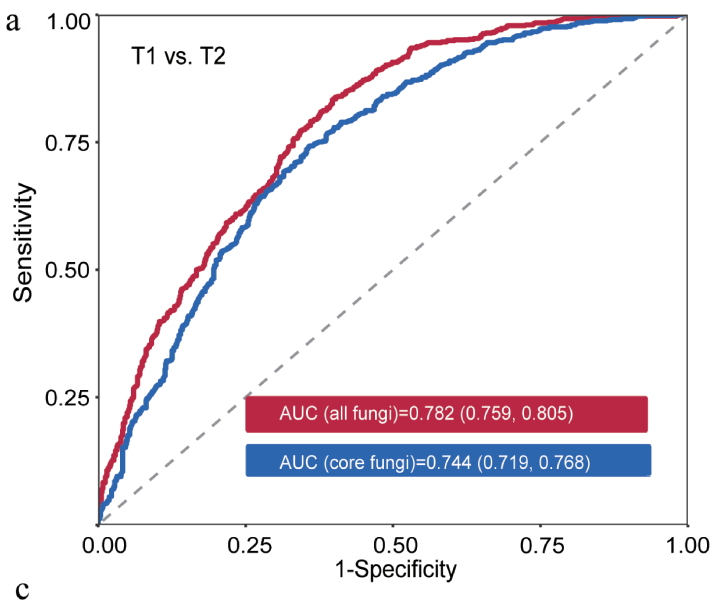
Data availability statement Data are available in a public, open access repository. The raw data of ITS2, 16S and metagenomic sequencing in this study have been deposited in the Genome Sequence Archive (GSA) (https://ngdc.cncb.ac.cn/gsa/) at accession number CRA014764; CRA014766; CRA014529. Analysis R codes, Stata codes as well as the annotation pipelines for gut fungi, gut bacteria and gut microbial function are available via https://github.com/nutrition-westlake/THSBC-gut_fungi.Other data sets generated during and/or analysed during the current study are available from the corresponding author upon reasonable request.

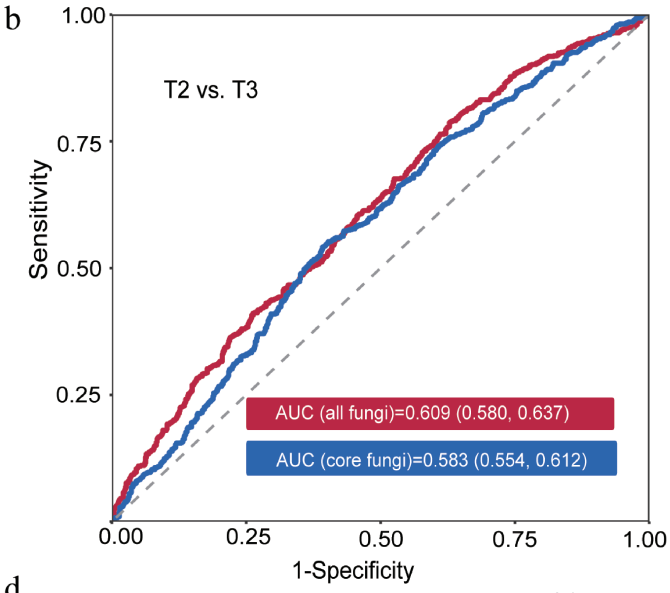
ORCID iDs
Yuanqing Fu http://orcid.org/0000-0002-3955-9376
Ju-Sheng Zheng http://orcid.org/0000-0001-6560-4890

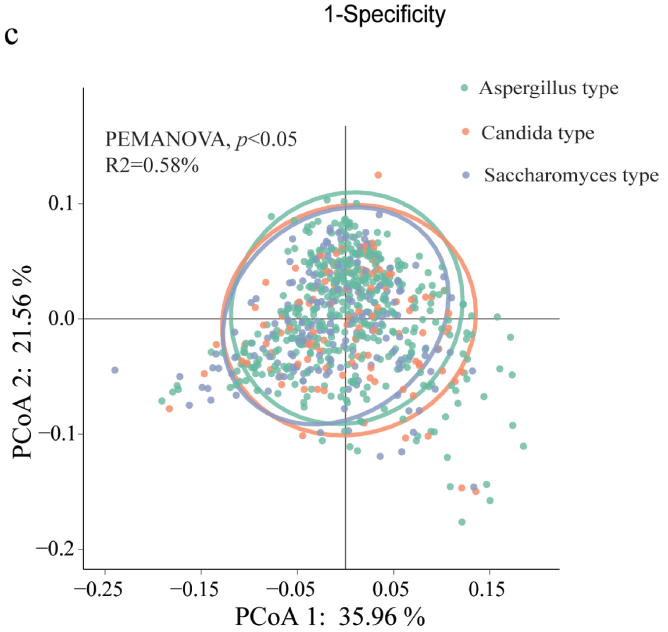
REFERENCES

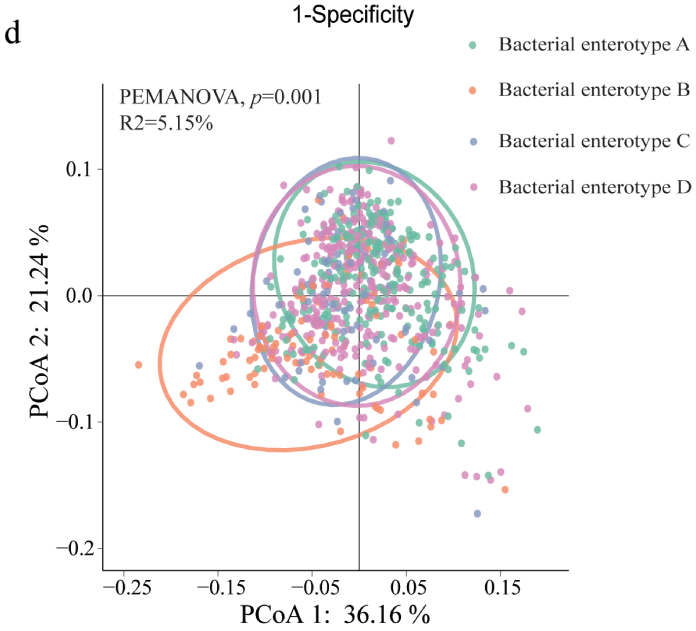
1 Newbern D, Freemark M. Placental hormones and the control of maternal metabolism and fetal growth. *Current Opinion in Endocrinology, Diabetes & Obesity* 2011;18:409–16.
2 Mor G, Cardenas I. The immune system in pregnancy: a unique complexity. *American J Rep Immunol* 2010;63:425–33.
3 Evans JM, Morris LS, Marchesi JR. The gut Microbiome: the role of a virtual organ in the Endocrinology of the host. *J Endocrinol* 2013;218:R37–47.
4 Koren O, Goodrich JK, Cullender TC, *et al*. Host remodeling of the gut Microbiome and metabolic changes during pregnancy. *Cell* 2012;150:470–80.
5 Underhill DM, Iliev ID. The Mycobiota: interactions between Commensal fungi and the host immune system. *Nat Rev Immunol* 2014;14:405–16.
6 Hamad I, Raoult D, Bittar F. Repertory of Eukaryotes (Eukaryome) in the human gastrointestinal tract: Taxonomy and detection methods. *Parasite Immunol* 2016;38:12–36.
7 Gouba N, Drancourt M. Digestive tract Mycobiota: a source of infection. *Med Mal Infect* 2015;45:9–16.
8 Zheng J-S, Wahlqvist M. Regulobiosis: a regulatory and food system-sensitive role for fungal Symbionts in human evolution and Ecobiology. *Asia Pac J Clin Nutr* 2020;29:9–15.
9 Shuai M, Fu Y, Zhong H-L, *et al*. Mapping the human gut Mycobiome in middle-aged and elderly adults: Multiomics insights and implications for host metabolic health. *Gut* 2022;71:1812–20.
10 Ferrocino I, Ponzo V, Pellegrini M, *et al*. Mycobiota composition and changes across pregnancy in patients with gestational diabetes mellitus (GDM). *Sci Rep* 2022;12:9192.
11 Vavreckova M, Galanova N, Kostovcik M, *et al*. Specific gut bacterial and fungal Microbiota pattern in the first half of pregnancy is linked to the development of gestational diabetes mellitus in the cohort including obese women. *Front Endocrinol (Lausanne)* 2022;13:970825.
12 Wu N, Mo H, Mu Q, *et al*. The gut Mycobiome characterization of gestational diabetes mellitus and its association with dietary intervention. *Front Microbiol* 2022;13:1–13.
13 Limon JJ, Tang J, Li D, *et al*. Malassezia is associated with Crohn's disease and exacerbates colitis in Mouse models. *Cell Host Microbe* 2019;25:377–88.
14 Zuo T, Wong SH, Cheung CP, *et al*. Gut fungal Dysbiosis correlates with reduced efficacy of fecal Microbiota transplantation in Clostridium difficile infection. *Nat Commun* 2018;9:3663.
15 Lai S, Yan Y, Pu Y, *et al*. Enterotypes of the human gut Mycobiome. *Microbiome* 2023;11:179.
16 Nash AK, Auchtung TA, Wong MC, *et al*. The gut Mycobiome of the human Microbiome project healthy cohort. *Microbiome* 2017;5:153.
17 Stanislawski MA, Dabelea D, Wagner BD, *et al*. Pre-pregnancy weight, gestational weight gain, and the gut Microbiota of mothers and their infants. *Microbiome* 2017;5:113.
18 Cömert TK, Akpinar F, Erkaya S, *et al*. The effect of pre-pregnancy obesity on gut and Meconium Microbiome and relationship with fetal growth. *The Journal of Maternal-Fetal & Neonatal Medicine* 2022;35:10629–37.
19 Liu T, Jia F, Differding MK, *et al*. Pre-pregnancy body mass index and gut Microbiota of mothers and children 5 years postpartum. *Int J Obes* 2023;47:807–16.
20 Cortez RV, Taddei CR, Sparvoli LG, *et al*. Microbiome and its relation to gestational diabetes. *Endocrine* 2019;64:254–64.
21 Liu H, Pan L-L, Lv S, *et al*. Alterations of gut Microbiota and blood Lipidome in gestational diabetes mellitus with hyperlipidemia. *Front Physiol* 2019;10:1015.
22 Tu X, Duan C, Lin B, *et al*. Characteristics of the gut Microbiota in pregnant women with fetal growth restriction. *BMC Pregnancy Childbirth* 2022;22:297.
23 Fernandez-Gonzalez S, Ortiz-Arrabal O, Torrecillas A, *et al*. Study of the fetal and maternal Microbiota in pregnant women with Intrauterine growth restriction and its relationship with inflammatory biomarkers: a case-control study protocol. *Medicine* 2020;99:e22722.
24 Kuang Y-S, Lu J-H, Li S-H, *et al*. Connections between the human gut Microbiome and gestational diabetes mellitus. *Gigascience* 2017;6:1–12.
25 Bertagnolli M, Luu TM, Lewandowski AJ, *et al*. Preterm birth and hypertension: is there a link *Curr Hypertens Rep* 2016;18:28.
26 Preda A, Iliescu D-G, Comănescu A, *et al*. Gestational diabetes and Preterm birth: what do we know? our experience and mini-review of the literature. *J Clin Med* 2023;12:1–10.
27 Kc K, Shakya S, Zhang H. Gestational diabetes mellitus and Macrosomia: a literature review. *Ann Nutr Metab* 2015;66 Suppl 2:14–20.
28 Demir M, Lang S, Hartmann P, *et al*. The fecal Mycobiome in non-alcoholic fatty liver disease. *J Hepatol* 2022;76:788–99.
29 Prabhu S, In A, Balakrishnan D. Dental perspective on Mucormycosis in COVID-19: a literature review. *Curr Oral Health Rep* 2022;9:211–4.
30 Mueller KD, Zhang H, Serrano CR, *et al*. Gastrointestinal Microbiota alteration induced by Mucor Circinelloides in a murine model. *J Microbiol* 2019;57:509–20.
31 Snelson M, de Pasquale C, Ekinci EI, *et al*. Gut Microbiome, Prebiotics, intestinal permeability and diabetes complications. *Best Pract Res Clin Endocrinol Metab* 2021;35:101507.
32 Mønsted MØ, Falck ND, Pedersen K, *et al*. Intestinal permeability in type 1 diabetes: an updated comprehensive overview. *J Autoimmun* 2021;122:102674.
33 Fazili ABA, Shah AM, Zan X, *et al*. Mucor Circinelloides: a model organism for Oleaginous fungi and its potential applications in bioactive lipid production. *Microb Cell Fact* 2022;21:29.
34 Mar Rodríguez M, Pérez D, Javier Chaves F, *et al*. Obesity changes the human gut Mycobiome. *Sci Rep* 2015;5:14600.
35 Paterson MJ, Oh S, Underhill DM. Host–microbe interactions: Commensal fungi in the gut. *Curr Opin Microbiol* 2017;40:131–7.
36 Zajc J, Kogej T, Galinski EA, *et al*. Osmoadaptation strategy of the most Halophilic fungus, Wallemia Ichthyophaga, growing Optimally at Salinities above 15% Nacl. *Appl Environ Microbiol* 2014;80:247–56.
37 Zalar P, Sybren de Hoog G, Schroers H-J, *et al*. Taxonomy and Phylogeny of the Xerophilic genus Wallemia (Wallemiomycetes and Wallemiales, Cl. et ord. Nov. *Antonie Van Leeuwenhoek* 2005;87:311–28.
38 Jančič S, Frisvad JC, Kocev D, *et al*. Production of secondary metabolites in extreme environments: Food- and airborne Wallemia Spp. produce toxic metabolites at Hypersaline conditions. *PLoS One* 2016;11:e0169116.
39 Wood GM, Mann PJ, Lewis DF, *et al*. Studies on a toxic metabolite from the mould Wallemia. *Food Addit Contam* 1990;7:69–77.
40 Skalski JH, Limon JJ, Sharma P, *et al*. Expansion of Commensal fungus Wallemia Mellicola in the gastrointestinal Mycobiota enhances the severity of allergic airway disease in mice. *PLoS Pathog* 2018;14:e1007260.


a
PEMANOVA, p=0.001
T1
T2
PCoA 2: 15.93 %
b
PEMANOVA, p=0.203
T2
T3
PCoA 2: 16.22 %
PCoA 1: 29.32 %
c
T1_inter
Intra
T2_Inter
Fungi
Bacteria
Probability density (%)
Bray-Curtis distance
d
Intra-individual distance
Unstable enterotype
Stable enterotype


a
T1 vs. T2
Sensitivity
1-Specificity
AUC (all fungi)=0.782 (0.759, 0.805)
AUC (core fungi)=0.744 (0.719, 0.768)


b
T2 vs. T3
Sensitivity
1-Specificity
AUC (all fungi)=0.609 (0.580, 0.637)
AUC (core fungi)=0.583 (0.554, 0.612)


c
Aspergillus type
Candida type
Saccharomyces type
PEMANOVA, p<0.05
R2=0.58%
PCoA 2: 21.56 %
PCoA 1: 35.96 %


d
Bacterial enterotype A
Bacterial enterotype B
Bacterial enterotype C
Bacterial enterotype D
PEMANOVA, p=0.001
R2=5.15%
PCoA 2: 21.24 %
PCoA 1: 36.16 %

Supplemental Materials

Landscape of the gut mycobiome dynamics during pregnancy and its relationship with host metabolism and pregnancy health

This file includes:

Methods

(1) Study population

(2) Sample collection

(3) Questionnaires and clinical data collection

(4) Bioinformatic analyses

(5) Statistical analysis

Supplemental Figure S1 to S2

(1) Supplemental Figure S1. Supplementary Table S1: Characteristics of demographic, clinical and dietary factors stratified by early-pregnancy fungal enterotypes.

(2) Supplemental Figure S2. The association between changes in consumption of food groups during pregnancy and alterations in gut fungal richness.

Fu Y, *et al. Gut* 2024;0:1–11. doi: 10.1136/gutjnl-2024-332260

Methods

Study population

The current study was based on the Tongji-Huaxi-Shuangliu Birth Cohort (THSBC), in which pregnant women aged 18 to 41 years were recruited during the early pregnancy when they presented to antenatal care clinics in in a local maternal and child health hospital during their early pregnancy. Exclusion criteria were 1) receiving infertility treatment (e.g., in vitro fertilization or intrauterine insemination); 2) reporting severe chronic or infectious diseases (e.g., cancer, HIV infection, or tuberculosis); or 3) were unable to or refused to sign the informed consent. The THSBC study was approved by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology (No. [2017](S225)-1), and informed consent was obtained from all participants.

In the present analysis, we included 4800 participants who had available ITS2 sequencing data, dietary information and clinical records during their pregnancy. This dataset enables us to comprehensively profile the gut mycobiome among pregnant women and investigate potential determinants contributing to the variations of gut mycobiome. To examine how pregnancy impacts the gut mycobiome over time and investigate their potential associations with host metabolism, we established a sub-cohort of 1059 participants, which included 514 women who gave birth to preterm (n=240), low birthweight (n=137), or macrosomia (n=216) infants, as well as 545 randomly selected participants who did not experience the above three adverse pregnancy outcomes.

ITS2 sequencing was performed for all the 4800 participants, while the shotgun metagenomics sequencing was performed for T1 samples within the established sub-cohort (n=1059). Additionally, within the sub-cohort, 750 and 748 participants had ITS2 and 16S sequencing data available, respectively, for all trimesters. We also repeatedly measured serum metabolome throughout each trimester of pregnancy for participants in this selected sub-cohort using an LC-ESI-MS/MS system.

Fu Y, *et al. Gut* 2024;0:1–11. doi: 10.1136/gutjnl-2024-332260

Sample collection

Stool samples were collected and stored in ice boxes at the hospital or home by the participants under instructions and then transferred to the hospital to store at −40 °C within 24 hours. A detailed standard operating procedure was given to the pregnant women for instructions on fecal sample collection, temporary storage, and transportation to the hospital. Stocks of frozen fecal samples were transported by dry ice every 2 to 3 months to the laboratory in Huazhong University of Science and Technology and stored at -80 °C before further processing.

Questionnaires and clinical data collection

All participants completed a set of structured questionnaires on sociodemographic information, lifestyle and behaviors (such as cigarette smoking and alcohol drinking), diet (including recent daily intakes of major food groups), history of pregnancy and births, history of diseases and medications, and family history of diseases. Anthropometric and blood pressure measurements were collected on site using devices according to standard protocols. Body mass index (BMI) was calculated by dividing the weight in kilograms by the square of height in meters. Underweight, overweight and obesity was defined as BMI<18.5, BMI ≥ 24 kg/m^2 and BMI ≥ 28 kg/m^2, respectively. Pregnancy complications (e.g., gestational diabetes mellitus) and birth outcomes were extracted from the electronic clinical records.

Bioinformatic analyses

Gut mycobiome analysis using ITS2 rRNA gene sequencing data

Microbial DNA was extracted using the E.Z.N.A.® soil DNA Kit (Omega Bio-tek, Norcross, GA, U.S.) according to manufacturer's protocols. The final DNA concentration and purification were determined by NanoDrop 2000 UV-vis spectrophotometer (Thermo Scientific, Wilmington, USA), and DNA quality was checked by 1% agarose gel electrophoresis. The ITS2 hypervariable regions of the fungal ITS rRNA gene were amplified with primers ITS3F: GCATCGATGAAGAACGCAGC and ITS4R: TCCTCCGCTTATTGATATGC by

thermocycler PCR system (GeneAmp 9700, ABI, USA). Purified amplicons were pooled in equimolar and paired-end sequenced (2 × 250) on an Illumina NovaSeq platform (Illumina, San Diego, USA) according to the standard protocols by Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China).

The mean sequencing depth and its standard deviation for all samples were 118,177 and 7955, respectively. It is worth noting that within the sub-cohort analyzed longitudinally, the number of sequencing reads remained relatively consistent across trimesters. Specifically, the mean (SD) sequencing depths were 117,906 (8350), 118,882 (6125), and 116,688 (9198) for samples collected during the first, second, and third trimester of pregnancy, respectively. The demultiplexed ITS2 sequences were denoised and grouped into amplicon sequence variants (ASVs; i.e., 100% exact sequence match) using DADA2.[1] During the process, marker gene Illumina sequence data and low-quality regions of the sequences were detected and filtered. We trimmed 28 bases (primer and barcode) from the beginning of the sequences. We also truncated the sequences at the 245 bases as the quality dropped around position 245 (median of quality score <30). The ASV features that were presented in only one sample were excluded as suggested by the Qiime2 tutorial, based on the suspicion that these may not represent real biological diversity but rather PCR or sequencing errors. The individual ASVs were taxonomically classified based on the UNITE (version 8.2, 99%) database using the VSEARCH tool wrapped in QIIME2 (version 2021.2).[2] α-diversity analysis was conducted through the q2-diversity plugin at the sampling depth of 10000. α-diversity was estimated by Shannon's diversity index (or Shannon; a quantitative measure of community richness and evenness), Observed Features (or Richness; a qualitative measure of community richness), and Faith's PD (or Faith's Phylogenetic Diversity; a qualitative measure of community richness that incorporates phylogenetic relationships between the observed features).

Gut bacteria analysis using 16S rRNA gene sequencing data

For the 16S analysis, raw sequencing reads were merge-paired, quality filtered and

analyzed using QIIME2 (version 2021.2). As described above, we used DADA2 denoised-paired plugin in QIIME2 to process the fastq files. We filtered the features that were present in only a single sample. The taxonomies of ASVs were subsequently determined using the Naive Bayes classifier trained on the Sliva_138 99% reference database. α-diversity analysis was conducted at the sampling depth of 10000. A diversity of the gut bacteria was estimated by the indices the same as ITS2 data.

Microbial functional profiling using metagenome data

Microbial DNA extractions were carried out by a standardized CTAB procedure. DNA concentration was measured using Qubit dsDNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, CA, USA). For DNA library preparation, a total amount of 1μg DNA per sample was used. In addition, the NEBNext Ultra DNA Library Prep Kit (NEB, USA) was used following manufacturer's recommendations and index codes were added to attribute sequences to each sample. The DNA samples were fragmented by sonication to a size of approximately 350 bp. Then, the DNA fragments were end-polished, A-tailed, and ligated with the full-length adaptor for Illumina sequencing with further PCR amplification. Thereafter, PCR products were purified (AMPure XP system) and libraries were analyzed for size distribution by Agilent2100 Bioanalyzer and quantified using real-time PCR. The clustering of the index-coded samples was performed on a cBot Cluster Generation System according to the manufacturer's instructions. Lastly, sequencing was performed using the Illumina NovaSeq platform at Shanghai Personal Biotechnology Co. Ltd. (Shanghai, China) and 150 bp paired-end reads were generated.

Next, raw sequencing reads were first quality-controlled with KneadData toolkit (v0.10.0): 1) to trim the reads by quality score from the 5′ end and 3′ end with a quality threshold of 20; 2) removed read pairs when either read was < 50 bp, contained "N" bases or quality score mean below 30; and 3) deduplicate the reads. Reads aligning to the human genome (H. sapiens, UCSC hg38) were removed via KneadData integrated with Bowtie2 tool (v2.4.5).

Functional profiling was performed with HUMAnN3 (v3.0.1), which maps sample reads against the sample-specific reference database to quantify gene presence and abundance in a species-stratified manner, with unmapped reads further used in a translated search against Uniref90 to include taxonomically unclassified but functionally distinct gene family abundances. We extracted the Uniref90 gene families of gut bacteria for downstream analyses. The Uniref90 gene families were then converted into relative abundances of unstratified pathway.

Serum metabolomics profiling

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/) at Wuhan Metware Biotechnology Co., Ltd. (Wuhan, China). LIT and triple quadrupole (QQQ) scans were acquired on a triple quadrupole-linear ion trap mass spectrometer (QTRAP), QTRAP® LC-MS/MS System, equipped with an ESI Turbo Ion-Spray interface, operating in positive and negative ion mode and controlled by Analyst 1.6.3 software (Sciex). Instrument tuning and mass calibration were performed with 10 and 100 μmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as MRM experiments with collision gas (nitrogen) set to 5 psi. Declustering potential (DP) and collision energy (CE) for individual MRM transitions was done with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period. The mass spectrum data were processed using Software Analyst 1.6.3. The metabolite identification was conducted by referencing standards in self-built metware database and public databases. The identified metabolites were matched with the parent ion mass-to-charge ratio, the fragment ion mass-to-charge ratio as well as retention time of their corresponding standards. The accuracy of metabolite characterization was classified into three levels, depending on the presence of isotope internal standards or matching score with the secondary mass spectrometry. The matching score>0.7 indicates the level 1 accuracy of metabolite characterization, while 0.5-0.7 and <0.5 indicate the level 2 and level 3 accuracy, respectively.

Statistical analysis

Determinants of gut fungal composition among pregnant women

The gut mycobiome data were analyzed at the genus level. We investigated the determinants of the gut fungal composition using the data collected from 4800 participants during the first trimester, with each participant contributing only 1 stool sample. We used the *vegdist* function in the R package *vegan* to calculate the gut fungal Bray-Curtis dissimilarity matrix. The contribution of 20 environmental variables (including demographics, physiologic traits, diseases, and habitual dietary intakes) to fungal community variation was determined by PERMANOVA analysis using the function adonis2 in *vegan*.[3] We applied a complete data analysis strategy which excluded 19 participants with missing values for at least one of the environmental variables. The 4800 samples included in this analysis were sequenced in two separate batches. Therefore, we included the batch information as a covariate in the model, to adjust for potential batch effects. The p value was determined through 999 permutations.

Gut fungal and bacterial enterotype clustering

The fecal samples (T1, n=4800; T2, n=890; T3, n=850) of ITS2 amplification were clustered into fungal enterotypes by using a partitioning around medoid (PAM) clustering method as those previously described.[4] Briefly, the samples were grouped into clusters with partitioning around medoid (PAM) based on the between-sample Bray–Curtis distance calculated at genus-level. The optimal number of clusters was determined by the silhouette index. The driver genus of each enterotype was determined as the genus with the highest relative abundance in the enterotype. The fecal samples of 16S amplification were clustered into bacterial enterotypes by using the method as that for bacterial enterotype.

Dynamics of within-sample α diversity throughout each trimester of pregnancy

This analysis was conducted in the established sub-cohort of 750 participants who had available gut fungi α diversity data for each trimester of pregnancy. We utilized

paired t-tests to assess the statistical significance of the changes in fungal α diversity between T1 and T2, as well as between T2 and T3 independently.

Among the 750 participants included in the study to profile changes in gut fungal richness from T1 to T3, data on the estimation of consumption changes in food groups from T1 to T3 based on the FFQ were available for 639 participants. These food groups consisted of rice, steamed bread, noodles, vegetables, meat, eggs, milk, and fruit. The weight of each food group consumed per day was quantified based on the FFQ. We applied a linear regression model to estimate the association between the changes in richness from T1 to T3 and the respective food group. As covariates, we incorporated age, pre-pregnancy BMI, interval time between sample collections, parity, and gravidity. An FDR<0.05 was considered statistically significant.

Loss rate calculations and discriminative genera identification

Utilizing the available repeated measurements of gut fungi throughout pregnancy within the designated sub-cohort (n=750), we elucidated the loss rate for each fungal genus as the host underwent progression from T1 to T3. The loss rate for each fungal genus was determined by quantifying the proportion of the decline in frequency observed between T1 and T3.

To assess the gut fungi enriched or depleted during early or late pregnancy, we conducted an analysis using paired t test analysis. This analysis was based on the examination of 465 genera that were detected at either T1 or T3. We transformed the taxa data using the centered log-ratio (CLR) method to address the compositional nature of the mycobiome data before we perform the paired t test analysis. To determine statistical significance, a false discovery rate (FDR)-adjusted p-value threshold of less than 0.05 was used.

Additionally, we used these 465 fungal genera to construct a machine learning framework of LightGBM for predicting the trimester that the samples belong to.[5] The construction of prediction model was based on Scikit-learn (v0.15.2), and ten-

fold cross validation (CV) was applied. To evaluate binary classification performance, receiver operating characteristic (ROC) curve analyses were conducted using the R package pROC. We used the SHAP (Shapley Additive exPlanations) to interpret predictions and the importance of each fungal genus to the prediction model is represented using Shapley values.[6]

Quantification of intra-individual gut fungi compositional alterations

We applied the *vegdist* function from the R package *vegan* to calculate the Bray-Curtis distance based on gut fungal genus-level composition to assess intra-individual and inter-individual dissimilarities, respectively.[3] The sub-cohort, consisting of 750 participants, was included in this analysis, as they had relevant data throughout each trimester of pregnancy. To determine the intra-individual distance, paired data was used, with the gut mycobiome data at T1 serving as the reference. Subsequently, a Bray-Curtis distance value was calculated for each participant in the sub-cohort, reflecting the extent of gut fungal compositional alterations.

We fitted a multivariate regression model to examine the associations of pre-pregnancy overweight status (category variable) or pre-pregnancy BMI (continuous variable, z-score transformed) with the extent of gut fungal compositional alteration within the established sub-cohort. The model was adjusted for potential confounders including age, time interval between sample collection, parity, antibiotics use and pregnancy complications. To examine the potential influence of gut fungal compositional alteration on adverse birth outcomes, multivariate regression models were constructed for preterm delivery, low birthweight, and macrosomia. The aforementioned potential confounders, along with the extent of gut fungal compositional alteration, were considered as exposure variables in these models.

With regard to the divergence between individuals, the inter-individual distance was evaluated at various time points (i.e., T1, T2 or T3), separately. At each time point (e.g., T1 or T3), we calculate the average Bray-Curtis distance for each participant compared to all other participants. Thus, at each time point, each participant

possessed a distance value to reflect the similarity of her gut fungal composition with others.

Dynamic trajectory of the core gut fungal genera and their relationship with host health status

We conducted a longitudinal trajectory analysis for each core fungal genus in the established sub-cohort comprising 750 participants. For each core fungi, every participant had three measurements recorded at different time points, namely T1, T2, and T3. Therefore, the trajectory of a genus within an individual participant could be characterized by a vector consisting of three component values corresponding to these time points. Thereafter, we applied the *vegdist* function from the R package *vegan* to calculate the Canberra distance metric.[3] We performed PERMANOVA analysis using the function adonis2 in *vegan*, to assess the association of between-individual variation in the trajectory of each fungal genus with pre-pregnancy overweight status or adverse birth outcomes. The p value was determined through 999 permutations and an FDR-adjusted p value of less than 0.05 was considered indicative of statistical significance.

Network analysis among gut fungal enterotype, functional pathways and host serum metabolites

The network analysis was conducted among the participants who had available metagenome data during the first trimester of pregnancy. After excluding 35 participants who had antibiotics exposure within 2weeks before stool sample collection, this analysis included a total of 1001 women. We firstly performed Kruskal-Wallis test to identify pathways whose distribution varied across fungal enterotypes. Thereafter, we performed post-hoc pair-wise comparison to defined which enterotype was enrich with the identified pathways. We fitted multivariate regression models to examine the associations of identified pathways with host serum metabolites. Here, we adjusted for potential confounders including age, gestation week, parity and pre-pregnancy BMI. Both the fungi and pathway data were

standardized using z-score before the regression analysis. An FDR-adjusted p value of less than 0.05 was considered indicative of statistical significance.

Covarying relationship between gut fungi alterations and host metabolic changes

We investigated the relationship between alterations in gut fungi and changes in host metabolism among 709 participants. These participants had available gut mycobiome sequencing data as well as serum metabolome data during each trimester of pregnancy. To address the compositional nature of the mycobiome data, we first applied the centered log-ratio (CLR) method to transform the taxa data. Next, we calculated the changes in each core fungus from T1 to T3 for each participant. Additionally, we calculated the changes in signal intensity of each serum metabolite from T1 to T3 for each participant. This enabled us to construct a matrix of gut fungi alterations and a matrix of host metabolic changes. To investigate the overall relationship between gut fungi alterations and host metabolic changes, we conducted Procrustes analysis in R using the 'vegan' R package. Procrustes. The p-value was generated based on 999 permutations.[3]

Furthermore, we explored the covarying relationship between individual fungal genera and individual serum metabolites. For this analysis, we applied pairwise Spearman correlation analysis to each genus-serum metabolite pair in the aforementioned dataset. We considered a false discovery rate (FDR)-adjusted p-value of less than 0.05 as indicative of statistical significance.

Pre-pregnancy overweight status impacts the metabolic changes during pregnancy

For those participants with available serum metabolomics data throughout each trimester of pregnancy, we had constructed a matrix of host metabolic changes. Based on this matrix, the *vegdist* function from the R package *vegan* was utilized to calculate the Canberra distance metric.[3] To assess the contribution of pre-pregnancy overweight to the variation in metabolic changes between individuals, a

PERMANOVA analysis was conducted using the *adonis2* function from the vegan package.[3] The significance of the results was determined using 999 permutations, and a false discovery rate (FDR)-adjusted p-value of less than 0.05 was considered statistically significant.

To identify the distinct metabolic alterations in underweight and overweight pregnant women, we classified the participants into three groups based on their pre-pregnancy weight status: underweight, normal weight, and overweight/obese. Paired t-tests were conducted for each serum metabolite measured at T1 and T3 within each group. A FDR-adjusted p-value of less than 0.05 was considered statistically significant. Metabolic changes that were observed as significant solely among the underweight group, but not in the other groups, were defined as unique metabolic changes in pregnant women who were underweight prior to pregnancy. Similarly, significant metabolic changes observed solely among the overweight/obese group were defined as unique metabolic changes in pregnant women who were overweight/obese prior to pregnancy.

Clinical impact of the gut mycobiome during pregnancy

Logistic regression was used to evaluate the association between each core fungal genus, specifically measured during the first trimester of pregnancy, and the occurrence of pregnancy complications (GDM and PIH), as well as adverse birth outcomes (e.g., preterm delivery, macrosomia, and low birthweight). The model included age, pre-pregnancy BMI, parity as covariates. Specifically, when examining GDM or PIH as exposures, the model was adjusted for these variables. Additionally, when fungal genera were included as exposures, the fungal genera data were z-score standardized and the model was further adjusted for the gestational week corresponding to stool sample collection and the batch of sequencing.

The analysis pertaining to pregnancy complications was conducted among 4606 participants who possessed gut fungi sequencing data at T1 and information regarding

pregnancy complications. Likewise, the analysis concerning birth outcomes was performed among 4656 participants who had gut fungi sequencing data at T1 and information on birth outcomes. In both analyses, a complete data analysis strategy was implemented for variables with an exceptionally low occurrence of missing values, which consequently excluded 3 participants. For the variable "gestational week at stool sample collection," which was missing for 278 participants, a multiple imputation strategy was employed to impute these missing values. As a result, the logistic regression model for pregnancy complications consisted of 4603 participants, while the logistic regression model for adverse birth outcomes encompassed 4653 participants.

For pregnancy complications that exhibited significant associations with both gut fungi and adverse birth outcomes, we conducted mediation analysis to investigate the potential mediation effect of pregnancy complications on the link between the mycobiome and adverse birth outcomes. All statistical analyses were performed using Stata version 15 or R version 4.0.2.

References

1 Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3.

2 Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:852–7.

3. Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Sólymos, P., Stevens, M.H.H. & Wagner, H., vegan: Community Ecology Package, 2012, Software. http://CRAN.R-project.org/package=vegan.

4. Arumugam M, Raes J, Pelletier E, et al. Enterotypes of the human gut microbiome. Nature. 2011;473:174–80.

5. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu T.Y. (2017) LightGBM: a highly efficient gradient boosting decision tree. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems.

3149–3157.

6. Lundberg, S.M., Lee, S.I. (2017) A unified approach to interpreting model predictions. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. 4768–4777.

Table S1: Characteristics of demographic, clinical and dietary factors stratified by early-pregnancy fungal enterotypes

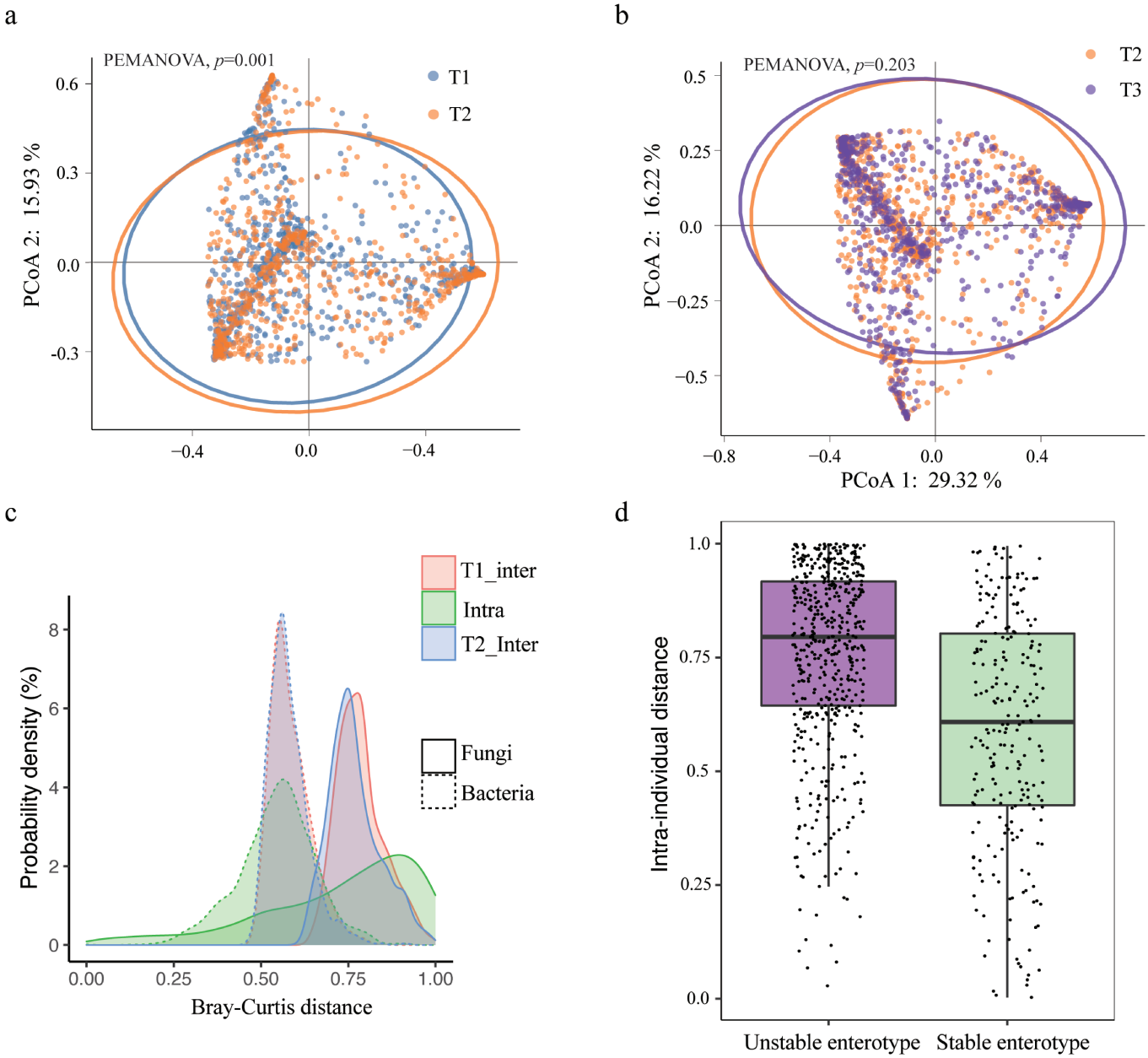
	Candida-dominated enterotype (n=907)	*Aspergillus*-dominated enterotype (n=2723)	*Saccharomyces*-dominated enterotype (n=1170)	*p**
Age (year)	26.4 (3.6)	26.3 (3.6)	26.5 (3.5)	0.246
Pre-pregnancy BMI	20.8 (2.9)	20.9 (2.9)	21.0 (3.1)	0.343
Gestation week (wk)	10.2 (2.1)	10.2 (2.0)	10.1 (2.0)	0.374
Time to delivery (day)	198.2 (16.1)	198.4 (15.5)	198.5 (15.6)	0.940
Parity				0.013
Primiparous	54.0%	57.3%	60.5%	
Overweight status				0.249
Overweight or obese	21.8	19.2	19.7	
Normal weight	65.1	67.6	65.3	
Underweight	13.1	13.3	15	
Steam bread consumption				<0.001
Ever (during the past 1 year)	46.9%	47.3%	54.2%	
Drinking				0.070
Ever (during the past 1 year)	22.4%	22.0%	18.9%	
Tea consumption				0.024
Ever (during the past 1 year)	34.0%	33.7%	29.5%	
Coffee consumption				0.057
Ever (during the past 1 year)	24.6%	26.6%	23.1%	
Milk consumption				0.564
never	19.2%	21.0%	19.7%	
<1 / day	42.9%	42.9%	42.1%	
⩾1/day	38.0%	36.1%	38.2%	

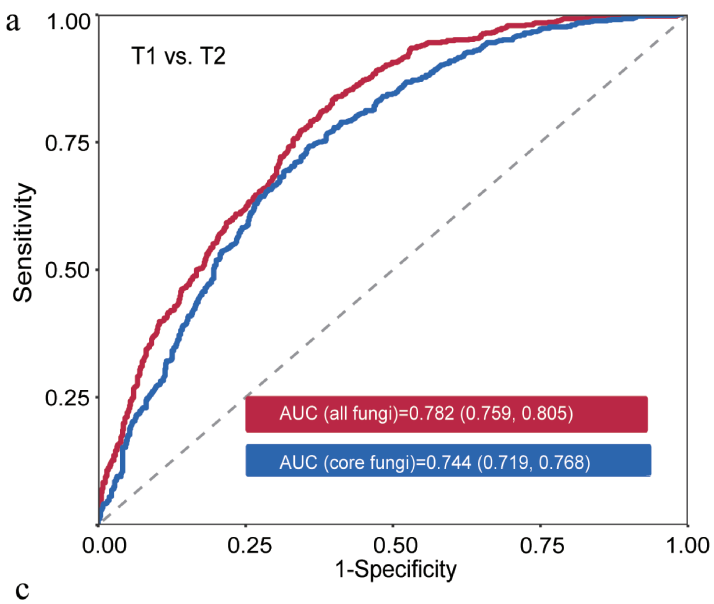
* One-way ANOVA was applied to examine the significance of difference between groups for continuous variables, while chi-square test was used for category variables.

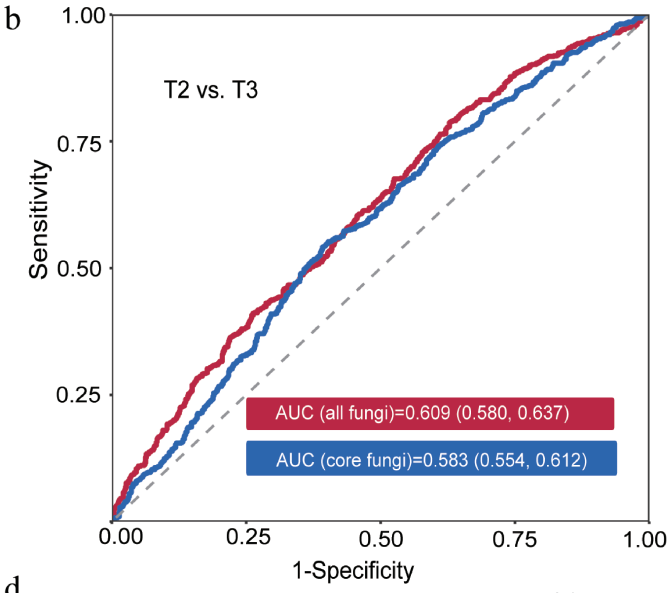
Table S2. The association between changes in consumption of food groups during pregnancy and alterations in gut fungal richness*

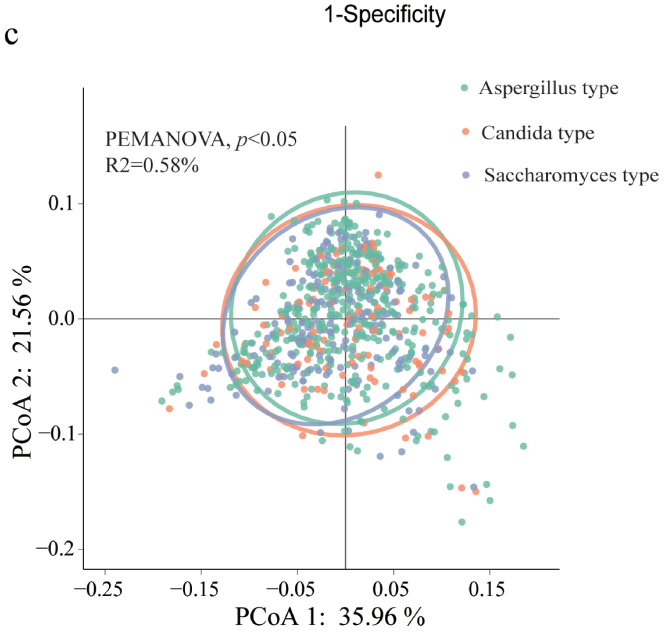
Exposure (changes in consumption of food groups)	Beta coefficient	95% CI	p	FDR
Rice (per 50g/day)	-3.20	(-5.75, -0.64)	0.014	0.11
Steamed bread (per 1 bread/day)	-6.38	(-12.9, 0.14)	0.055	0.15
Noodle (per 50g/day)	-5.06	(-10.63, 0.50)	0.074	0.15
Vegetables (per 50g/day)	1.20	(-9.82, 3.22)	0.243	0.39
Meat (per 50g/day)	1.45	(-3.48, 6.37)	0.564	0.64
Egg (per 1 egg/day)	7.33	(0.26, 14.41)	0.042	0.15
Milk (per 50g/day)	0.01	(-0.01, 0.03)	0.51	0.64
Fruit (per 50g/day)	-0.30	(-1.89, 1.29)	0.712	0.71

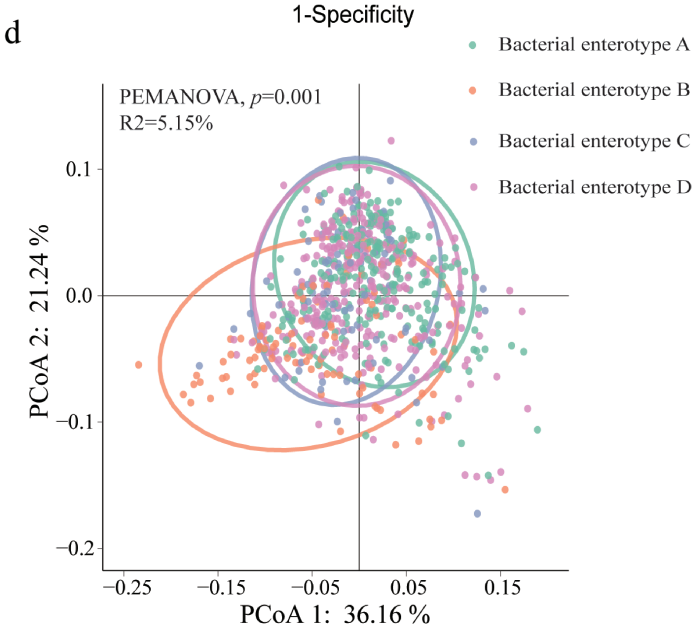
* Covariates included in the regression model: age, pre-pregnancy BMI, interval time between sample collections, parity, and gravidity.
639 out of the 750 with available data on the estimation of consumption changes in food groups from T1 to T3 based on the FFQ were included.


a
PEMANOVA, p=0.001
T1
T2
PCoA 2: 15.93 %
b
PEMANOVA, p=0.203
T2
T3
PCoA 2: 16.22 %
PCoA 1: 29.32 %
c
T1_inter
Intra
T2_Inter
Fungi
Bacteria
Probability density (%)
Bray-Curtis distance
d
Intra-individual distance
Unstable enterotype
Stable enterotype


a
T1 vs. T2
Sensitivity
1-Specificity
AUC (all fungi)=0.782 (0.759, 0.805)
AUC (core fungi)=0.744 (0.719, 0.768)


b
T2 vs. T3
Sensitivity
1-Specificity
AUC (all fungi)=0.609 (0.580, 0.637)
AUC (core fungi)=0.583 (0.554, 0.612)


c
Aspergillus type
Candida type
Saccharomyces type
PEMANOVA, p<0.05
R2=0.58%
PCoA 2: 21.56 %
PCoA 1: 35.96 %


d
Bacterial enterotype A
Bacterial enterotype B
Bacterial enterotype C
Bacterial enterotype D
PEMANOVA, p=0.001
R2=5.15%
PCoA 2: 21.24 %
PCoA 1: 36.16 %

Supplemental Materials

Landscape of the gut mycobiome dynamics during pregnancy and its relationship with host metabolism and pregnancy health

This file includes:

Methods

(1) Study population

(2) Sample collection

(3) Questionnaires and clinical data collection

(4) Bioinformatic analyses

(5) Statistical analysis

Supplemental Figure S1 to S2

(1) Supplemental Figure S1. Supplementary Table S1: Characteristics of demographic, clinical and dietary factors stratified by early-pregnancy fungal enterotypes.

(2) Supplemental Figure S2. The association between changes in consumption of food groups during pregnancy and alterations in gut fungal richness.

Fu Y, *et al. Gut* 2024;0:1–11. doi: 10.1136/gutjnl-2024-332260

Methods

Study population

The current study was based on the Tongji-Huaxi-Shuangliu Birth Cohort (THSBC), in which pregnant women aged 18 to 41 years were recruited during the early pregnancy when they presented to antenatal care clinics in in a local maternal and child health hospital during their early pregnancy. Exclusion criteria were 1) receiving infertility treatment (e.g., in vitro fertilization or intrauterine insemination); 2) reporting severe chronic or infectious diseases (e.g., cancer, HIV infection, or tuberculosis); or 3) were unable to or refused to sign the informed consent. The THSBC study was approved by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology (No. [2017](S225)-1), and informed consent was obtained from all participants.

In the present analysis, we included 4800 participants who had available ITS2 sequencing data, dietary information and clinical records during their pregnancy. This dataset enables us to comprehensively profile the gut mycobiome among pregnant women and investigate potential determinants contributing to the variations of gut mycobiome. To examine how pregnancy impacts the gut mycobiome over time and investigate their potential associations with host metabolism, we established a sub-cohort of 1059 participants, which included 514 women who gave birth to preterm (n=240), low birthweight (n=137), or macrosomia (n=216) infants, as well as 545 randomly selected participants who did not experience the above three adverse pregnancy outcomes.

ITS2 sequencing was performed for all the 4800 participants, while the shotgun metagenomics sequencing was performed for T1 samples within the established sub-cohort (n=1059). Additionally, within the sub-cohort, 750 and 748 participants had ITS2 and 16S sequencing data available, respectively, for all trimesters. We also repeatedly measured serum metabolome throughout each trimester of pregnancy for participants in this selected sub-cohort using an LC-ESI-MS/MS system.

Fu Y, *et al. Gut* 2024;0:1–11. doi: 10.1136/gutjnl-2024-332260

Sample collection

Stool samples were collected and stored in ice boxes at the hospital or home by the participants under instructions and then transferred to the hospital to store at −40 °C within 24 hours. A detailed standard operating procedure was given to the pregnant women for instructions on fecal sample collection, temporary storage, and transportation to the hospital. Stocks of frozen fecal samples were transported by dry ice every 2 to 3 months to the laboratory in Huazhong University of Science and Technology and stored at -80 °C before further processing.

Questionnaires and clinical data collection

All participants completed a set of structured questionnaires on sociodemographic information, lifestyle and behaviors (such as cigarette smoking and alcohol drinking), diet (including recent daily intakes of major food groups), history of pregnancy and births, history of diseases and medications, and family history of diseases. Anthropometric and blood pressure measurements were collected on site using devices according to standard protocols. Body mass index (BMI) was calculated by dividing the weight in kilograms by the square of height in meters. Underweight, overweight and obesity was defined as BMI<18.5, BMI ≥ 24 kg/m^2 and BMI ≥ 28 kg/m^2, respectively. Pregnancy complications (e.g., gestational diabetes mellitus) and birth outcomes were extracted from the electronic clinical records.

Bioinformatic analyses

Gut mycobiome analysis using ITS2 rRNA gene sequencing data

Microbial DNA was extracted using the E.Z.N.A.® soil DNA Kit (Omega Bio-tek, Norcross, GA, U.S.) according to manufacturer's protocols. The final DNA concentration and purification were determined by NanoDrop 2000 UV-vis spectrophotometer (Thermo Scientific, Wilmington, USA), and DNA quality was checked by 1% agarose gel electrophoresis. The ITS2 hypervariable regions of the fungal ITS rRNA gene were amplified with primers ITS3F: GCATCGATGAAGAACGCAGC and ITS4R: TCCTCCGCTTATTGATATGC by

thermocycler PCR system (GeneAmp 9700, ABI, USA). Purified amplicons were pooled in equimolar and paired-end sequenced (2 × 250) on an Illumina NovaSeq platform (Illumina, San Diego, USA) according to the standard protocols by Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China).

The mean sequencing depth and its standard deviation for all samples were 118,177 and 7955, respectively. It is worth noting that within the sub-cohort analyzed longitudinally, the number of sequencing reads remained relatively consistent across trimesters. Specifically, the mean (SD) sequencing depths were 117,906 (8350), 118,882 (6125), and 116,688 (9198) for samples collected during the first, second, and third trimester of pregnancy, respectively. The demultiplexed ITS2 sequences were denoised and grouped into amplicon sequence variants (ASVs; i.e., 100% exact sequence match) using DADA2.[1] During the process, marker gene Illumina sequence data and low-quality regions of the sequences were detected and filtered. We trimmed 28 bases (primer and barcode) from the beginning of the sequences. We also truncated the sequences at the 245 bases as the quality dropped around position 245 (median of quality score <30). The ASV features that were presented in only one sample were excluded as suggested by the Qiime2 tutorial, based on the suspicion that these may not represent real biological diversity but rather PCR or sequencing errors. The individual ASVs were taxonomically classified based on the UNITE (version 8.2, 99%) database using the VSEARCH tool wrapped in QIIME2 (version 2021.2).[2] α-diversity analysis was conducted through the q2-diversity plugin at the sampling depth of 10000. α-diversity was estimated by Shannon's diversity index (or Shannon; a quantitative measure of community richness and evenness), Observed Features (or Richness; a qualitative measure of community richness), and Faith's PD (or Faith's Phylogenetic Diversity; a qualitative measure of community richness that incorporates phylogenetic relationships between the observed features).

Gut bacteria analysis using 16S rRNA gene sequencing data

For the 16S analysis, raw sequencing reads were merge-paired, quality filtered and

analyzed using QIIME2 (version 2021.2). As described above, we used DADA2 denoised-paired plugin in QIIME2 to process the fastq files. We filtered the features that were present in only a single sample. The taxonomies of ASVs were subsequently determined using the Naive Bayes classifier trained on the Sliva_138 99% reference database. α-diversity analysis was conducted at the sampling depth of 10000. A diversity of the gut bacteria was estimated by the indices the same as ITS2 data.

Microbial functional profiling using metagenome data

Microbial DNA extractions were carried out by a standardized CTAB procedure. DNA concentration was measured using Qubit dsDNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, CA, USA). For DNA library preparation, a total amount of 1μg DNA per sample was used. In addition, the NEBNext Ultra DNA Library Prep Kit (NEB, USA) was used following manufacturer's recommendations and index codes were added to attribute sequences to each sample. The DNA samples were fragmented by sonication to a size of approximately 350 bp. Then, the DNA fragments were end-polished, A-tailed, and ligated with the full-length adaptor for Illumina sequencing with further PCR amplification. Thereafter, PCR products were purified (AMPure XP system) and libraries were analyzed for size distribution by Agilent2100 Bioanalyzer and quantified using real-time PCR. The clustering of the index-coded samples was performed on a cBot Cluster Generation System according to the manufacturer's instructions. Lastly, sequencing was performed using the Illumina NovaSeq platform at Shanghai Personal Biotechnology Co. Ltd. (Shanghai, China) and 150 bp paired-end reads were generated.

Next, raw sequencing reads were first quality-controlled with KneadData toolkit (v0.10.0): 1) to trim the reads by quality score from the 5′ end and 3′ end with a quality threshold of 20; 2) removed read pairs when either read was < 50 bp, contained "N" bases or quality score mean below 30; and 3) deduplicate the reads. Reads aligning to the human genome (H. sapiens, UCSC hg38) were removed via KneadData integrated with Bowtie2 tool (v2.4.5).

Functional profiling was performed with HUMAnN3 (v3.0.1), which maps sample reads against the sample-specific reference database to quantify gene presence and abundance in a species-stratified manner, with unmapped reads further used in a translated search against Uniref90 to include taxonomically unclassified but functionally distinct gene family abundances. We extracted the Uniref90 gene families of gut bacteria for downstream analyses. The Uniref90 gene families were then converted into relative abundances of unstratified pathway.

Serum metabolomics profiling

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/) at Wuhan Metware Biotechnology Co., Ltd. (Wuhan, China). LIT and triple quadrupole (QQQ) scans were acquired on a triple quadrupole-linear ion trap mass spectrometer (QTRAP), QTRAP® LC-MS/MS System, equipped with an ESI Turbo Ion-Spray interface, operating in positive and negative ion mode and controlled by Analyst 1.6.3 software (Sciex). Instrument tuning and mass calibration were performed with 10 and 100 μmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as MRM experiments with collision gas (nitrogen) set to 5 psi. Declustering potential (DP) and collision energy (CE) for individual MRM transitions was done with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period. The mass spectrum data were processed using Software Analyst 1.6.3. The metabolite identification was conducted by referencing standards in self-built metware database and public databases. The identified metabolites were matched with the parent ion mass-to-charge ratio, the fragment ion mass-to-charge ratio as well as retention time of their corresponding standards. The accuracy of metabolite characterization was classified into three levels, depending on the presence of isotope internal standards or matching score with the secondary mass spectrometry. The matching score>0.7 indicates the level 1 accuracy of metabolite characterization, while 0.5-0.7 and <0.5 indicate the level 2 and level 3 accuracy, respectively.

Statistical analysis

Determinants of gut fungal composition among pregnant women

The gut mycobiome data were analyzed at the genus level. We investigated the determinants of the gut fungal composition using the data collected from 4800 participants during the first trimester, with each participant contributing only 1 stool sample. We used the *vegdist* function in the R package *vegan* to calculate the gut fungal Bray-Curtis dissimilarity matrix. The contribution of 20 environmental variables (including demographics, physiologic traits, diseases, and habitual dietary intakes) to fungal community variation was determined by PERMANOVA analysis using the function adonis2 in *vegan*.[3] We applied a complete data analysis strategy which excluded 19 participants with missing values for at least one of the environmental variables. The 4800 samples included in this analysis were sequenced in two separate batches. Therefore, we included the batch information as a covariate in the model, to adjust for potential batch effects. The p value was determined through 999 permutations.

Gut fungal and bacterial enterotype clustering

The fecal samples (T1, n=4800; T2, n=890; T3, n=850) of ITS2 amplification were clustered into fungal enterotypes by using a partitioning around medoid (PAM) clustering method as those previously described.[4] Briefly, the samples were grouped into clusters with partitioning around medoid (PAM) based on the between-sample Bray–Curtis distance calculated at genus-level. The optimal number of clusters was determined by the silhouette index. The driver genus of each enterotype was determined as the genus with the highest relative abundance in the enterotype. The fecal samples of 16S amplification were clustered into bacterial enterotypes by using the method as that for bacterial enterotype.

Dynamics of within-sample α diversity throughout each trimester of pregnancy

This analysis was conducted in the established sub-cohort of 750 participants who had available gut fungi α diversity data for each trimester of pregnancy. We utilized

paired t-tests to assess the statistical significance of the changes in fungal α diversity between T1 and T2, as well as between T2 and T3 independently.

Among the 750 participants included in the study to profile changes in gut fungal richness from T1 to T3, data on the estimation of consumption changes in food groups from T1 to T3 based on the FFQ were available for 639 participants. These food groups consisted of rice, steamed bread, noodles, vegetables, meat, eggs, milk, and fruit. The weight of each food group consumed per day was quantified based on the FFQ. We applied a linear regression model to estimate the association between the changes in richness from T1 to T3 and the respective food group. As covariates, we incorporated age, pre-pregnancy BMI, interval time between sample collections, parity, and gravidity. An FDR<0.05 was considered statistically significant.

Loss rate calculations and discriminative genera identification

Utilizing the available repeated measurements of gut fungi throughout pregnancy within the designated sub-cohort (n=750), we elucidated the loss rate for each fungal genus as the host underwent progression from T1 to T3. The loss rate for each fungal genus was determined by quantifying the proportion of the decline in frequency observed between T1 and T3.

To assess the gut fungi enriched or depleted during early or late pregnancy, we conducted an analysis using paired t test analysis. This analysis was based on the examination of 465 genera that were detected at either T1 or T3. We transformed the taxa data using the centered log-ratio (CLR) method to address the compositional nature of the mycobiome data before we perform the paired t test analysis. To determine statistical significance, a false discovery rate (FDR)-adjusted p-value threshold of less than 0.05 was used.

Additionally, we used these 465 fungal genera to construct a machine learning framework of LightGBM for predicting the trimester that the samples belong to.[5] The construction of prediction model was based on Scikit-learn (v0.15.2), and ten-

fold cross validation (CV) was applied. To evaluate binary classification performance, receiver operating characteristic (ROC) curve analyses were conducted using the R package pROC. We used the SHAP (Shapley Additive exPlanations) to interpret predictions and the importance of each fungal genus to the prediction model is represented using Shapley values.[6]

Quantification of intra-individual gut fungi compositional alterations

We applied the *vegdist* function from the R package *vegan* to calculate the Bray-Curtis distance based on gut fungal genus-level composition to assess intra-individual and inter-individual dissimilarities, respectively.[3] The sub-cohort, consisting of 750 participants, was included in this analysis, as they had relevant data throughout each trimester of pregnancy. To determine the intra-individual distance, paired data was used, with the gut mycobiome data at T1 serving as the reference. Subsequently, a Bray-Curtis distance value was calculated for each participant in the sub-cohort, reflecting the extent of gut fungal compositional alterations.

We fitted a multivariate regression model to examine the associations of pre-pregnancy overweight status (category variable) or pre-pregnancy BMI (continuous variable, z-score transformed) with the extent of gut fungal compositional alteration within the established sub-cohort. The model was adjusted for potential confounders including age, time interval between sample collection, parity, antibiotics use and pregnancy complications. To examine the potential influence of gut fungal compositional alteration on adverse birth outcomes, multivariate regression models were constructed for preterm delivery, low birthweight, and macrosomia. The aforementioned potential confounders, along with the extent of gut fungal compositional alteration, were considered as exposure variables in these models.

With regard to the divergence between individuals, the inter-individual distance was evaluated at various time points (i.e., T1, T2 or T3), separately. At each time point (e.g., T1 or T3), we calculate the average Bray-Curtis distance for each participant compared to all other participants. Thus, at each time point, each participant

possessed a distance value to reflect the similarity of her gut fungal composition with others.

Dynamic trajectory of the core gut fungal genera and their relationship with host health status

We conducted a longitudinal trajectory analysis for each core fungal genus in the established sub-cohort comprising 750 participants. For each core fungi, every participant had three measurements recorded at different time points, namely T1, T2, and T3. Therefore, the trajectory of a genus within an individual participant could be characterized by a vector consisting of three component values corresponding to these time points. Thereafter, we applied the *vegdist* function from the R package *vegan* to calculate the Canberra distance metric.[3] We performed PERMANOVA analysis using the function adonis2 in *vegan*, to assess the association of between-individual variation in the trajectory of each fungal genus with pre-pregnancy overweight status or adverse birth outcomes. The p value was determined through 999 permutations and an FDR-adjusted p value of less than 0.05 was considered indicative of statistical significance.

Network analysis among gut fungal enterotype, functional pathways and host serum metabolites

The network analysis was conducted among the participants who had available metagenome data during the first trimester of pregnancy. After excluding 35 participants who had antibiotics exposure within 2weeks before stool sample collection, this analysis included a total of 1001 women. We firstly performed Kruskal-Wallis test to identify pathways whose distribution varied across fungal enterotypes. Thereafter, we performed post-hoc pair-wise comparison to defined which enterotype was enrich with the identified pathways. We fitted multivariate regression models to examine the associations of identified pathways with host serum metabolites. Here, we adjusted for potential confounders including age, gestation week, parity and pre-pregnancy BMI. Both the fungi and pathway data were

standardized using z-score before the regression analysis. An FDR-adjusted p value of less than 0.05 was considered indicative of statistical significance.

Covarying relationship between gut fungi alterations and host metabolic changes

We investigated the relationship between alterations in gut fungi and changes in host metabolism among 709 participants. These participants had available gut mycobiome sequencing data as well as serum metabolome data during each trimester of pregnancy. To address the compositional nature of the mycobiome data, we first applied the centered log-ratio (CLR) method to transform the taxa data. Next, we calculated the changes in each core fungus from T1 to T3 for each participant. Additionally, we calculated the changes in signal intensity of each serum metabolite from T1 to T3 for each participant. This enabled us to construct a matrix of gut fungi alterations and a matrix of host metabolic changes. To investigate the overall relationship between gut fungi alterations and host metabolic changes, we conducted Procrustes analysis in R using the 'vegan' R package. Procrustes. The p-value was generated based on 999 permutations.[3]

Furthermore, we explored the covarying relationship between individual fungal genera and individual serum metabolites. For this analysis, we applied pairwise Spearman correlation analysis to each genus-serum metabolite pair in the aforementioned dataset. We considered a false discovery rate (FDR)-adjusted p-value of less than 0.05 as indicative of statistical significance.

Pre-pregnancy overweight status impacts the metabolic changes during pregnancy

For those participants with available serum metabolomics data throughout each trimester of pregnancy, we had constructed a matrix of host metabolic changes. Based on this matrix, the *vegdist* function from the R package *vegan* was utilized to calculate the Canberra distance metric.[3] To assess the contribution of pre-pregnancy overweight to the variation in metabolic changes between individuals, a

PERMANOVA analysis was conducted using the *adonis2* function from the vegan package.[3] The significance of the results was determined using 999 permutations, and a false discovery rate (FDR)-adjusted p-value of less than 0.05 was considered statistically significant.

To identify the distinct metabolic alterations in underweight and overweight pregnant women, we classified the participants into three groups based on their pre-pregnancy weight status: underweight, normal weight, and overweight/obese. Paired t-tests were conducted for each serum metabolite measured at T1 and T3 within each group. A FDR-adjusted p-value of less than 0.05 was considered statistically significant. Metabolic changes that were observed as significant solely among the underweight group, but not in the other groups, were defined as unique metabolic changes in pregnant women who were underweight prior to pregnancy. Similarly, significant metabolic changes observed solely among the overweight/obese group were defined as unique metabolic changes in pregnant women who were overweight/obese prior to pregnancy.

Clinical impact of the gut mycobiome during pregnancy

Logistic regression was used to evaluate the association between each core fungal genus, specifically measured during the first trimester of pregnancy, and the occurrence of pregnancy complications (GDM and PIH), as well as adverse birth outcomes (e.g., preterm delivery, macrosomia, and low birthweight). The model included age, pre-pregnancy BMI, parity as covariates. Specifically, when examining GDM or PIH as exposures, the model was adjusted for these variables. Additionally, when fungal genera were included as exposures, the fungal genera data were z-score standardized and the model was further adjusted for the gestational week corresponding to stool sample collection and the batch of sequencing.

The analysis pertaining to pregnancy complications was conducted among 4606 participants who possessed gut fungi sequencing data at T1 and information regarding

pregnancy complications. Likewise, the analysis concerning birth outcomes was performed among 4656 participants who had gut fungi sequencing data at T1 and information on birth outcomes. In both analyses, a complete data analysis strategy was implemented for variables with an exceptionally low occurrence of missing values, which consequently excluded 3 participants. For the variable "gestational week at stool sample collection," which was missing for 278 participants, a multiple imputation strategy was employed to impute these missing values. As a result, the logistic regression model for pregnancy complications consisted of 4603 participants, while the logistic regression model for adverse birth outcomes encompassed 4653 participants.

For pregnancy complications that exhibited significant associations with both gut fungi and adverse birth outcomes, we conducted mediation analysis to investigate the potential mediation effect of pregnancy complications on the link between the mycobiome and adverse birth outcomes. All statistical analyses were performed using Stata version 15 or R version 4.0.2.

References

1 Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3.

2 Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:852–7.

3. Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Sólymos, P., Stevens, M.H.H. & Wagner, H., vegan: Community Ecology Package, 2012, Software. http://CRAN.R-project.org/package=vegan.

4. Arumugam M, Raes J, Pelletier E, et al. Enterotypes of the human gut microbiome. Nature. 2011;473:174–80.

5. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu T.Y. (2017) LightGBM: a highly efficient gradient boosting decision tree. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems.

3149–3157.

6. Lundberg, S.M., Lee, S.I. (2017) A unified approach to interpreting model predictions. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. 4768–4777.

Table S1: Characteristics of demographic, clinical and dietary factors stratified by early-pregnancy fungal enterotypes

	Candida-dominated enterotype (n=907)	*Aspergillus*-dominated enterotype (n=2723)	*Saccharomyces*-dominated enterotype (n=1170)	*p**
Age (year)	26.4 (3.6)	26.3 (3.6)	26.5 (3.5)	0.246
Pre-pregnancy BMI	20.8 (2.9)	20.9 (2.9)	21.0 (3.1)	0.343
Gestation week (wk)	10.2 (2.1)	10.2 (2.0)	10.1 (2.0)	0.374
Time to delivery (day)	198.2 (16.1)	198.4 (15.5)	198.5 (15.6)	0.940
Parity				0.013
Primiparous	54.0%	57.3%	60.5%	
Overweight status				0.249
Overweight or obese	21.8	19.2	19.7	
Normal weight	65.1	67.6	65.3	
Underweight	13.1	13.3	15	
Steam bread consumption				<0.001
Ever (during the past 1 year)	46.9%	47.3%	54.2%	
Drinking				0.070
Ever (during the past 1 year)	22.4%	22.0%	18.9%	
Tea consumption				0.024
Ever (during the past 1 year)	34.0%	33.7%	29.5%	
Coffee consumption				0.057
Ever (during the past 1 year)	24.6%	26.6%	23.1%	
Milk consumption				0.564
never	19.2%	21.0%	19.7%	
<1 / day	42.9%	42.9%	42.1%	
⩾1/day	38.0%	36.1%	38.2%	

* One-way ANOVA was applied to examine the significance of difference between groups for continuous variables, while chi-square test was used for category variables.

Table S2. The association between changes in consumption of food groups during pregnancy and alterations in gut fungal richness*

Exposure (changes in consumption of food groups)	Beta coefficient	95% CI	p	FDR
Rice (per 50g/day)	-3.20	(-5.75, -0.64)	0.014	0.11
Steamed bread (per 1 bread/day)	-6.38	(-12.9, 0.14)	0.055	0.15
Noodle (per 50g/day)	-5.06	(-10.63, 0.50)	0.074	0.15
Vegetables (per 50g/day)	1.20	(-9.82, 3.22)	0.243	0.39
Meat (per 50g/day)	1.45	(-3.48, 6.37)	0.564	0.64
Egg (per 1 egg/day)	7.33	(0.26, 14.41)	0.042	0.15
Milk (per 50g/day)	0.01	(-0.01, 0.03)	0.51	0.64
Fruit (per 50g/day)	-0.30	(-1.89, 1.29)	0.712	0.71

* Covariates included in the regression model: age, pre-pregnancy BMI, interval time between sample collections, parity, and gravidity.
639 out of the 750 with available data on the estimation of consumption changes in food groups from T1 to T3 based on the FFQ were included.