

Discovery of metabolic biomarkers for gestational diabetes mellitus in a Chinese population

Wenqian Lu

Southern Medical University

Mingjuan Luo

Southern Medical University

Xiangnan Fang

Southern Medical University

Rong Zhang

Shanghai Jiao Tong University Affiliated Sixth People's Hospital

Mengyang Tang

Southern Medical University

Xiangtian Yu

Shanghai Jiao Tong University Affiliated Sixth People's Hospital

Cheng Hu (✉ alfredhc@sjtu.edu.cn)

Shanghai Diabetes Institute, Shanghai Key Laboratory of Diabetic Mellitus, Shanghai Jiao Tong University Affiliated Sixth People's Hospital <https://orcid.org/0000-0003-4314-2386>

Research Article

Keywords: Gestational diabetes mellitus, Metabolic biomarkers, Pregnancy, Pathogenesis

Posted Date: April 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-109665/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Nutrition & Metabolism on August 21st, 2021. See the published version at <https://doi.org/10.1186/s12986-021-00606-8>.

Abstract

Background: Gestational diabetes mellitus (GDM), one of the most common pregnancy complications, can lead to morbidity and mortality in both the mother and the infant. Metabolomics has provided new insights into the pathology of GDM and systemic analysis of GDM with metabolites is required for providing more clues for GDM diagnosis and mechanism research. This study aims to reveal metabolic differences between normal pregnant women and GDM patients in the second- and third-trimester stages and to confirm the clinical relevance of these new findings.

Methods: Metabolites were quantitated with the serum samples of 200 healthy pregnant women and 200 GDM women in the second trimester, 199 normal controls, and 199 GDM patients in the third trimester. Both function and pathway analyses were applied to explore biological roles involved in the two sets of metabolites. Then the trimester stage-specific GDM metabolite biomarkers were identified by combining machine learning approaches, and the logistic regression models were constructed to evaluate predictive efficiency. Finally, the weighted gene co-expression network analysis method was used to further capture the associations between metabolite modules with biomarkers and clinical indices.

Results: This study revealed that 57 differentially expressed metabolites (DEMs) were discovered in the second-trimester group, among which the most significant one was 3-methyl-2-oxovaleric acid. Similarly, 72 DEMs were found in the third-trimester group, and the most significant metabolites were ketoleucine and alpha-ketoisovaleric acid. These DEMs were mainly involved in the metabolism pathway of amino acids, fatty acids and bile acids. The logistic regression models for selected metabolite biomarkers achieved the area under the curve values of 0.807 and 0.81 for the second- and third-trimester groups. Furthermore, significant associations were found between DEMs/biomarkers and GDM-related indices.

Conclusions: Metabolic differences between healthy pregnant women and GDM patients were found. Associations between biomarkers and clinical indices were also investigated, which may provide insights into pathology of GDM.

Introduction

Gestational diabetes mellitus (GDM), defined as diabetes diagnosed during pregnancy, affects approximately 15% of pregnant women globally [1]. Several risk factors are highly correlated with the development of GDM, including maternal obesity, advanced age, family history of diabetes mellitus, and history of abnormal glucose metabolism [2, 3]. Studies have found that GDM can lead to several short-term and long-term complications, including obesity, impaired glucose metabolism, and cardiovascular diseases, for both mothers and infants [4]. It was reported that in women with previous GDM the risk of developing diabetes after delivery was more than seven times that in women with normal glucose tolerance (NGT). For the offspring, they are more than two to eight times as likely to develop obesity, the metabolic syndrome, type 2 diabetes, and impaired insulin sensitivity and secretion [5]. Although the pathogenesis of GDM is still not well understood, metabolomics has introduced new insights into the pathology of GDM and has yielded potential biomarkers related to GDM [6].

Metabolomics is the comprehensive analysis of low molecular weight compounds, known as metabolites, in biological systems. As the end products of metabolic processes, metabolites can reflect the internal physiological status of the organism that changes in response to environmental factors. Metabolic profiling is a useful

approach for qualitative and quantitative metabolite studies of cell, bio-fluids, and tissue [7]. In recent decades, metabolomics has been applied in various aspects, including disease diagnosis and treatment, metabolic pathway elucidation, biomarker discovery, and drug safety evaluation [8, 9]. For example, an untargeted metabolomics study has revealed a large number of pregnancy-related metabolic pathways and metabolites, five of which were able to predict gestational age in high accordance with ultrasound [10].

Many studies have been conducted for identifying biomarkers of GDM and for exploring the underlying mechanisms of its development. Akturk et al. observed that asymmetric dimethylarginine was elevated in women with GDM during late pregnancy [11]. Cetin et al. found that in GDM pregnancies, valine, methionine, phenylalanine, isoleucine, leucine, ornithine, glutamate, proline, and alanine were increased while glutamine was significantly decreased [12]. Chen et al. revealed that 2-aminobutyric acid was associated with an increased likelihood of GDM in China [13]. Moreover, 26 serum metabolites investigated by Liu et al. contributed to GDM, including 1-methyladenosine, homovanillic acid sulfate, and glucosamine compared with healthy pregnant women. These identified biomarkers are involved in some metabolic pathways that mainly participate in lipid, carbohydrate, and amino acid metabolisms [14]. Although these investigations of metabolite profiles have identified branched-chain amino acids (BCAAs), aromatic amino acids, sulfur-containing amino acids, and other metabolites, their findings still lack consistency. Therefore, systemic analysis of the metabolites of GDM is necessary for obtaining more clues for GDM diagnosis and mechanism research.

In this study, we aimed to explore the metabolic difference between GDM women and normal pregnant women in the second and third trimester, using ultra-performance liquid chromatography coupled to tandem mass spectrometry (UPLC-MS/MS) system. Additionally, we expected to reveal the association of clinical indices with differentially expressed metabolites (DEMs) found in such metabolomics analysis.

Materials And Methods

Study participant recruitment

All the samples were obtained from 200 healthy pregnant women and 200 GDM women in the second trimester, 199 normal controls, and 199 GDM patients in the third trimester at the University of Hong Kong-Shenzhen Hospital from 2016 to 2018. GDM patients were matched 1:1 with normal pregnant women. Matching was based on maternal age (± 3 years), pregestational BMI (± 3 kg/m²), and gestational week (± 3 weeks). Participants with cancer, kidney disease, heart disease, hepatic disease, other metabolic diseases or patients using medications that might affect glucolipid metabolism were excluded. The diagnostic criteria for GDM women were based on the standards recommended by the International Association of the Diabetes and Pregnancy Study Group. Glucose (75g) was used for 2 h for conducting oral glucose tolerance test (OGTT). GDM was defined when fasting plasma glucose (FPG) was ≥ 5.1 mmol/L or 1-h plasma glucose was ≥ 10.0 mmol/L or 2-h plasma glucose was ≥ 8.5 mmol/L. All values for the OGTT less than the thresholds were considered normal. The study was approved by the ethics committee of the University of Hong Kong-Shenzhen Hospital. It was conducted according to relevant regulations, and informed consent was signed by every participant.

Sample collection and serum metabolomics

Age, height, and weight were recorded for every participant. General background information, including family history of diabetes, reproductive history, and medical history, were collected. Information about pregnancy

outcomes was acquired after delivery, from the maternal and infant medical records in the hospital. Body mass index (BMI) was calculated as body weight (in kg)/height squared (in m²). Serum samples obtained from 24 to the end of 27 gestational weeks were defined as the second-trimester group, while those after 28 gestational weeks were defined as the third-trimester group. Blood samples were drawn in the morning after an overnight fast through the antecubital vein. Glucose levels were measured using the hexokinase method on a Roche Cobas 701 analyzer (Roche, Ltd, Basel, Switzerland). HbA1c values were tested using an Arkray HA-8160 analyzer (Arkray, Ltd, Kyoto, Japan). Serum lipids, including total triglyceride (TG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) levels were measured with a Siemens ADVIA2400 fully automated chemistry analyzer (Siemens AG, Munich, Germany).

Metabolomic profiling was performed using Metabo-Profile (Shanghai, China). Serum samples were stored at -80 °C until analysis. Approximately 120 µL of ice-cold methanol with partial internal standards was added to 25 µL of samples for extracting the metabolites. Internal standards include several metabolite standards, such as arginine, hippuric acid, carnitine, and citric acid. After vortexing for 5 mins, the mixture was centrifuged for 30 mins at 4000g; 30 µL of supernatant was transferred to a 96-well plate, and 20 µL of freshly prepared derivative reagents (3-Nitrophenylhydrazine) was added to each well. After derivatization for 60 mins at 30 °C, 330 µL of ice-cold 50% methanol solution was added for dilution. The samples were stored at -20 °C for 20 mins and were centrifuged at 4000g for 30 mins at 4 °C. Approximately 135 µL of supernatant was mixed with 10 µL of internal standards in each well of a new 96-well plate. Serial dilutions of derivatized stock standards were added to the left wells, and the plate was ready for analysis.

Chromatographic separation was performed on an ACQUITY UPLC BEH C18 VanGuard pre-column (2.1×5 mm, 1.7 µm) and an ACQUITY UPLC BEH C18 analytical column (2.1×100 mm, 1.7 µm). The temperature of the column and the sample manager was set at 40 °C and 10 °C, respectively. The mobile phase A was water with 0.1% formic acid, while B was a mixture of acetonitrile and isopropanol (70:30). Gradient conditions were 0–1 min, 5% B; 1–11 mins, 5%–78% B; 11–13.5 mins, 78%–95% B; 13.5–14 mins, 95%–100% B; 14–16 mins, 100% B; 16–16.1 mins, 100%–5% B; 16.1–18 mins, 5% B. The flow rate was 0.4 mL/min with a 5 µL injection volume.

The mass spectrometer was operated in positive electrospray ionization (ESI+) mode with a capillary voltage of 1.5 kV as well as the negative electrospray ionization (ESI-) mode with a capillary voltage of 2 kV. The temperature of the ion source and desolvation was 150 °C and 550 °C, respectively. The desolvation gas flow was set at 1000 L/h. Raw data files generated by UPLC-MS/MS were processed using Masslynx software (v4.1, Waters, Milford, MA, USA) for performing peak integration, calibration, and quantitation for each metabolite. An overview of the research design and analysis workflow is presented in Figure 1.

Bioinformatics analysis and metabolic biomarker recognition

For the analysis of clinical characteristics, participants were divided into four groups: the second-trimester GDM group (group 1), second-trimester normal group (group 2), third-trimester GDM group (group 3), and third-trimester normal group (group 4). Clinical characteristics between the same gestational-stage groups (group1 vs. group 2 and group3 vs. group 4) were compared using the Student's paired *t*-test or signed-rank test for continuous variables and chi-square tests for categorical variables with SAS 9.3 (SAS Institute, Cary, NC, USA). After normality testing for continuous variables, the data was presented as mean ± standard deviation (SD) or median (interquartile range). The level of significance was set at *p* < 0.05.

To observe changes in metabolites between the groups, the univariate and multivariate analyses were conducted, including differential expression analysis using *t*-test, fold-change with volcano plot, principal component analysis (PCA), partial least square discriminant analysis (PLS-DA), and orthogonal partial least square discriminant analysis (OPLS-DA). Data processing was performed using the iMAP platform (v1.0, Metabo-Profile, Shanghai, China). $p < 0.05$ and $|\log_2\text{FC}| > 0$ in the univariate analysis and variable importance in the projection (VIP) > 1 in the multivariate analysis were the standards for screening DEMs. To investigate the biological meaning of DEMs, enrichment analysis was performed with the small molecule pathway database (SMPDB) with $p < 0.05$ regarded as the significance level, and pathway analysis was also performed using p values to account for significance and false discovery. A threshold of $p < 0.05$ was used for considering the function or pathway to be impactful. Metabolites were further analyzed using random forest (RF) for identifying potential biomarkers based on which logistic regression (LR) models were constructed for evaluating predictive efficiency. Area under the curve (AUC) for receiver operating characteristic (ROC) curves were calculated for assessing the performance of potential biomarkers with LR for GDM. The determination of sensitivity and specificity at the optimal cut-off point was defined by the minimum distance to the top-left corner. The weighted gene co-expression network analysis (WGCNA) [15] was used for inferring the association between metabolite modules and clinical indices.

Results

Clinical characteristic of study participants

The clinical characteristics of all participants are presented in **Table 1**. Compared with normal pregnant women, pregnant women who developed GDM had higher FPG, 1-h, and 2-h blood glucose levels after an OGTT and smaller changes in BMI during pregnancy, in both the second and third trimesters. HbA1c, TG, TC, and LDL were statistically different between the case and control groups in the second trimester, which were consistent with previous research conclusions. TC, delivery gestational week, neonatal weight, and neonatal length were statistically different between GDM patients and healthy pregnant women in the third trimester. Women with GDM were more likely to have a family history of diabetes. Age, BMI, and gestational week were similar between the two groups in each trimester.

Metabolomics profiling of study participants

As shown in the metabolomics profiles (**Figures 2a, b**), the case group (GDM) did not separate from the control group (normal) in the second and third trimester with PCA; moreover, PLS-DA showed a relatively clear discrimination for both the trimesters (**Figures 2c, d**). Finally, the results of OPLS-DA (**Figures 2e, f**) indicated the possibility of evaluating the differences between GDM patients and normal controls with metabolite abundance. The results showed that R^2 and Q^2 of the OPLS-DA model in the second trimester were 0.347 and 0.165, while in the third trimester were 0.324 and 0.201, respectively. Permutation tests ($n = 200$) were employed for validating the predictive ability of the built OPLS-DA models (**Figure S1**). The R^2 and Q^2 values derived from the permuted data were lower than the original values, which demonstrated that the OPLS-DA model did not overfit. The VIP values of the OPLS-DA model and criteria, including $|\log_2\text{FC}| > 0$ with $p < 0.05$, were further employed for determining the DEMs. A volcano plot provided a quick way to display the differences in metabolite expression levels between normal pregnant women and GDM patients with statistical significance (**Figure 3**). As shown in the volcano plot, of 200 metabolites found in this study, 57 metabolites in the second-trimester group (**Table S1**) and 72 metabolites in the third-trimester group (**Table S2**) were considered DEMs. The top five DEMs in the

second-trimester group were 3-methyl-2-oxovaleric acid, 3-hydroxybutyric acid, palmitic acid, alpha-hydroxyisobutyric acid, and acetic acid. As for the third-trimester group, they were ketoleucine, alpha-ketoisovaleric acid, pyruvic acid, L-tryptophan, and succinic acid. Most DEMs were down-regulated during the second trimester (**Figure 3a**), while many DEMs were positively-regulated during the third trimester (**Figure 3c**), suggesting different metabolomic dysfunction at different trimester stages.

Metabolic enrichment of biological function and pathway relevant to GDM

The DEMs for each comparison group were evaluated using enrichment analysis with SMPDB (**Figures 4a, b**). In the second-trimester group, the alpha linolenic acid and linoleic acid metabolism pathways had the highest fold enrichment, lowest p value ($p < 0.001$), and FDR of < 0.1 . Other significant functions included beta oxidation of very long chain fatty acids and valine-leucine-isoleucine degradation (**Figure 4a and Table 2, Table S3**). In the third-trimester group, functions such as urea cycle, ammonia recycling, glycine and serine metabolism, valine-leucine-isoleucine degradation, arginine and proline metabolism, alanine metabolism, glutamate metabolism, aspartate metabolism, glucose-alanine cycle, phenylalanine and tyrosine metabolism, and carnitine synthesis were significantly associated with the corresponding DEMs (**Figure 4b and Table 3, Table S4**).

Pathway analysis was also performed for investigating the function of DEMs. 32 pathways were observed, nine of which were significantly enriched in the second-trimester group (**Figure 4c, Table S5**), among which two pathways, valine-leucine-isoleucine biosynthesis and valine-leucine-isoleucine degradation, played key roles in reflecting the changes in metabolites. For the third-trimester group, 48 pathways were found, of which 21 were significantly enriched with DEMs (**Figure 4d, Table S6**).

Furthermore, remarkable differences also exist between different trimester stages on function and pathway levels (**Figures 4 and S2**), suggesting that stage-specific biomarkers and diagnostic models should be considered.

Selection of potential metabolic biomarkers for GDM

After observing metabolomics differences between the groups and reliable functional enrichment analysis, it was necessary to establish a diagnostic model for predicting the presence of GDM in pregnant women and for selecting potential metabolite characteristics with the importance determined using machine learning algorithms [16, 17].

For the second-trimester group, samples were first divided into training data with 70% samples and test data with 30% samples. The RF model was then learned on the training data for obtaining the importance score for each metabolite, based on which the candidate metabolite biomarkers were selected with top large importance scores and with overlaps to DEMs as much as possible. Next, the LR model was constructed on the training data with such metabolite biomarkers (**Table S7**), with a training AUC of 0.808. Finally, the second-trimester group-specific LR model was validated on the test data and achieved a testing AUC of 0.807 (**Figure 5a**). Similarly, for the third-trimester group, the metabolite biomarkers (**Table S7**) were selected using RF, and an LR model was built, which had learning performance as AUC 0.819 and validation performance as AUC 0.810 on the test data (**Figure 5b**). In addition, almost all identified biomarker candidates were DEMs (**Table 4**).

Clinical relevance of metabolic biomarkers for GDM

Finally, to confirm the clinical relevance of metabolic biomarkers associated with GDM, WGCNA was performed for inferring the association between metabolite modules and clinical indices. As shown in **Figure 6a**, four modules were detected for the second-trimester group. Module turquoise was significantly associated with GDM (i.e., group index) and many other important clinical indices, including pre-pregnancy BMI, OGTT, TC, TG, and LDL. Additionally, this module significantly included 11 DEMs ($P = 9.99e^{-04}$), compared to module blue containing three DEMs ($P = 0.666$), module brown containing two DEMs ($P = 0.580$), and module grey containing 41 DEMs ($P = 0.868$) (**Figure 6c**). Similarly, five modules were found for the third-trimester group (**Figure 6b**), where the module turquoise contains 21 DEMs ($P = 6.737e^{-08}$), module blue contains five DEMs ($P = 0.944$), module brown contains six DEMs ($P = 0.090$), module grey contains 37 DEMs ($P = 0.992$), and module yellow contains three DEMs ($P = 0.606$). At this time, module turquoise was associated with group, OGTT, TC, TG, and LDL; and module yellow was positively associated with the group, pre-pregnancy BMI, and OGTT. It should be noted that the biomarker candidates are almost DEMs (**Table 4**); thus, they have similar contribution to particular modules and their associations with GDM.

Discussion

Using UPLC-MS/MS for metabolomics analysis, detailed information was obtained on the metabolic changes in normal pregnant women and GDM patients in our study. The changes in serum metabolites were further investigated using univariate and multivariate statistical analyses. Thirteen candidate metabolite biomarkers found in the second-trimester group and thirteen found in the third-trimester group contributed to GDM when compared with healthy pregnant women. According to metabolic enrichment and pathway analyses, valine-leucine-isoleucine degradation in the second-trimester group was consistently found in both analyses. In the third trimester, valine-leucine-isoleucine degradation and glycine, serine, arginine, proline, alanine, glutamate, aspartate, and phenylalanine metabolisms were consistent in both the analyses. Metabolic biomarkers have been found by RF, and LR models based on which showed high predictive efficiency. Furthermore, these biomarkers demonstrate remarkable relationship with clinical indices.

BCAAs, which consist of valine, leucine, and isoleucine, were higher in the GDM group than in the NGT group in our study. Although some studies have found that the levels of BCAAs did not differ significantly between GDM and NGT [18, 19], several studies have shown that elevated BCAAs in GDM patients might serve as biomarkers for GDM [20-22]. BCAAs may contribute to GDM through activating a mammalian target of rapamycin complex 1 (mTORC1) to cause insulin resistance [23, 24]. For example, leucine promotes the translocation of inactive mTORC1 to lysosomal compartments and interacts with the AMP kinase pathway, which leads to the downstream activation of mTORC1 by interacting with the TSC1/TSC2 complex [24].

In our study, glycine and serine metabolisms were lower in the GDM group, which is consistent with findings of previous studies. Takashina et al. observed that fasting glucose and 2-h plasma glucose levels or the homeostasis model assessment of insulin resistance negatively correlated with glycine levels, and the homeostasis model assessment for the β -cell function index negatively correlated with glycine and serine levels [25]. Moreover, oral glycine has been reported to increase insulin secretion without affecting insulin sensitivity [26].

Additionally, arginine and proline levels were also lower in the GDM group. It has been reported that arginine and its metabolites promote insulin secretion [27] and improve insulin resistance in humans [28]. Arginine plays

multiple beneficial roles against metabolic abnormalities, but it might also induce oxidative stress [29]. Proline is absorbed and metabolized into glutamine, which may enter the tricarboxylic cycle and ultimately be converted into glucose. One study demonstrated that the ingestion of proline with glucose attenuated the glucose area response without affecting insulin response and decreased glucagon levels compared to glucose alone [30].

Alpha linolenic acid is a precursor of polyunsaturated fatty acids, which mainly contain omega-6 and omega-3 fatty acids. A meta-analysis by Zhong et al. investigated the efficacy of omega-3 fatty acid for GDM, revealing that omega-3 fatty acids supplementation in GDM patients could reduce FPG and HOMA-IR score [31]. Omega-3 fatty acids were involved in the mechanism of increasing β -oxidation of fatty acids, improving antioxidant functions and insulin action, and reducing lipogenesis [32].

Of note, 3-hydroxybutyric acid was selected as a potential metabolic biomarker in both the trimester groups. As a classic ketone body, the levels of 3-hydroxybutyric acid increases because of the oxidation of free fatty acids and excess acetyl-CoA. ATP production from fatty acids and carbohydrate oxidation happens out of control, resulting in increased acetyl-CoA levels. A study in diabetic rats showed that inefficient utilization and mobilization of glucose may contribute to the elevation of 3-hydroxybutyric acid levels [33].

In the WGCNA, carnitine was included in the module grey for both the trimesters. Previous studies have shown that using 2 g/day of L-carnitine resulted in a reduction of TC and LDL, and its mechanism may be related to the phenomena of insulin resistance and lipotoxicity [34]. However, a higher dose of L-carnitine had different effects, since it contributed to the elevation of TG, apolipoprotein-A1, and apolipoprotein-B100 levels [35]. While BCAAs are in the module grey in the second-trimester group, isoleucine belongs to the module grey, leucine and valine are in the module turquoise for the third-trimester group. It is widely accepted that BCAAs transaminase helps in the conversion of isoleucine and valine into branched-chain α -ketoacids, which are further transformed into propionyl-CoA by the branched-chain α -ketoacid dehydrogenase complex. Propionyl-CoA can become methylmalonyl-CoA with relevant carboxylase. Methylmalonyl-CoA mutase (MUT) is an enzyme that catalyzes the conversion of methylmalonyl-CoA to succinyl-CoA [36]. Based on experiments with mice, decreased *Mut* expression led to higher body weight, hyperinsulinemia, elevated fasting glucose and increased triglyceride [37].

Actually, this study has some limitations. First, in a cross-sectional study design, metabolites were only measured at one point; thus, further prospective cohort studies are needed for establishing the dynamic association of these metabolites with GDM. Second, while the second trimester starts at week 14 of pregnancy and lasts through the end of week 27, participants were recruited from 24 weeks to the end of 27 gestational weeks for the second-trimester group; thus, important information may have missed out. Third, the precise molecular mechanisms underlying the development of GDM remain unclear and mechanistic studies need to be conducted for clarifying the exact roles of these discovered metabolites in GDM.

Conclusion

In conclusion, with the analysis of serum samples, our study suggested that specific metabolomic profile existed among GDM patients. Several key metabolites, such as glycine, serine, proline, and 3-hydroxybutyric acid, are associated with GDM in the second or third trimester, and potential biomarkers for GDM have been identified. These metabolites mainly participated in fatty acid and amino acid metabolism, which may shed light on the

pathology of GDM. However, further research is needed to confirm our findings and explore the underlying molecular mechanisms.

Abbreviations

AUC: Area under the curve

BCAAs: Branched-chain amino acids

BMI: Body mass index

DEMs: Differentially expressed metabolites

FPG: Fasting plasma glucose

GDM: Gestational diabetes mellitus

HDL-C: High-density lipoprotein cholesterol

LDL-C: Low-density lipoprotein cholesterol

LR: Logistic regression

mTORC1: mammalian Target of Rapamycin Complex 1

NGT: Normal glucose tolerance

OGTT: Oral glucose tolerance test

OPLS-DA: Orthogonal partial least square discriminant analysis

PCA: Principal component analysis

PLS-DA: Partial least square discriminant analysis

RF: Random forest

ROC: Receiver operating characteristic

SMPDB: Small molecule pathway database

TC: Total cholesterol

TG: Total triglyceride

UPLC-MS/MS: Ultra-performance liquid chromatography coupled to tandem mass spectrometry

VIP: Variable Important for the Projection

WGCNA: Weighted gene co-expression network analysis

Supporting Information

Table S1. List of DEMs in the second trimester group. (xlsx)

Table S2. List of DEMs in the third trimester group. (xlsx)

Table S3. Function enrichments of DEMs in the second trimester group. (xlsx)

Table S4. Function enrichments of DEMs in the third trimester group. (xlsx)

Table S5. Pathway enrichments of DEMs in the second trimester group. (xlsx)

Table S6. Pathway enrichments of DEMs in the third trimester group. (xlsx)

Table S7. Candidate metabolite biomarkers in the second- and third-trimester groups. (xlsx)

Figure S1. Permutation test in the second trimester group (a) and the third trimester group (b). (tif)

Figure S2. Overlap statistic of different findings in two trimester groups. (pdf)

Declarations

Ethics approval and consent to participate

The study was approved by the ethics committee of the University of Hong Kong-Shenzhen Hospital. It was conducted according to the Declaration of Helsinki, and informed consent was signed by every participant.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author (Cheng Hu) on reasonable request.

Competing interests

The authors have no conflicts of interest to declare.

Funding

This study was supported by the National Natural Science Foundation of China (No. 61803360 and No. 81974118), the Shanghai Municipal Science and Technology Major Project (Grant No. 2017SHZDZX01), the Shanghai Municipal Education Commission—Gaofeng Clinical Medicine (Grant No. 20152527), and the Shanghai Science and Technology Innovation Action Plan (Grant No. 20XD1433300).

Authors' contributions

CH and XY supervised the study. ML, MT, and XF collected and checked the data. CH and WL designed the experiments and analyzed the data. RZ, XF, and ML helped to carry out the experiments. XY, WL, and MT undertook the data analysis. WL and XY prepared and revised the manuscript. All authors read and approved the final manuscript. WL and XF contributed equally.

Acknowledgements

We thank Metabo-profile Biotechnology (Shanghai, China) for this metabolomics profiling and metabolomic analysis.

References

1. Zhu Y, Zhang C: **Prevalence of Gestational Diabetes and Risk of Progression to Type 2 Diabetes: a Global Perspective.** *Curr Diab Rep* 2016, **16**:7.
2. Jovanovic L, Pettitt DJ: **Gestational diabetes mellitus.** *JAMA* 2001, **286**:2516-2518.
3. The L: **Gestational diabetes in England: cause for concern.** *Lancet* 2019, **393**:1262.
4. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P: **Gestational diabetes mellitus.** *Nat Rev Dis Primers* 2019, **5**:47.
5. Damm P, Houshmand-Oeregaard A, Kelstrup L, Lauenborg J, Mathiesen ER, Clausen TD: **Gestational diabetes mellitus and long-term consequences for mother and offspring: a view from Denmark.** *Diabetologia* 2016, **59**:1396-1399.
6. Huynh J, Xiong G, Bentley-Lewis R: **A systematic review of metabolite profiling in gestational diabetes mellitus.** *Diabetologia* 2014, **57**:2453-2464.
7. Nicholson JK, Lindon JC: **Systems biology: Metabonomics.** *Nature* 2008, **455**:1054-1056.
8. Vogeser M, Seger C: **A decade of HPLC-MS/MS in the routine clinical laboratory—goals for further developments.** *Clin Biochem* 2008, **41**:649-662.
9. Hsieh Y: **HPLC-MS/MS in drug metabolism and pharmacokinetic screening.** *Expert Opin Drug Metab Toxicol* 2008, **4**:93-101.
10. Liang L, Rasmussen MH, Piening B, Shen X, Chen S, Rost H, Snyder JK, Tibshirani R, Skotte L, Lee NC, et al: **Metabolic Dynamics and Prediction of Gestational Age and Time to Delivery in Pregnant Women.** *Cell* 2020, **181**:1680-1692 e1615.
11. Akturk M, Altinova A, Mert I, Dincel A, Sargin A, Buyukkagnici U, Arslan M, Danisman N: **Asymmetric dimethylarginine concentrations are elevated in women with gestational diabetes.** *Endocrine* 2010, **38**:134-141.
12. Cetin I, de Santis MS, Taricco E, Radaelli T, Teng C, Ronzoni S, Spada E, Milani S, Pardi G: **Maternal and fetal amino acid concentrations in normal pregnancies and in pregnancies with gestational diabetes mellitus.** *Am J Obstet Gynecol* 2005, **192**:610-617.
13. Chen X, de Seymour JV, Han TL, Xia Y, Chen C, Zhang T, Zhang H, Baker PN: **Metabolomic biomarkers and novel dietary factors associated with gestational diabetes in China.** *Metabolomics* 2018, **14**:149.
14. Liu T, Li J, Xu F, Wang M, Ding S, Xu H, Dong F: **Comprehensive analysis of serum metabolites in gestational diabetes mellitus by UPLC/Q-TOF-MS.** *Anal Bioanal Chem* 2016, **408**:1125-1135.

15. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis.** *BMC Bioinformatics* 2008, **9**:559.
16. Yu XT, Wang ZJ, Zeng T: **Essential gene expression pattern of head and neck squamous cell carcinoma revealed by tumor-specific expression rule based on single-cell RNA sequencing.** *Biochimica Et Biophysica Acta-Molecular Basis of Disease* 2020, **1866**.
17. Yu XT, Gan ZJ, Xu Y, Wan SB, Li M, Ding SJ, Zeng T: **Identifying Essential Methylation Patterns and Genes Associated With Stroke.** *Ieee Access* 2020, **8**:96669-96676.
18. Bentley-Lewis R, Huynh J, Xiong G, Lee H, Wenger J, Clish C, Nathan D, Thadhani R, Gerszten R: **Metabolomic profiling in the prediction of gestational diabetes mellitus.** *Diabetologia* 2015, **58**:1329-1332.
19. Pappa KI, Vlachos G, Theodora M, Roubelaki M, Angelidou K, Antsaklis A: **Intermediate metabolism in association with the amino acid profile during the third trimester of normal pregnancy and diet-controlled gestational diabetes.** *Am J Obstet Gynecol* 2007, **196**:65 e61-65.
20. Park S, Park JY, Lee JH, Kim SH: **Plasma levels of lysine, tyrosine, and valine during pregnancy are independent risk factors of insulin resistance and gestational diabetes.** *Metab Syndr Relat Disord* 2015, **13**:64-70.
21. Andersson-Hall U, Gustavsson C, Pedersen A, Malmodin D, Joelsson L, Holmang A: **Higher Concentrations of BCAAs and 3-HIB Are Associated with Insulin Resistance in the Transition from Gestational Diabetes to Type 2 Diabetes.** *J Diabetes Res* 2018, **2018**:4207067.
22. Fitch WL, King JC: **Plasma amino acid, glucose, and insulin responses to moderate-protein and high-protein test meals in pregnant, nonpregnant, and gestational diabetic women.** *Am J Clin Nutr* 1987, **46**:243-249.
23. Kimball SR, Jefferson LS: **Signaling pathways and molecular mechanisms through which branched-chain amino acids mediate translational control of protein synthesis.** *J Nutr* 2006, **136**:227S-231S.
24. O'Connell TM: **The complex role of branched chain amino acids in diabetes and cancer.** *Metabolites* 2013, **3**:931-945.
25. Takashina C, Tsujino I, Watanabe T, Sakaue S, Ikeda D, Yamada A, Sato T, Ohira H, Otsuka Y, Oyama-Manabe N, et al: **Associations among the plasma amino acid profile, obesity, and glucose metabolism in Japanese adults with normal glucose tolerance.** *Nutr Metab (Lond)* 2016, **13**:5.
26. Gonzalez-Ortiz M, Medina-Santillan R, Martinez-Abundis E, von Drateln CR: **Effect of glycine on insulin secretion and action in healthy first-degree relatives of type 2 diabetes mellitus patients.** *Horm Metab Res* 2001, **33**:358-360.
27. Schmidt HH, Warner TD, Ishii K, Sheng H, Murad F: **Insulin secretion from pancreatic B cells caused by L-arginine-derived nitrogen oxides.** *Science* 1992, **255**:721-723.
28. Monti LD, Galluccio E, Villa V, Fontana B, Spadoni S, Piatti PM: **Decreased diabetes risk over 9 year after 18-month oral L-arginine treatment in middle-aged subjects with impaired glucose tolerance and metabolic syndrome (extension evaluation of L-arginine study).** *Eur J Nutr* 2018, **57**:2805-2817.
29. Carvalho AM, Oliveira AA, Loureiro AP, Gattas GJ, Fisberg RM, Marchioni DM: **Arginine intake is associated with oxidative stress in a general population.** *Nutrition* 2017, **33**:211-215.
30. Nuttall FQ, Gannon MC, Jordan K: **The metabolic response to ingestion of proline with and without glucose.** *Metabolism* 2004, **53**:241-246.

31. Zhong N, Wang J: **The efficacy of omega-3 fatty acid for gestational diabetes: a meta-analysis of randomized controlled trials.** *Gynecol Endocrinol* 2019, **35**:4-9.
32. Backes J, Anzalone D, Hilleman D, Catini J: **The clinical relevance of omega-3 fatty acids in the management of hypertriglyceridemia.** *Lipids Health Dis* 2016, **15**:118.
33. Oshida Y, Iwao N, Ohsawa I, Sato J, Nakao T, Sato Y: **Effect of insulin on intramuscular 3-hydroxybutyrate levels in diabetic rats.** *Horm Metab Res* 1998, **30**:70-71.
34. Vidal-Casariiego A, Burgos-Pelaez R, Martinez-Faedo C, Calvo-Gracia F, Valero-Zanuy MA, Luengo-Perez LM, Cuerda-Compes C: **Metabolic effects of L-carnitine on type 2 diabetes mellitus: systematic review and meta-analysis.** *Exp Clin Endocrinol Diabetes* 2013, **121**:234-238.
35. Rahbar AR, Shakerhosseini R, Saadat N, Taleban F, Pordal A, Gollestan B: **Effect of L-carnitine on plasma glycemic and lipidemic profile in patients with type II diabetes mellitus.** *Eur J Clin Nutr* 2005, **59**:592-596.
36. Sivanand S, Vander Heiden MG: **Emerging Roles for Branched-Chain Amino Acid Metabolism in Cancer.** *Cancer Cell* 2020, **37**:147-156.
37. Lerin C, Goldfine AB, Boes T, Liu M, Kasif S, Dreyfuss JM, De Sousa-Coelho AL, Daher G, Manoli I, Sysol JR, et al: **Defects in muscle branched-chain amino acid oxidation contribute to impaired lipid metabolism.** *Mol Metab* 2016, **5**:926-936.

Tables

Table 1

Characteristics of the NGT and GDM groups in the second and third trimesters

	Subjects in the second trimester			Subjects in the third trimester		
	NGT(n=200)	GDM(n=200)	P	NGT(n=199)	GDM(n=199)	P
Age (years)	29(27,30)	29(27,30)	0.117	31(29,34)	31(29,35)	0.208
Pregestational BMI (kg/m ²)	20.33±2.58	20.72±2.77	0.151	21.33±2.26	21.44±2.39	0.665
Changes of BMI (kg/m ²)	5.26±1.31	4.63±1.47	<0.001	5.6±1.46	4.68±1.61	<0.001
Gestational age (week)	25.71(24.86,26.43)	25.86(25,26.5)	0.925	29(28,31.1)	29.29(28,31.1)	0.395
FBG (mmol/L)	4.4±0.32	4.56±0.38	<0.001	4.49±0.25	4.63±0.33	<0.001
1h-PG (mmol/L)	7.22±1.29	9.67±1.34	<0.001	7.35±1.37	9.94±1.51	<0.001
2h-PG (mmol/L)	6.27±0.92	8.5±1.35	<0.001	6.44±1.06	8.79±1.56	<0.001
HbA1c			<0.001			0.962
mmol/mol	32.47±2.08	33.45±2.53		32.93±3.52	32.94±3.14	
%	5.12±0.19	5.21±0.23		5.16±0.32	5.16±0.29	
Total cholesterol (mmol/L)	6.13±1.05	5.83±1.11	0.005	6.64±1.22	6.13±1.16	<0.001
Triglycerides (mmol/L)	1.93±0.65	2.33±0.86	<0.001	2.91±1.38	2.94±1.78	0.850
LDL-C (mmol/L)	2.91±0.74	3.24±0.96	<0.001	3.29±0.89	3.40±0.91	0.233
HDL-C (mmol/L)	2.05±0.37	2.02±0.42	0.488	1.98±0.39	1.95±0.40	0.395
Delivery gestational age (week)	39(38,40)	39(38,40)	0.120	39(39,40)	39(38,40)	0.004
Neonatal weight (kg)	3.29±0.38	3.23±0.46	0.127	3.38±0.39	3.26±0.41	0.005
Neonatal length (cm)	50(50,50)	50(50,50)	0.214	50(50,51)	50(50,50)	0.010
Family history of diabetes, n(%)	5(2.5)	48(24)	<0.001	27(13.6)	53(26.6)	0.001

Data are presented as means±SD, median (interquartile range) or n (%)

NGT, normal glucose tolerance; GDM, gestational diabetes mellitus; FBG, fasting blood glucose; 1h-PG, one hour postprandial glucose; 2h-PG, two hours postprandial glucose; LDL-C, low-density lipoprotein cholesterol; HDL-C,

high-density lipoprotein cholesterol

Table 2

Enrichment analysis of the second trimester group

Pathway Associated Metabolite sets	Enriched compounds	Total	Hits	Raw <i>p</i>	Holm <i>p</i>	FDR
Alpha Linolenic Acid and Linoleic Acid Metabolism	Linoleic acid; Alpha-Linolenic acid; Docosahexaenoic acid; Adrenic acid; 8,11,14-Eicosatrienoic acid	19	5	0.000693	0.0679	0.0679
Beta Oxidation of Very Long Chain Fatty Acids	L-Carnitine; L-Acetylcarnitine; Caprylic acid	17	3	0.0292	1	1
Valine-Leucine-Isoleucine Degradation	Alpha-ketoisovaleric acid; L-Valine; L-Isoleucine; 3-Methyl-2-oxovaleric acid; L-Leucine; Ketoleucine;	60	6	0.0314	1	1

Pathway associated metabolite sets with p value ≤ 0.05 are shown in the table. Total means total number of metabolites in the metabolite set; Hits mean number of metabolites from GDM patients in the metabolite set; Raw p refers to original p value in the enrichment analysis; Holm p refers to adjusted raw p value by Holm-Bonferroni method; FDR is the false discovery rate.

Table 3

Enrichment analysis of the third trimester group

Pathway Associated Metabolite sets	Enriched compounds	Total	Hits	Raw <i>p</i>	Holm <i>p</i>	FDR
Urea Cycle	Pyruvic acid; Oxoglutaric acid; L-Alanine; L-Arginine;L-Glutamine Citrulline	29	6	0.000325	0.0318	0.0176
Ammonia Recycling	Glycine; L-Histidine; L-Glutamine; L-Serine; Oxoglutaric acid; Pyruvic acid	32	6	0.000574	0.0557	0.0176
Glycine and Serine Metabolism	Glycine;Guanidoacetic acid;Pyruvic acid;L-Alanine;L-Serine;L-Arginine Oxoglutaric acid; L-Methionine	59	8	0.000638	0.0612	0.0176
Valine-Leucine-Isoleucine Degradation	Alpha-ketoisovaleric acid; L-Valine; Methylmalonic acid; Oxoglutaric acid; Succinic acid; Ketoleucine; 3-Methyl-2-oxovaleric acid; L-Leucine	60	8	0.000718	0.0682	0.0176
Arginine and Proline Metabolism	Glycine; Guanidoacetic acid; L-Proline; Oxoglutaric acid; Succinic acid; L-Arginine; Citrulline	53	7	0.00173	0.163	0.034
Alanine Metabolism	Glycine; Oxoglutaric acid; Pyruvic acid; L-Alanine	17	4	0.00221	0.206	0.0362
Glutamate Metabolism	Glycine; L-Alanine; Oxoglutaric acid; Pyruvic acid; Succinic acid; L-Glutamine	49	6	0.00571	0.525	0.0747
Aspartate Metabolism	Oxoglutaric acid; L-Arginine; L-Glutamine; N-Acetyl-L-aspartic acid; Citrulline	35	5	0.00609	0.555	0.0747
Glucose-Alanine Cycle	L-Alanine; Oxoglutaric acid; Pyruvic acid	13	3	0.00896	0.807	0.0976
Phenylalanine and Tyrosine Metabolism	L-Tyrosine; Phenylpyruvic acid; 4-Hydroxyphenylpyruvic acid; Oxoglutaric acid	28	4	0.0144	1	0.141
Carnitine Synthesis	Glycine; Oxoglutaric acid; Succinic	22	3	0.0388	1	0.346

Pathway associated metabolite sets with p value ≤ 0.05 are shown in the table.

Table 4

Candidate metabolite biomarkers associated with GDM in the second and third trimester

Name	Second trimester	Third trimester
3-Methyl-2-oxovaleric acid	grey*	grey*
D-Gluconolactone	grey*	
D-Glucose	blue*	
3-Hydroxybutyric acid	grey*	
Alpha-Hydroxyisobutyric acid	grey*	
Isobutyric acid	grey*	turquoise*
Isovaleric acid	grey*	turquoise*
Octanoic acid	brown*	
Glycocholic acid	grey*	
Nonanoic acid	grey*	
Myristic acid	turquoise*	
DHA	turquoise*	
Palmitic acid	turquoise*	
Glycylproline		turquoise*
Alpha-Ketoisovaleric acid		grey*
Ketoleucine		grey*
Acetic acid		grey*
Caproic acid		turquoise*
Heptanoic acid		turquoise*
Pyruvic acid		yellow*
Arachidonic acid		blue*
Adrenic acid		blue
Citramalic acid		turquoise*

Figures

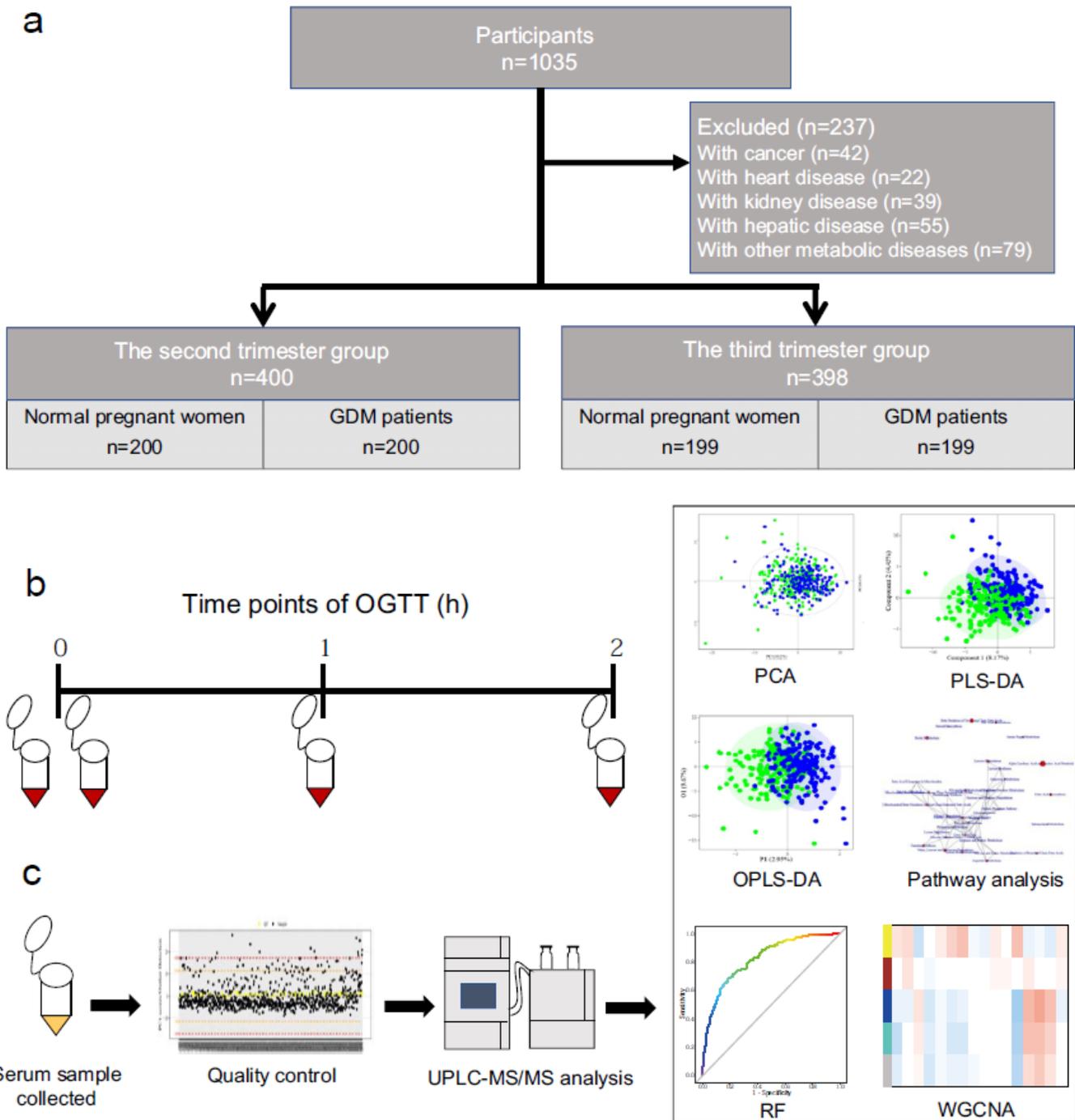


Figure 1

The design and analysis workflow of the study. (a) Participant selection. (b) Time points when blood samples were collected during OGTT. One blood sample at the zero-time point was obtained before OGTT for the serum sample. (c) Metabolic analysis pipeline.

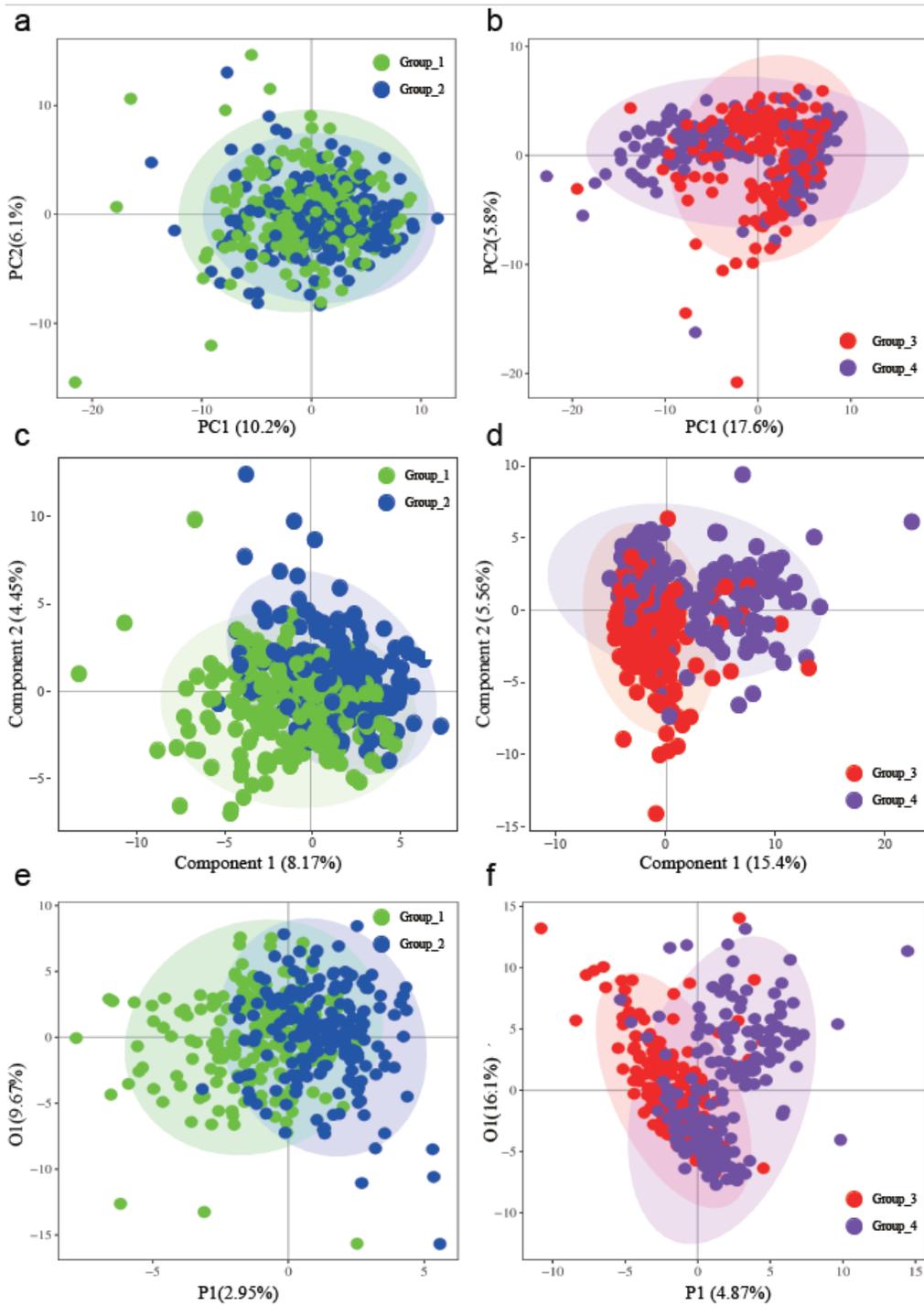


Figure 2

Discrimination analysis of metabolomic profiles. (a) PCA model in the second trimester. (b) PCA model in the third trimester. (c) PLS-DA model in the second trimester. (d) PLS-DA model in the third trimester. (e) OPLS-DA model in the second trimester. (f) OPLS-DA model in the third trimester. Blue refers to the control group while green refers to the case group in the second trimester, and purple refers to normal pregnant women while red refers to GDM patients in the third trimester.

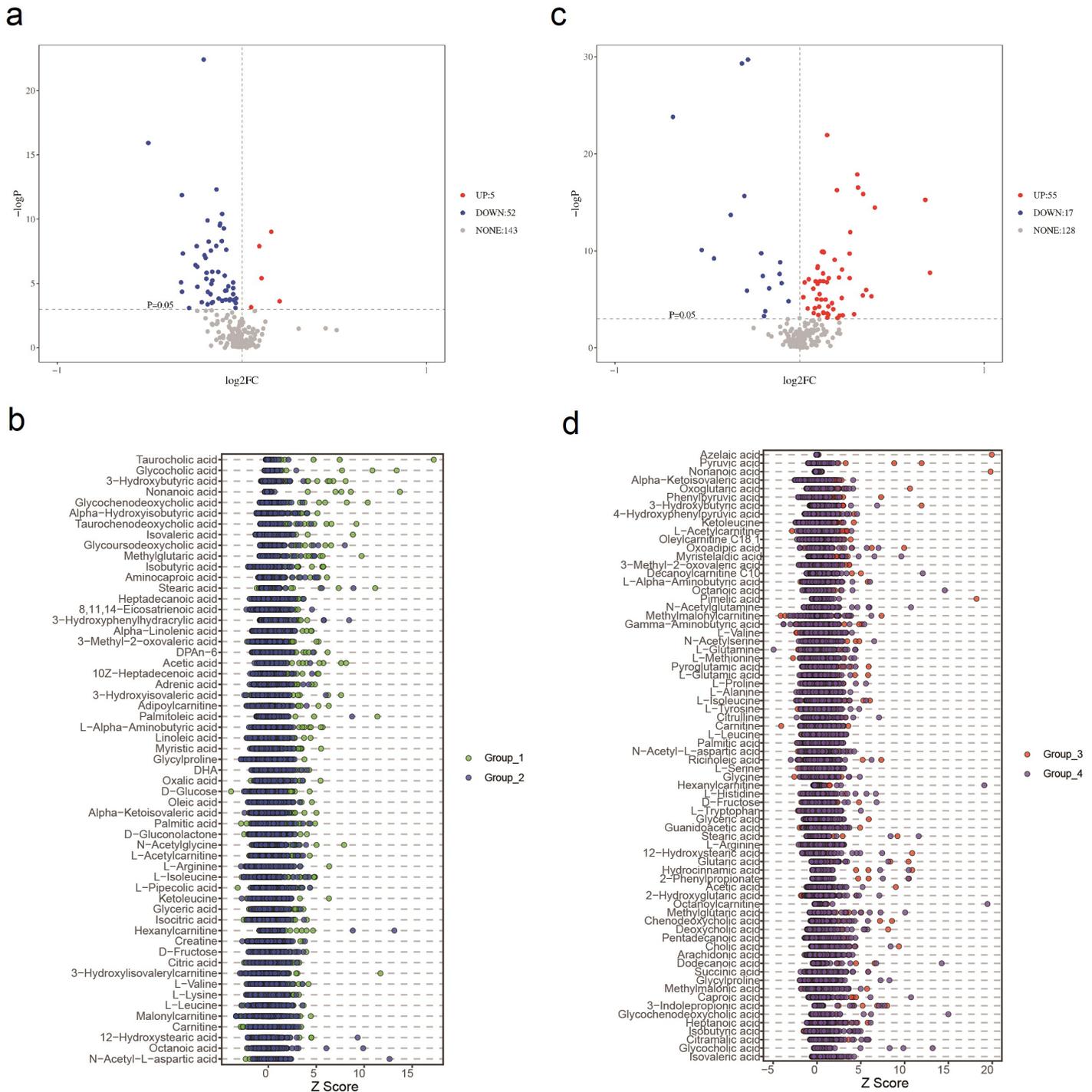


Figure 3

Differential expression analysis of metabolomic profiles. (a) Volcano plot of DEMs in the second-trimester group. (b) Test statistics of DEMs in the second-trimester group. (c) Volcano plot of DEMs in the third-trimester group. (d) Test statistics of DEMs in the third-trimester group.

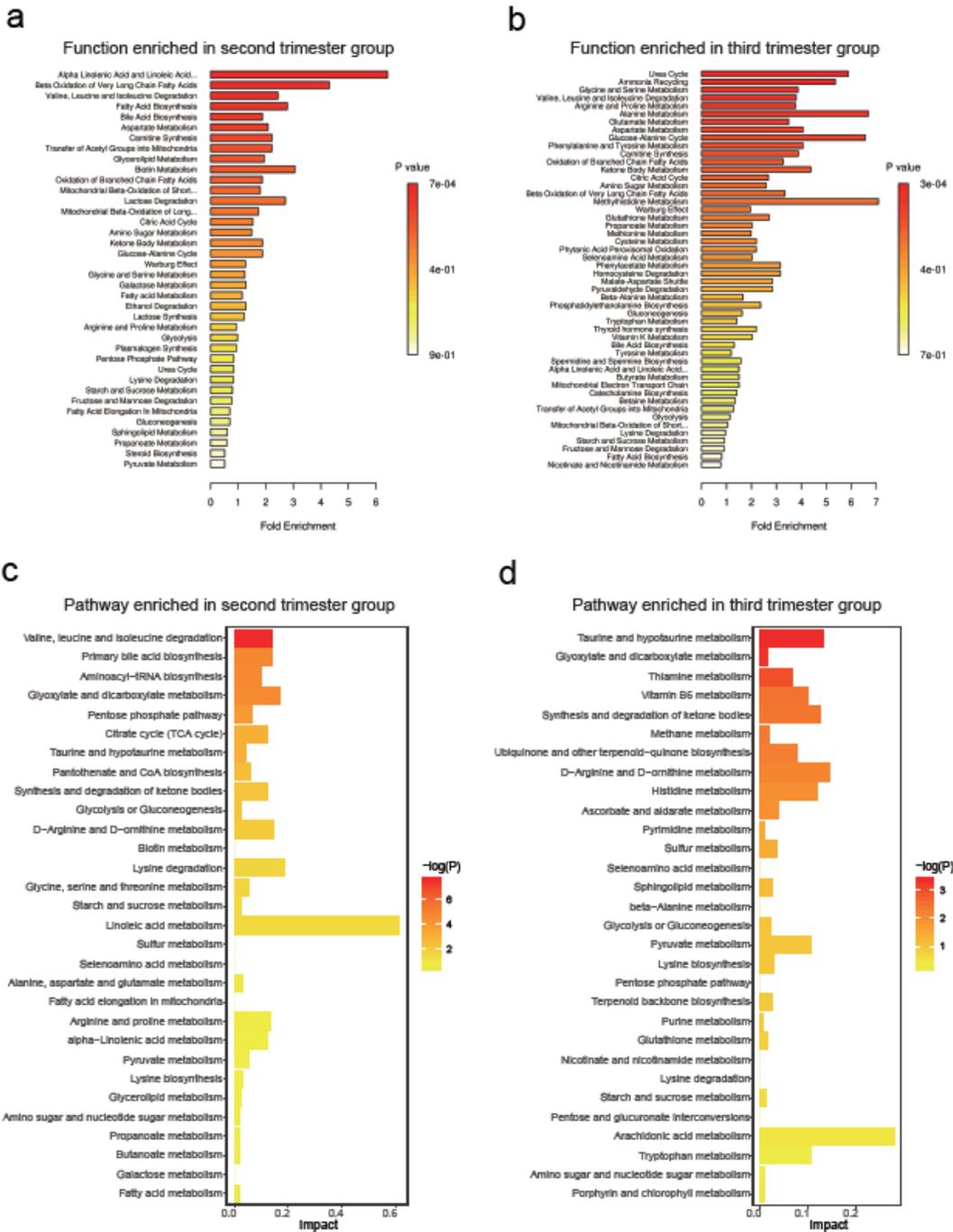


Figure 4

The enrichment analysis of metabolomic profiles. (a) Function enrichment for DEMs in the second-trimester group. (b) Function enrichment for DEMs in the third-trimester group. (c) Pathway enrichment for DEMs in the second- trimester group. (d) Pathway enrichment for DEMs in the third-trimester group.

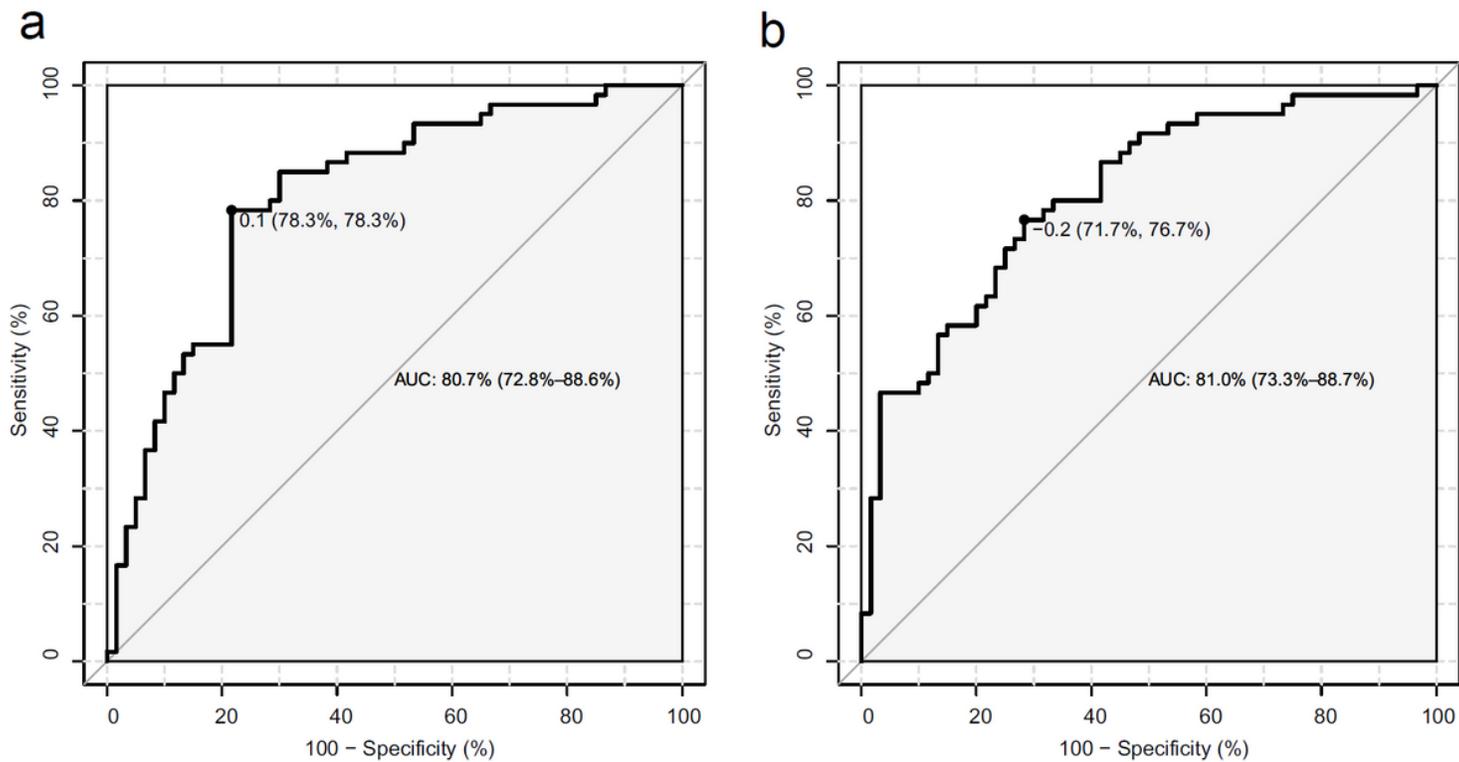


Figure 5

Validation ROC curves of candidate metabolite biomarkers. (a) ROC curve in the second-trimester group. (b) ROC curve in the third-trimester group.

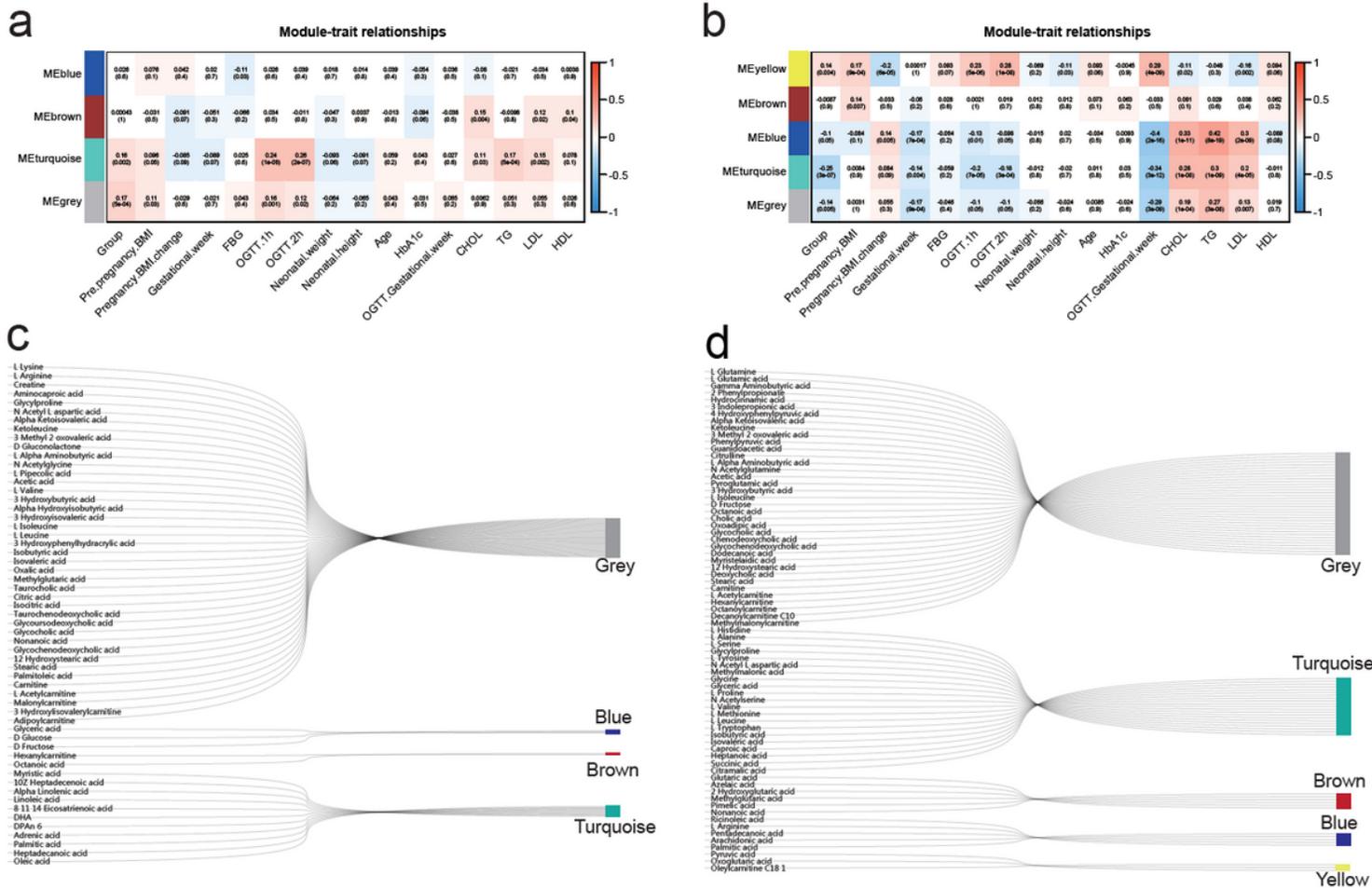


Figure 6

Clinical relevance of DEMs related to GDM. (a) The WGCNA for metabolite module and clinical indices in the second-trimester group. (b) The WGCNA for metabolite module and clinical indices in the third-trimester group. (c) DEMs' distribution in metabolite modules in the second-trimester group. (d) DEMs' distribution in metabolite modules in the third-trimester group.

Supplementary Files

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

- [LuFigureS1.tif](#)
- [LuFigureS2.pdf](#)
- [LuTableS1.xlsx](#)
- [LuTableS2.xlsx](#)
- [LuTableS3.xlsx](#)
- [LuTableS4.xlsx](#)
- [LuTableS5.xlsx](#)
- [LuTableS6.xlsx](#)

- [LuTableS7.xlsx](#)