

# Untargeted Metabolomic Approach to Study the Serum Metabolites in Women with Polycystic Ovary Syndrome

**Ying YU**

Jiangsu University

**PANLI TAN**

Chinese Medicine Hospital of Zhejiang

**ZHENCHAO ZHUANG**

Chinese Medicine Hospital of Zhejiang

**ZHEJIONG WANG**

Chinese Medicine Hospital of Zhejiang

**LINCHAO ZHU**

Chinese Medicine Hospital of Zhejiang

**RUYI QIU**

Chinese Medicine Hospital of Zhejiang

**HUAXI XU** (✉ [xu\\_huaxi@163.com](mailto:xu_huaxi@163.com))

Jiangsu University

## Research Article

**Keywords:** metabolomics analysis, UPLC-HRMS, polycystic ovary syndrome, serum metabolites

**Posted Date:** January 18th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-144365/v1>

**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 BMC Medical Genomics on August 20th, 2021. See the published version at <https://doi.org/10.1186/s12920-021-01058-y>.

## Abstract

**Background:** Polycystic ovary syndrome (PCOS) is a kind of common endocrine syndrome as well as metabolic disorder, which harms the reproductive system and the overall body metabolism of the patients. This study, we aimed to investigate the difference of serum metabolic profiles of the patients with PCOS compared to the healthy controls.

**Material and methods:** a total of 62 study subjects including 31 PCOS patients and 31 matched healthy female controls were recruited, the clinical characteristics data were recorded at the time of recruitment, the laboratory biochemical data were detected. Then, we utilized metabolomics approach by UPLC-HRMS technology to study the metabolic changes between PCOS and controls.

**Results:** The metabolomics analysis showed that there were 68 were downregulated and 78 were upregulated metabolites in PCOS patients serum compared to the controls. These metabolites mainly belong to Triacylglycerols, Glycerophosphocholines, Acylcarnitines, Diacylglycerols, Peptides, Amino acids, Glycerophosphoethanolamines and fatty acid. Pathway analysis showed that these metabolites were enriched in pathways including Glycerophospholipid metabolism, Fatty acid degradation, Fatty acid biosynthesis, Ether lipid metabolism etc. Diagnosis value assessment by ROC analysis showed that the changed metabolites including Leu-Ala/Ile-Ala, 3-(4-Hydroxyphenyl) propionic acid, Ile-Val/Leu-Val, Gly-Val/Val-Gly, aspartic acid, DG(34:2)\_DG(16:0/18:2), DG(34:1)\_DG(16:0/18:1), Phe-Trp, DG(36:1)\_DG(18:0/18:1), Leu-Leu/Leu-Ile had higher AUC values, indicated significant role in PCOS.

**Conclusion:** The present study characterized the difference of metabolic and related pathway profiles in PCOS patients, this finding hope to provide potential metabolic marker for the prognosis and diagnosis of this disease.

## Background

Around the world, approximately 15 ~ 20% of the women in childbearing age are affected with polycystic ovary syndrome (PCOS) according to the Rotterdam criteria [1]. This is one of the most common endocrine syndrome as well as metabolic disorder, which is mainly characterized with hyperandrogenism and insulin resistance (IR). The main manifestations of the PCOS patients are menstrual cycle irregular, oligo-ovulation and/or polycystic ovarian morphology, hyperandrogenism induced hirsutism and acne, and IR induced obesity [2]. But the diagnosis of PCOS is always a controversial issue and the criteria are continues to be updated [3–5]. Except for the impair on ovarian function and the overall body metabolism, the thereby caused anovulatory infertility and recurrent pregnancy loss also have tremendous harm to PCOS patients. In addition, due to the dysfunction of ovary and metabolism, long term consequences such as gynecological cancer, hypertension, atherosclerosis and type 2 diabetes mellitus (T2DM), cardiovascular disease (CVD) are also seems to have high incidences in PCOS women than normal populations [6–8]. In light of these risks, there is a strong need of finding reliable biochemical or molecular marker, which would enable to make the accurate diagnosis and effective therapy of PCOS.

Yet, we know surprising little about the exact mechanisms underlying the pathophysiology of PCOS, and this results in the limited available or effective therapies that ameliorate symptoms of PCOS or improve the thereby induced metabolic complications [9]. And shockingly, half of all women with PCOS are thought to remain undiagnosed. Genomic, proteomic, and metabolomic approaches to study the pathogenesis of diseases have been introduced to various diseases researching. Metabolomics involves the comprehensive characterization of metabolites and metabolism in biological systems, and is widely applied to better disease diagnose, understand the potential mechanisms, identify novel drug targets, customize drug treatments and monitor therapeutic outcomes [10]. Untargeted metabolomic approach, known as metabolic fingerprinting, mainly focuses on the identification and quantification of as many as possible low-molecular-weight compounds present in tested samples. This approach is commonly applied to uncover metabolic profiles, metabolic markers as well as to reveal new insights into the mechanisms underlying the pathogenesis of human diseases, including PCOS [11].

In this study, based on metabolomics approach by utilizing an ultra-performance liquid chromatography-high resolution mass spectrometry (UPLC-HRMS) technology, we aim to characterize the metabolic fingerprints of the patients with PCOS and healthy controls, and also in hope of identifying potential metabolic marker for the prognosis and diagnosis of this disease.

## Material And Methods

### Study subjects

All of the PCOS patients and healthy controls were recruited from the Zhejiang Provincial Hospital of Chinese Medicine (Hangzhou, China). This study was approved by the Ethics Committee of Zhejiang Provincial Hospital of Chinese Medicine. The signed informed consents were obtained from all the participants prior to inclusion in this study.

According to the Rotterdam criteria, 2003, PCOS patients can be diagnosed if two of the three criteria are present after excluding congenital adrenal hyperplasia, Cushing's syndrome, androgen secreting tumors, or other related disorders. The three criteria are (1) oligo- and/or anovulation; (2) clinical and/or biochemical signs of hyperandrogenism (clinical manifestations of hyperandrogenism include presence of acne, hirsutism, and androgenic alopecia); (3) polycystic ovaries by ultrasound examination: presence of 12 or more follicles in each ovary measuring 2–9 mm in diameter and/or ovarian volume >10 cm<sup>3</sup>.

The inclusion criteria for PCOS cases in this study were: diagnosed with PCOS according to the Rotterdam criteria, 2003 [4]; adolescent females (18–40 years old); had at least 2 years of menstrual history. Exclusion criteria: had received any androgenic drug or sex steroid therapy in the past 3 months before the study; current pregnant, delivery or miscarriage within the preceding 3 months; congenital adrenal hyperplasia, androgen-secreting tumors, and other diseases with hyperandrogenism, thyroid dysfunction, hyperprolactinemia, cardiovascular diseases, diabetes or any chronic diseases. Controls group were healthy female volunteers, 18–40 years old, regular menstrual cycles and normal androgen levels, no PCOS and IR, and no evident disease was detected in them during

the course of the study. According to the above mentioned inclusion/exclusion criteria, a total of 31 PCOS patients and 31 health participants were included from December of 2018 to April of 2019 in the present study.

The clinical characteristics data of the enrolled participants were recorded at the time of recruitment. After fasting for 8 h, blood sample from each participant was collected. The serum samples were stored at -80 °C for subsequent assay.

## Clinical laboratory tests

Serum concentrations of Fasting glucose, Fasting insulin, follicle-stimulating hormone (FSH), luteinizing hormone (LH), estradiol (E2), prolactin (PRL), testosterone (T), progesterone (P), TC, TG, HDL-c, LDL-c in all PCOS patients and control participants were detected by Immulite 2000 analyzer (Siemens Healthcare Diagnostics Products Ltd., UK) using two site chemiluminescent immunometric assays.

## Sample preparation and metabolite extraction

*The polar metabolome extraction:* After thawed at 4°C, a 100 µL serum samples were added with 400 µL methanol-acetonitrile (1:1, v:v; including isotope internal standard tryptophan -d5, cetylic acid-[13C]12), centrifuged at 15,000 g for 15 min. Then a 200 µL supernatants were dried under low temperature vacuum (Thermo Scientific, USA) to obtained the sample for UPLC-HRMS analysis. Before analysis, the samples were redissolved with 100 µL 10% methanol (including multiple internal standards).

*The lipidomic metabolome extraction:* After thawed at 4°C, a 50 µL serum samples were added with 300 µL methanol (including internal standards: Ceramide (d18:1/17:0)□PC(17:0/17:0)□TG(15:0/15:0/15:0)), swirled for 120 s, and added with 900 µL MTBE, 250 µL ultrapure water. After vortex mixed and vibrated at room temperature for 15 min, solution was placed under 4°C, 30 min for stratify. Then a 900 µL supernatants were transferred into EP pipe and dried under low temperature vacuum (Thermo Scientific, USA) to obtained the sample for UPLC-HRMS analysis. Before analysis, the samples were redissolved with 600 µL acetonitrile-isopropanol mixture.

## UPLC-HRMS instrumentation and measurement conditions

Untargeted metabolomics analysis was conducted by using three different analytical methods (M1-3) on a Ultimate 3000 ultra-high performance liquid chromatograph coupled with Q ExactiveTM quadrupole-Orbitrap high resolution mass spectrometer (UPLC-HRMS) system (Thermo Scientific, USA).

### UPLC system

Untargeted metabolomics analysis was conducted by using three different analytical methods (M1-3). Method 1 and 2 (M1, 2) was used for the polar metabolome extracts analysis on UPLC-HRMS system with positive and negative ionization detection, respectively. Metabolites were separated by an AcquityTM HSS C18 column (Waters Co., USA, 2.1 × 100 mm) for M1, and eluted by 0.1% formate/water (A) and acetonitrile (B) in linear gradient from 2% organic mobile phase to 98% in 10 min. Furthermore, other mobile phases consisting of water and ammonium acetonitrile/methanol both containing ammonium bicarbonate buffer salt were employed to eluted metabolites separated on an AcquityTM BEH C18 column (Waters Co., USA, 1.7 µm, 2.1 × 100 mm), the gradient was used as follow: from 0~10 min, 2% organic phase ramped to 100%, and from 10~15 min, column washing and equilibrating. Untargeted lipidomic analysis was operated based on Method 3, the chromatographic separation conditions were maintained under positive and negative ionization detection mode, respectively. The used column was an Accucore C30 core-shell column, the mobile phase were 60% acetonitrile in water (A) and 10% acetonitrile in isopropanol (B) both containing 10 mM ammonium formate and 0.1% formate. The separation gradient was optimized as follow: initial 10% B, ramping to 50% in 5 min, and further increasing to 100% in 23 min, the rest 7 min for column washing and equilibration. For Method 1~3, the flow rate was 0.4 mL/min, injection volume was 5 µL, and column temperature was 50°C.

### Mass spectrometer system

For Method 1~2, the quadrupole-Orbitrap mass spectrometer was all operated under identical ionization parameters with a heated electrospray ionization source except ionization voltage including sheath gas 45 arb, aux gas 10 arb, heater temperature 355°C, capillary temperature 320°C and S-Lens RF level 55%. The metabolome extracts were profiled with full scan mode under 70,000 FWHM resolution with AGC 1E6 and 200 ms max injection time. The scan range was 70~1000 m/z. QC samples were repeatedly injected to acquired Top 10 data dependent MS2 spectra (full scan-ddMS2) for comprehensive metabolite and lipid structural annotation. 17,500 FWHM resolution setting were used for full MS/MS data acquisition. Apex trigger, dynamic exclusion and isotope exclusion was turned on, precursor isolation window as set at 1.0 Da. Stepped normalized collision energy was employed for collision induced disassociation of metabolite using ultra-pure nitrogen as fragmentation gas. All the data acquired as centroid format. For Method 3, the ionized lipid molecules were detected using the same parameters as previous description 6.3.1. 300-2000 m/z lipid extracts were profiled with the same parameters as the metabolome used. Lipid were structurally identified through acquiring data dependent MS2 spectra, the key settings included 70,000 FWHM full scan resolution, 17,500 FWHM MS/MS resolution, loop count 10, AGC target 3e6, maximum injection time 200 ms and 80 ms for full scan and MS/MS respectively, dynamic exclusion 8 s. Stepped normalized collision energy 25%+40% and 35% were employed for positive and negative mode after optimization.

## Metabolomics data analysis

The full scan and data-dependent MS2 metabolic profiles data were further processed with Compound Discoverer software for comprehensive component extraction. The polar metabolites were structural annotated through searching acquired MS2 against a local proprietary iPhenomeTM SMOL high resolution MS/MS spectrum library created using authentic standards, NIST 17 Tandem MS/MS library (National Institute of Standards and Technology), local version MoNA (MassBank of North America), as well as mzCloud library (Thermo Scientific, USA). Besides, exact m/z of MS1 spectra was searched against a local KEGG, HMDB metabolite chemical database. For metabolite identification or structural annotation, mass accuracy of precursor within  $\pm$  5 ppm was prerequisite, meanwhile, isotopic information including at least 1 isotopes within 10 ppm and fit score of relative isotopic abundance pattern 70% were introduced to confirm the chemical formula in addition to exact mass. Furthermore, retention time information as well as high resolution MS/MS spectra similarity was employed to strictly confirm the structural annotation of metabolites. The area under curve (AUC) values as extracted as quantitative information of metabolites with XCalibur Quan Browser information, all peak areas data for the annotated metabolites were exported into Excel software for trim and organization before statistics (Microsoft, USA). And on the other hand, untargeted lipidomics data was processed with LipidSearch software including peak picking, lipid identification. The acquired MS2 spectra were searching against in silico predicted spectra of diverse phospholipid, neutral glycerolipid, spingolipid, neutral glycosphingolipids, glycosphingolipids, steroids, fatty ester, etc. The mass accuracy for precursor and MS/MS product ions searching were 5 ppm and 5 mDa, respectively. The MS/MS similarity score threshold was set at 5. The potential ionization adduct including hydrogen, sodium, ammonium for positive and hydrogen loss, formate and acetate adduct for negative mode. The lipid identification was strictly manually checked and investigated one-by-one to eliminate false positive chiefly basing on peak shape, adduct ions behavior, fragmentation pattern, and chromatographic behavior.

## Statistical analysis

All the clinical data were computed using SPSS18.0 version software. An unpaired, two-tailed student *t* test was performed on clinical biochemical data, the chi-square test was used for comparison of categorical variables. *p*-value < 0.05 was considered to be statistically significant. The metabolome and lipidome data deriving from different measurements were normalized to sample weight used prior to further process, respectively. Then, the resultant quantitative information from foregoing methods were merged together and those detected with multiple methods was excluded to guaranteed uniqueness of metabolite and lipid, and then Log10 tranformed for final statistical analysis. The principal component analysis was conducted with SIMCA-P software (Umetrics, Sweden), and other univariate analysis including independent sample *t*-test and *p* value FDR adjust, as well as metabolic pathway analysis was conduct on MetaboAnalyst website.

## Results

### Clinical characteristics and biochemical data of the study subjects

The Clinical characteristics and biochemical data of the study subjects were collected and analyzed, and presented in Table 1. In this study, the study subjects included 31 healthy controls and 31 PCOS women. There is no statistically differences for the age, BMI between the two groups (*p* value > 0.05). For biochemical data, the levels of fasting glucose, LH, T, TG, LDL-c and LH/FSH ratio were significantly higher in PCOS patients than those in controls, the levels of PRL, HDL-c were significantly lower in PCOS patients than those in controls (*p* value < 0.05).

Table 1  
Clinical characteristic and biochemical data of the study subjects

	Control (n = 31)	PCOS (n = 31)	<i>p</i> -value
Age [years]	24.52 $\pm$ 2.31	24.20 $\pm$ 4.49	0.750
BMI [kg/m <sup>2</sup> ]	20.48 $\pm$ 2.67	22.27 $\pm$ 3.56	0.081
Fasting glucose [mmol/L]	4.68 $\pm$ 0.42	5.25 $\pm$ 1.20	0.026 < 0.05
Fasting insulin [pmol/L]	6.9 $\pm$ 3.38	14.10 $\pm$ 13.83	0.007 < 0.05
FSH [IU/L]	5.15 $\pm$ 1.32	5.59 $\pm$ 2.86	0.440
LH [IU/L]	5.85 $\pm$ 2.74	9.82 $\pm$ 8.57	0.017 < 0.05
LH/FSH	1.14 $\pm$ 0.50	1.73 $\pm$ 0.97	0.003 < 0.01
PRL [mIU/L]	466.83 $\pm$ 231.05	309.64 $\pm$ 158.59	0.003 < 0.01
E2 [pmol/L]	209.95 $\pm$ 127.20	220.25 $\pm$ 246.80	0.837
T [nmol/L]	1.12 $\pm$ 0.40	1.65 $\pm$ 0.67	< 0.001
P [nmol/L]	0.93 $\pm$ 0.36	1.05 $\pm$ 0.80	0.451
TC [mmol/L]	4.44 $\pm$ 0.63	4.65 $\pm$ 0.76	0.306
TG [mmol/L]	0.75 $\pm$ 0.29	1.13 $\pm$ 0.51	0.002 < 0.05
HDL-c [mmol/L]	2.23 $\pm$ 0.52	1.46 $\pm$ 0.41	< 0.001
LDL-c [mmol/L]	1.69 $\pm$ 0.29	2.54 $\pm$ 0.62	< 0.001

## Multivariate statistical analysis

The multivariate statistical analysis were performed using SIMCAP software. The PCA (principal component analysis) analysis outlined the original distribution of metabolites in PCOS and control subjects. As shown in Fig. 1a, score plot of PCA suggested that there is no obvious outlier samples in two groups. The scatter plot classification in PCOS and control groups was observed in t[2] axis, but was failed to separate in t[1] axis. Hence, OPLS-DA model was further applied. As the results indicated in Fig. 1b, PCOS samples could be clearly distinguished from the healthy control samples. The models possessed satisfactory fit of  $R^2 = 0.93$ ,  $Q^2 = 0.70$ , which indicated the significant discrimination of the serum metabolomics signature between the two study groups. Figure 1c showed the result of permutation test made of OPLS-DA model, permutation plots of OPLS-DA model repeated 999 times verified the validation of the analysis. S plot of the OPLS-DA model indicated the influence of metabolite expression level on metabolic phenotype classification (Fig. 1d).

## Significant changed metabolomics identification by UPLC-HRMS

The variables with FDR adjusted p value  $< 0.05$  was selected as remarkable significance in the OPLS-DA model. And these metabolites were identified through the Compound Discoverer software for polar metabolites and LipidSearch software for lipidomics with the online database (like HMDB) or libraries. As a result, a total of 146 significantly changed metabolites were identified and selected as potential biomarkers of PCOS for subsequent analysis. Volcano plot showed that compared to the controls group, among these 146 significantly varied metabolites, 68 were downregulated, 78 were upregulated (Fig. 2a, Table 2). Heatmap of these 146 significantly changed metabolites in 31 PCOS samples and 31 control samples indicated that these metabolites have clustering correlativity in PCOS patients from healthy controls (Fig. 2b). After chemical structure classification of the identified 146 differential metabolites, Fig. 3 showed that the significantly changed metabolites mainly belonged to Triacylglycerol (36), Glycerophosphocholine (34), Acylcarnitine (15), Diacylglycerol (15), Peptide (10), Amino acid (8), Glycerophosphoethanolamine (6), Fatty acid (FA) (3), etc. And top of the changed percentage metabolites (changed metabolites/total metabolites in this classification) were Diacylglycerol (78.95 %), Choline (50.00 %), Acylcarnitine (48.39 %), Peptide (43.48 %), Nucleoside&nucleotide (27.27 %), Glycerophosphocholine (21.12 %), Neutral glycosphingolipid (18.18 %), Triacylglycerol (13.74 %), etc.

Table 2  
The details about the differential metabolites identified in PCOS patients

Class	Metabolite	HMDB ID	Detection Method	Mass accuracy (PPM)	FD ad p-v
Acylcarnitine	Palmitoylcarnitine(AcCa(16:0)	HMDB0000222	M1	-1.54	0.0
Acylcarnitine	Oleoylcarnitine(AcCa(18:1)	HMDB0005065	M1	-2.08	0.0
Acylcarnitine	Undecanoylcarnitine(AcCa(11:0)	HMDB0013321	M1	-1.68	0.0
Acylcarnitine	Isobutyryl-L-carnitine(AcCa(4:0)	HMDB0000736	M1	-0.47	0.0
Acylcarnitine	Tetradecanoylcarnitine(AcCa(14:0)	HMDB0005066	M1	-1.53	0.0
Acylcarnitine	Linoleyl carnitine(AcCa(18:2)	HMDB0006469	M1	-2.03	0.0
Acylcarnitine	Dodecenoylcarnitine(AcCa(12:1)	HMDB0013326	M1	-1.62	0.0
Acylcarnitine	4-Decenoylcarnitine(AcCa(10:1)	HMDB0013205	M1	-1.67	0.0
Acylcarnitine	Hexanoylcarnitine(AcCa(6:0)	HMDB0000705	M1	-0.55	0.0
Acylcarnitine	9-Hexadecenoylcarnitine(AcCa(16:1)	HMDB0013207	M1	-1.33	0.0
Acylcarnitine	Tetradecenoylcarnitine(AcCa(14:1)	HMDB0002014	M1	-1.30	0.0
Acylcarnitine	Dodecanoylcarnitine(AcCa(12:0)	HMDB0002250	M1	-1.50	0.0
Acylcarnitine	Octanoylcarnitine(AcCa(8:0)	HMDB0000791	M1	-1.52	0.0
Acylcarnitine	Tetradecadienylcarnitine(AcCa(14:2)	HMDB0013331	M1	-1.49	0.0
Acylcarnitine	Decanoylcarnitine(AcCa(10:0)	HMDB0000651	M1	-1.09	0.0
Amino acid	Aspartic acid	HMDB0000191	M2	3.33	0.0
Amino acid	L-Glutamic acid	HMDB0000148	M2	3.78	0.0
Amino acid	L-Cystine	HMDB0000192	M2	1.79	0.0
Amino acid	L-Phenylalanine	HMDB0000159	M1	-0.20	0.0
Amino acid	Acetylspermidine	HMDB0001276	M1	-0.92	0.0
Amino acid	Taurine	HMDB0000251	M2	3.78	0.0
Amino acid	3-(2-Hydroxyphenyl)propanoic acid	HMDB0033752	M2	3.36	0.0
Amino acid	Proline betaine	HMDB0004827	M1	-0.08	0.0
Carbohydrate	Glucaric acid	HMDB0000663	M2	2.69	0.0
Carbohydrate	L-Lactic acid	HMDB0000190	M2	3.88	0.0
Cholesterol ester	ChE(16:1)_ChE(16:1)	HMDB0000658	M3	-0.36	0.0
Choline	Glycerophosphocholine	HMDB0000086	M1	-0.92	0.0
Diacylglycerol	DG(34:2)_DG(16:0/18:2)	HMDB0007103	M3	2.54	0.0
Diacylglycerol	DG(36:4)_DG(16:0/20:4)	HMDB0007113	M3	2.35	0.0
Diacylglycerol	DG(34:1)_DG(16:0/18:1)	HMDB0007101	M3	3.49	0.0
Diacylglycerol	DG(36:1)_DG(18:0/18:1)	HMDB0007159	M3	2.88	0.0
Diacylglycerol	DG(34:3)_DG(16:1/18:2)	HMDB0007132	M3	0.49	0.0
Diacylglycerol	DG(36:2)_DG(18:0/18:2)	HMDB0007161	M3	1.49	0.0
Diacylglycerol	DG(34:2)_DG(16:1/18:1)	HMDB0007131	M3	2.14	0.0
Diacylglycerol	DG(40:7)_DG(18:1/22:6)	HMDB0007208	M3	-0.89	0.0
Diacylglycerol	DG(38:4)_DG(18:1/20:3)	HMDB0007198	M3	1.90	0.0
Diacylglycerol	DG(36:3)_DG(18:1/18:2)	HMDB0007219	M3	2.40	0.0
Diacylglycerol	DG(36:4)_DG(18:2/18:2)	HMDB0007248	M3	-0.73	0.0
Diacylglycerol	DG(36:2)_DG(18:1/18:1)	HMDB0007218	M3	3.21	0.0
Diacylglycerol	DG(40:6)_DG(18:1/22:5)	HMDB0007207	M3	1.92	0.0

Class	Metabolite	HMDB ID	Detection Method	Mass accuracy (PPM)	FD ad p-v
Diacylglycerol	DG(38:5)_DG(18:2/18:3)	HMDB0007112	M3	1.76	0.0
Diacylglycerol	DG(38:4)_DG(18:0/20:4)	HMDB0007170	M3	2.64	0.0
Fatty acid	Dihomo-alpha-linolenic acid (FFA(20:3n3))	HMDB0060039	M2	1.69	0.0
Fatty acid	Myristoleic acid isomer2	HMDB0002000	M2	2.46	0.0
Fatty acid	Myristoleic acid isomer1	HMDB0002000	M2	2.46	0.0
Glycerophosphocholine	LysoPC(0-18:1/0:0)	LMGP01060039	M1	-0.46	0.0
Glycerophosphocholine	LysoPC(16:0e)		M1	-1.00	0.0
Glycerophosphocholine	LysoPC(0:0/18:0)	HMDB0011128	M1	-0.65	0.0
Glycerophosphocholine	LysoPC(0-18:0/0:0)	HMDB0011149	M1	-0.63	0.0
Glycerophosphocholine	LysoPC(18:0/0:0)	HMDB0010384	M1	-0.42	0.0
Glycerophosphocholine	LysoPC(P-18:1/0:0)	HMDB0010408	M1	-0.41	0.0
Glycerophosphocholine	LysoPC(P-16:0/0:0)	HMDB0010407	M1	-0.96	0.0
Glycerophosphocholine	LysoPC(19:0)		M1	-1.02	0.0
Glycerophosphocholine	LysoPC(20:1/0:0)	HMDB0010391	M1	-1.33	0.0
Glycerophosphocholine	PC(34:2)_PC(16:0/18:2)	HMDB0007973	M3	-0.45	0.0
Glycerophosphocholine	PC(32:0e)_PC(32:0e)	LMGP01020029	M3	1.50	0.0
Glycerophosphocholine	PC(37:2)_PC(37:2)mixture	HMDB0008592	M3	2.08	0.0
Glycerophosphocholine	PC(32:0p)_PC(32:0p)	HMDB0011206	M3	1.11	0.0
Glycerophosphocholine	PC(36:2e)_PC(36:2e)	HMDB0013418	M3	2.44	0.0
Glycerophosphocholine	PC(36:1p)_PC(18:0p/18:1)	HMDB0008127	M3	2.44	0.0
Glycerophosphocholine	PC(34:0p)_PC(34:0p)	HMDB0011239	M3	1.82	0.0
Glycerophosphocholine	PC(38:6p)_PC(38:6p)	HMDB0011229	M3	-1.62	0.0
Glycerophosphocholine	PC(34:2p)_PC(34:2p)	HMDB0011211	M3	0.42	0.0
Glycerophosphocholine	PC(35:2)_PC(17:0/18:2)	LMGP01011505	M3	-2.68	0.0
Glycerophosphocholine	PC(34:1p)_PC(34:1p)	HMDB0011210	M3	1.37	0.0
Glycerophosphocholine	LysoPC(18:2/0:0)	HMDB0010386	M1	-0.44	0.0
Glycerophosphocholine	PC(33:2)_PC(33:2)	HMDB0007940	M3	-2.06	0.0
Glycerophosphocholine	PC(36:3)_PC(16:0/20:3)	HMDB0007980	M3	-2.02	0.0
Glycerophosphocholine	LysoPC(0:0/18:2)	HMDB0061700	M1	-0.44	0.0
Glycerophosphocholine	PC(42:3)_PC(18:1/24:2)		M3	0.18	0.0
Glycerophosphocholine	PC(36:3)_PC(18:1/18:2)	HMDB0008105	M3	-2.02	0.0
Glycerophosphocholine	PC(32:1p)_PC(32:1p)		M3	-2.43	0.0
Glycerophosphocholine	PC(40:7p)_PC(40:7p)	HMDB0011295	M3	-2.19	0.0
Glycerophosphocholine	PC(33:0e)_PC(18:0e/15:0)		M3	1.27	0.0
Glycerophosphocholine	PC(32:1e)_PC(32:1e)	HMDB0013404	M3	-0.67	0.0
Glycerophosphocholine	PC(35:2)_PC(35:2)		M3	0.99	0.0
Glycerophosphocholine	PC(35:3)_PC(35:3)		M3	-2.26	0.0
Glycerophosphocholine	PC(30:0e)_PC(30:0e)	HMDB0013341	M3	1.32	0.0
Glycerophosphocholine	PC(33:0p)_PC(33:0p)	HMDB0011238	M3	0.39	0.0
Glycerophosphoethanolamine	LysoPE(0:0/22:6)	HMDB0011496	M1	-0.76	0.0
Glycerophosphoethanolamine	LysoPE(0:0/18:2)	HMDB0011477	M1	-1.03	0.0
Glycerophosphoethanolamine	PE(40:5)_PE(18:1/22:4)	HMDB0009075	M3	-1.19	0.0

Class	Metabolite	HMDB ID	Detection Method	Mass accuracy (PPM)	FD ad p-v
Glycerophosphoethanolamine	LysoPE(20:4/0:0)	HMDB0011517	M1	-1.41	0.0
Glycerophosphoethanolamine	LysoPE(22:6/0:0)	HMDB0011526	M1	-0.88	0.0
Glycerophosphoethanolamine	LysoPE(18:2/0:0)	HMDB0011507	M1	-0.84	0.0
Microbial metabolites	3-(4-Hydroxyphenyl)propionic acid(Desaminotyrosine)	HMDB0002199	M2	-0.05	0.0
Neutral glycosphingolipid	CerG3GNAc1(d34:1)_Tetrahexosylceramide(d18:1/16:0)	HMDB0004960	M3	1.29	0.0
Neutral glycosphingolipid	CerG2(d42:1)_LacCer(d18:1/24:0)	HMDB0011595	M3	2.97	0.0
Neutral glycosphingolipid	CerG3(d42:2)_CerG3(d18:1/24:1)	HMDB0004883	M3	3.52	0.0
Neutral glycosphingolipid	CerG3(d42:2)_CerG3(d42:2)	HMDB0004883	M3	3.52	0.0
Nucleoside&nucleotide	Hypoxanthine	HMDB0000157	M1	0.35	0.0
Nucleoside&nucleotide	Xanthine	HMDB0000292	M1	-0.03	0.0
Nucleoside&nucleotide	Inosine	HMDB0000195	M2	0.44	0.0
Organic acid	Glutaric acid isomers	HMDB0000661	M2	3.91	0.0
Peptide	Gly-Val/Val-Gly	HMDB0028854/HMDB0029127	M2	3.53	0.0
Peptide	Glp-Gly	HMDB0061890	M2	3.62	0.0
Peptide	Phe-Phe	HMDB0013302	M2	2.02	0.0
Peptide	Phe-Trp	HMDB0029006	M1	-1.33	0.0
Peptide	Ser-Hyp/Hyp-Ser	HMDB0029040/HMDB0028872	M1	-0.35	0.0
Peptide	Glu-Ile	HMDB0028822	M1	-0.82	0.0
Peptide	Glu-Gln	HMDB0028817	M2	0.67	0.0
Peptide	Leu-Leu/Leu-Ile	HMDB0028933	M1	-0.61	0.0
Peptide	Ile-Val/Leu-Val	HMDB0028920/HMDB0028942	M1	-0.53	0.0
Peptide	Leu-Ala/Ile-Ala	HMDB0028922/HMDB0028900	M1	0.09	0.0
Sphingolipid	Sphingosine	HMDB0000252	M1	-1.44	0.0
Sphingolipid	Sphinganine 1-phosphate	HMDB0001383	M1	-1.74	0.0
Triacylglycerol	TG(46:0)_TG(16:0/14:0/16:0)	HMDB0010411	M3	-1.43	0.0
Triacylglycerol	TG(48:0)_TG(16:0/16:0/16:0)	HMDB0005356	M3	-0.85	0.0
Triacylglycerol	TG(50:0)_TG(18:0/16:0/16:0)	HMDB0108576	M3	-0.59	0.0
Triacylglycerol	TG(60:9)_TG(18:0/20:4/22:5)	HMDB0045200	M3	-0.79	0.0
Triacylglycerol	TG(52:1)_TG(18:0/16:0/18:1)	HMDB0010431	M3	0.02	0.0
Triacylglycerol	TG(48:1)_TG(16:0/14:0/18:1)	HMDB0010414	M3	-1.12	0.0
Triacylglycerol	TG(54:6)_TG(16:0/16:0/22:6)	HMDB0044613	M3	-0.58	0.0
Triacylglycerol	TG(54:7)_TG(18:1/14:0/22:6)	HMDB0049719	M3	0.61	0.0
Triacylglycerol	TG(56:6)_TG(18:0/16:0/22:6)	HMDB0044747	M3	-0.08	0.0
Triacylglycerol	TG(52:6)_TG(14:0/16:0/22:6)	HMDB0042903	M3	0.80	0.0
Triacylglycerol	TG(48:2)_TG(12:0/18:1/18:1)	LMGL03012670	M3	-1.31	0.0
Triacylglycerol	TG(48:2)_TG(16:0/14:0/18:2)	HMDB0010415	M3	-1.31	0.0
Triacylglycerol	TG(50:1)_TG(16:0/16:0/18:1)	HMDB0005360	M3	-0.42	0.0
Triacylglycerol	TG(51:1)_TG(16:0/17:0/18:1)	LMGL03010051	M3	3.30	0.0
Triacylglycerol	TG(54:4)_TG(18:0/16:0/20:4)	HMDB0044738	M3	-0.97	0.0
Triacylglycerol	TG(52:2)_TG(18:0/16:0/18:2)	HMDB0044734	M3	-0.09	0.0
Triacylglycerol	TG(52:4)_TG(16:0/16:0/20:4)	HMDB0005363	M3	-0.31	0.0
Triacylglycerol	TG(54:1)_TG(18:0/18:0/18:1)	HMDB0005395	M3	2.99	0.0

Class	Metabolite	HMDB ID	Detection Method	Mass accuracy (PPM)	FD ad p-v
Triacylglycerol	TG(47:2)_TG(14:0/15:0/18:2)	HMDB0043227	M3	0.36	0.0
Triacylglycerol	TG(51:0)_TG(18:0/16:0/17:0)	HMDB0108587	M3	0.02	0.0
Triacylglycerol	TG(49:1)_TG(15:0/16:0/18:1)	HMDB0043027	M3	-0.54	0.0
Triacylglycerol	TG(49:1)_TG(49:1)		M3	-0.54	0.0
Triacylglycerol	TG(54:2)_TG(18:0/18:0/18:2)	HMDB0005397	M3	2.83	0.0
Triacylglycerol	TG(56:4)_TG(18:0/18:0/20:4)	HMDB0044771	M3	-0.07	0.0
Triacylglycerol	TG(50:2)_TG(16:0/16:0/18:2)	HMDB0005362	M3	-1.69	0.0
Triacylglycerol	TG(52:0)_TG(18:0/16:0/18:0)	HMDB0044722	M3	2.83	0.0
Triacylglycerol	TG(53:1)_TG(15:0/22:0/16:1)	HMDB0043112	M3	-0.49	0.0
Triacylglycerol	TG(53:1)_TG(17:0/18:0/18:1)	LMGL03010119	M3	-0.49	0.0
Triacylglycerol	TG(56:7)_TG(16:0/18:1/22:6)	HMDB0044135	M3	-0.72	0.0
Triacylglycerol	TG(54:5)_TG(18:0/18:2/18:3)	HMDB0045301	M3	0.28	0.0
Triacylglycerol	TG(54:5)_TG(16:0/18:1/20:4)	HMDB0044098	M3	0.28	0.0
Triacylglycerol	TG(54:2)_TG(18:0/18:1/18:1)	HMDB0005403	M3	-0.41	0.0
Triacylglycerol	TG(54:6)_TG(16:0/16:1/22:5)	HMDB0044591	M3	0.04	0.0
Triacylglycerol	TG(54:6)_TG(18:0/18:3/18:3)	HMDB0052887	M3	0.04	0.0
Triacylglycerol	TG(54:6)_TG(16:0/18:1/20:5)	HMDB0044133	M3	0.04	0.0
Triacylglycerol	TG(56:5)_TG(18:1/18:1/20:3)	HMDB0049883	M3	-0.64	0.0
Unannotated	X310-1-MZ310RT517		M1	-1.13	0.0
Unannotated	X310-2-MZ310RT526		M1	-2.61	0.0
Unannotated	X310-1-MZ310RT307		M1	-1.63	0.0
Unannotated	X310-2-MZ310RT320		M1	-1.73	0.0

## Metabolite enrichment and Metabolic pathway analysis

Based on these identified metabolites, metabolic pathway analysis (MetPA) analysis was performed (Fig. 4a, Table 3). In Fig. 4a, -log (p value) and pathway impact were the X and Y axes of the bubble diagram. It could be observed that these metabolites were significantly enriched in metabolic pathways including Glycerophospholipid metabolism, Sphingolipid metabolism, Phenylalanine, tyrosine and tryptophan biosynthesis, Arginine biosynthesis, Histidine metabolism, Ether lipid metabolism. Furthermore, metabolites set enrichment analysis (MSEA) was also performed based on the Metabolites Set in KEGG database (Fig. 4b, Table 4). As shown in Fig. 4b, -log (p value) and fold change were the X and Y axes of the bubble diagram. MSEA analysis showed that Purine metabolism, Porphyrin and chlorophyll metabolism, Fatty acid degradation, Taurine and hypotaurine metabolism, Phenylalanine metabolism, Phenylalanine, tyrosine and tryptophan biosynthesis, Fatty acid biosynthesis, etc., were involved in metabolic pathways of these metabolites enriched.

Table 3  
The pathways of the differential metabolites enriched

Pathway name	Match status	p-value	-log (p)	Impact
Glycerophospholipid metabolism	4/36	0.00555	5.1945	0.26445
Sphingolipid metabolism	3/21	0.00826	4.7961	0.06896
Arginine biosynthesis	2/14	0.03220	3.4358	0.11675
Histidine metabolism	2/16	0.04139	3.1848	0.00000
Ether lipid metabolism	2/20	0.06227	2.7763	0.14458
Aminoacyl-tRNA biosynthesis	3/48	0.07376	2.6069	0.00000
Phenylalanine, tyrosine and tryptophan biosynthesis	1/4	0.08013	2.5241	0.50000
Linoleic acid metabolism	1/5	0.09917	2.3109	0.00000
Alanine, aspartate and glutamate metabolism	2/28	0.11195	2.1897	0.42068
Nitrogen metabolism	1/6	0.11783	2.1385	0.00000
D-Glutamine and D-glutamate metabolism	1/6	0.11783	2.1385	0.50000
Purine metabolism	3/65	0.14751	1.9139	0.03102
Ascorbate and aldarate metabolism	1/8	0.15403	1.8706	0.00000
Taurine and hypotaurine metabolism	1/8	0.15403	1.8706	0.42857
Phenylalanine metabolism	1/10	0.18879	1.6671	0.35714
alpha-Linolenic acid metabolism	1/13	0.23835	1.4340	0.00000
Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	1/14	0.25420	1.3696	0.00399
Butanoate metabolism	1/15	0.26974	1.3103	0.00000
Nicotinate and nicotinamide metabolism	1/15	0.26974	1.3103	0.00000
Glycerolipid metabolism	1/16	0.28496	1.2554	0.01402
Pantothenate and CoA biosynthesis	1/19	0.32881	1.1123	0.00000
beta-Alanine metabolism	1/21	0.35659	1.0312	0.00000
Pyruvate metabolism	1/22	0.37005	0.9941	0.08398
Glutathione metabolism	1/28	0.44529	0.8090	0.01966
Porphyrin and chlorophyll metabolism	1/30	0.46838	0.7585	0.00000
Glyoxylate and dicarboxylate metabolism	1/32	0.49054	0.7123	0.00000
Cysteine and methionine metabolism	1/33	0.50128	0.6906	0.00000
Arachidonic acid metabolism	1/36	0.53220	0.6307	0.00000
Arginine and proline metabolism	1/38	0.55177	0.5946	0.08600
Fatty acid degradation	1/39	0.56126	0.5776	0.00000
Steroid biosynthesis	1/42	0.58856	0.5301	0.00000
Primary bile acid biosynthesis	1/46	0.62242	0.4741	0.00758

Table 4  
MSEA pathway analysis of the differential metabolites enriched

Metabolite Set	Total	Hits	FDR
Purine metabolism	65	4	0.000000
Porphyrin and chlorophyll metabolism	30	3	0.000186
Taurine and hypotaurine metabolism	8	1	0.016056
Fatty acid degradation	39	1	0.040924
Phenylalanine metabolism	10	2	0.044818
Phenylalanine, tyrosine and tryptophan biosynthesis	4	2	0.044818
Fatty acid biosynthesis	47	1	0.070722
Primary bile acid biosynthesis	46	3	0.040924
Cysteine and methionine metabolism	33	3	0.050999
Pentose and glucuronate interconversions	18	1	0.205370
Valine, leucine and isoleucine degradation	40	3	0.205370
Sphingolipid metabolism	21	5	0.205370
Glutathione metabolism	28	2	0.205370
Arginine biosynthesis	14	3	0.205370
Valine, leucine and isoleucine biosynthesis	8	4	0.208410
Arginine and proline metabolism	38	4	0.205370
beta-Alanine metabolism	21	2	0.249570
Steroid hormone biosynthesis	85	3	0.252260
Selenocompound metabolism	20	1	0.480370
Aminoacyl-tRNA biosynthesis	48	17	0.443750
Caffeine metabolism	10	2	0.480370
Pantothenate and CoA biosynthesis	19	1	0.534600
Alanine, aspartate and glutamate metabolism	28	3	0.648530
alpha-Linolenic acid metabolism	13	1	0.648530
Tryptophan metabolism	41	4	0.780710
Ubiquinone and other terpenoid-quinone biosynthesis	9	1	0.649670
Tyrosine metabolism	42	1	0.649670
Glycine, serine and threonine metabolism	33	5	0.863790
Glyoxylate and dicarboxylate metabolism	32	3	0.803010
Pyrimidine metabolism	39	2	0.803010
Glycerophospholipid metabolism	36	1	0.788010
Butanoate metabolism	15	1	0.803010
Pentose phosphate pathway	22	1	0.803010
D-Glutamine and D-glutamate metabolism	6	1	0.824110
Nitrogen metabolism	6	1	0.824110
Lysine degradation	25	2	0.943180
Biotin metabolism	10	1	0.869120
Nicotinate and nicotinamide metabolism	15	1	0.943180
Histidine metabolism	16	1	0.959140
Galactose metabolism	27	1	0.959140
Glycerolipid metabolism	16	1	0.959140

## ROC curves of significant metabolites in PCOS patients and controls

In order to further distinguish PCOS from controls, ROC curves analysis was also conducted on these changed metabolites, the top 10 metabolites of highest AUC and their intensity in PCOS and controls were presented in Fig. 5. These metabolites were Leu-Ala/Ile-Ala (AUC = 1.00), 3-(4-Hydroxyphenyl)propionic acid (AUC = 0.998), Ile-Val/Leu-Val (AUC = 0.982), Gly-Val/Val-Gly (AUC = 0.982), aspartic acid (AUC = 0.968), DG(34:2)\_DG(16:0/18:2) (AUC = 0.951), DG(34:1)\_DG(16:0/18:1) (AUC = 0.938), Phe-Trp (AUC = 0.935), DG(36:1)\_DG(18:0/18:1) (AUC = 0.905), Leu-Leu/Leu-Ile (AUC = 0.904).

## Discussion

PCOS is a kind of common endocrine syndrome as well as metabolic disorder, which harms the reproductive system and the overall body metabolism of the patients [2]. In this study, we utilized metabolomics approach by UPLC-HRMS technology to study the metabolic changes of in 31 PCOS patients and 31 healthy controls. The metabolomics analysis showed that in PCOS patients serum, there were 146 significantly varied metabolites, among them 68 were downregulated, 78 were upregulated. These metabolites mainly belong to Triacylglycerols, Glycerophosphocholines, Acylcarnitines, Diacylglycerols, Peptides, Amino acids, Glycerophosphoethanolamines and FA. Pathway analysis showed that these metabolites were enriched in pathways including Glycerophospholipid metabolism, Fatty acid degradation, Fatty acid biosynthesis, Ether lipid metabolism etc. Diagnosis value assessment by ROC analysis showed that the changed metabolites including Leu-Ala/Ile-Ala, 3-(4-Hydroxyphenyl) propionic acid, Ile-Val/Leu-Val, Gly-Val/Val-Gly, aspartic acid, DG(34:2)\_DG(16:0/18:2), DG(34:1)\_DG(16:0/18:1), Phe-Trp, DG(36:1)\_DG(18:0/18:1), Leu-Leu/Leu-Ile had higher AUC values, indicated significant role in PCOS.

Genomic, proteomic, and metabolomic approaches have been introduced to various diseases research. Metabolomics enables to identify both the endogenous metabolites from downstream output of the genome and the exogenous metabolites from the upstream input from the environment, therefore allowing researchers to explore the nexus of gene–environment interactions and providing unique insights into the fundamental causes of disease [10,12]. To date, many metabolomic studies in PCOS have revealed the metabolic profiles and changes in PCOS patients under various conditions. In Zhang's study, they recruited 286 subjects and grouped into healthy women (C), those with hyperandrogenism (B1), those with IR but not obese (B2) and obese women with IR (B3) in PCOS in order to reveal metabolic profiling for women with hyperandrogenism and IR in PCOS [13]. The serum metabolomics analysis identified 59 differential metabolites, including palmitic acid, cholesterol, myo-inositol, D-allose, 1,5-anhydro-D-sorbitol, 1-monopalmitin, 1-monostearin, glycerol 1-phosphate, malic acid and citric acid, and these differential metabolites were related to biosynthesis of unsaturated fatty acids, and citrate cycle. Another study enrolled 10 PCOS patients and 10 healthy people and identified six biomarkers, L-Carnitine, LPE (22:5), Sphinganine, LPC (18:2), DHEAS and Glycocholic acid, these biomarkers belonged to metabolic pathway including lipid metabolism, carnitine metabolism, androgen metabolism, and bile acid metabolism [14]. Zhao's serum metabolomics study of PCOS found Significant increase in the levels of unsaturated free fatty acids, fatty acid amides, sulfated steroids, glycated amino acid and the decrease in levels of lysophosphatidylcholines, lysophosphatidylethanolamines, and free fatty acid (FFA) 18:1/FFA 18:0, FFA 20:3, dihydrotestosterone sulfate, glycated phenylalanine, and uridine were combined as a diagnostic biomarker for PCOS [15]. Murri et al reviewed the the metabolites and metabolic pathways associated with PCOS according to recent metabolomic studies, and reported that PCOS-associated metabolites were involved mostly in carbohydrate, fat, and protein metabolism [16]. These studies revealed the metabolic changes in PCOS patients, offered new insights into disease processes, and the discriminated metabolites may also help to provide reference for PCOS diagnosis and therapy. But limited sample sizes and the different metabolomic techniques used impose important limitations when aiming to integrate the results of the different studies conducted to date.

As well acknowledged that, PCOS is a kind of metabolic disorders, IR and thereby induced obesity are common symptoms in PCOS patients. Hence, lipid and lipoprotein metabolic abnormalities are accompanied with the PCOS progression [17]. In present study, for the 146 identified differential metabolites, there were 36 metabolites belong to Triacylglycerols, and the identified 36 metabolites were all upregulated in PCOS group. Triacylglycerols, also named as triglyceride (TG), are the main components of oils and fats. Overweight PCOS patients usually have lipid abnormalities, including higher level of serum TG. Cross sectional study on PCOS showed that subjects with PCOS demonstrated higher waist:hip ratio (WHR), T, TG, VLDL-cholesterol concentrations ( $P < 0.05$ ) [18]. Another study found that in PCOS women, there is a significant decrease TG following VitD supplementation for 8 weeks [19]. In addition, fatty acids were also included in the identified differential metabolites, the three FAs (Dihomo-alpha-linolenic acid, Myristoleic acid isomer 1, Myristoleic acid isomer 2) were all downregulated in PCOS group. But for the three FAs identified in this study, the previous publications about them in PCOS are limited. Dihomo-alpha-linolenic acid is a rare polyunsaturated fatty acid of the  $\omega$ -3 series, it accounts a small portion of the serum phospholipid fatty acids in normal humans but have significant role. PUFA could modulate hormonal and lipid profiles of the body, lowered TG and cholesterol levels, but patients with PCOS usually showe abnormal levels of PUFA metabolites. Study focused on differences in FA profiles of abdominal subcutaneous adipose tissue between pregnant women with and without PCOS found that total monounsaturated FA (MUFA) ( $p < 0.0004$ ) were higher in PCOS women than non-PCOS women, but total polyunsaturated FA (PUFA) was lower in PCOS than non-PCOS women ( $p < 0.004$ ) [20].  $\omega$ -3 PUFA supplementation has the positive modulation effect on the PCOS ovarian function, and potentiates the cellular development and steroid biosynthesis in PCOS [21]. Animal model study also showed that  $\omega$ -3 PUFA had effective role in improving lipid and hormonal profile, reducing blood glucose, weight gain and histopathological damages in PCOS rats [22]

As aforementioned, PCOS-associated metabolites were involved mostly in lipid and lipoprotein metabolic abnormalities. In present study, pathway analysis found that these identified differential metabolites were associated with the various pathways, especially including Glycerophospholipid metabolism, Sphingolipid metabolism, Phenylalanine metabolism, Ether lipid metabolism, Purine metabolism, Fatty acid degradation, Fatty acid biosynthesis, etc. Untargeted metabolomics approach on PCOS follicular fluid found significant abundance differences of glycerolipid, glycerophospholipids, sphingolipids, and carboxylic acids compared with healthy women, and these metabolism dysfunctions are contributed to decline the 2 pronuclei (PN) fertilization rate during IVF procedure [23]. Another LC-MS-based metabolomics showed that abnormalities of glycerophospholipid, glycerolipid and FA metabolisms were involved in the pathogenesis of PCOS and IR complications [24]. Amino acid metabolism is also a critical metabolism pathway of the body. In this study, except for the

identification of eight differential amino acids in PCOS, several related amino acid pathway were also identified, indicating the involvement of amino acid metabolism in PCOS., Zhao et al found that, in PCOS patients, the levels of phenylalanine, tyrosine and tryptophan are generally increased, and the ovulatory dysfunction of PCOS patients was associated with raised production of serine, threonine, phenylalanine, tyrosine and ornithine [25]. Fatty acid related pathways, including fatty acid degradation and biosynthesis were also found to be associated with the changed metabolites in PCOS of this study. And this was corresponded to the differential metabolites in PCOS compared to the healthy controls.

## Conclusion

In conclusion, in this study, we utilized metabolomics approach by UPLC-HRMS technology to study the metabolic changes of in 31 PCOS patients and 31 healthy controls. The metabolomics analysis showed that in PCOS patients serum, there were 146 significantly varied metabolites, among them 68 were downregulated, 78 were upregulated. These metabolites mainly belong to Triacylglycerols, Glycerophosphocholines, Acylcarnitines, Diacylglycerols, Peptides, Amino acids, Glycerophosphoethanolamines and FA. Pathway analysis showed that these metabolites were enriched in pathways including Glycerophospholipid metabolism, Fatty acid degradation, Fatty acid biosynthesis, Ether lipid metabolism etc. Diagnosis value assessment by ROC analysis showed that the changed metabolites including Leu-Ala/Ile-Ala, 3-(4-Hydroxyphenyl) propionic acid, Ile-Val/Leu-Val, Gly-Val/Val-Gly, aspartic acid, DG(34:2)\_DG(16:0/18:2), DG(34:1)\_DG(16:0/18:1), Phe-Trp, DG(36:1)\_DG(18:0/18:1), Leu-Leu/Leu-Ile had higher AUC values, indicated significant role in PCOS.

## Declarations

### Acknowledgement:

None.

### Funding:

None.

### Competing interest:

The authors declare that they have no conflict of interest.

## Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

### Conflict of publication:

Not applicable.

## Authors' contributions

Conception and design of the research: HX X Y Y

Acquisition of data: PL T RY Q

Analysis and interpretation of data: ZC Z

Statistical analysis: ZJ W

Drafting the manuscript: LC Z

Revision of manuscript for important intellectual content: ZJ W

All authors read and approved the final manuscript

## Ethics approval and consent to participate:

The study was approved by the Ethics Committee of Zhejiang Provincial Hospital of Chinese Medicine. [2020-KL-155-02]. All the participants signed informed consent forms before the start of the study. All protocols are carried out in accordance with relevant guidelines and regulations of "The Regulations of Ethical Reviews of Biomedical Research Involving Human Subjects".

## References

1. Empowering women with PCOS. *Lancet Diabetes Endocrinol.* 2019;10(10):737.
2. Escobar-Morreale HF. Polycystic ovary syndrome: definition, aetiology, diagnosis and treatment. *Nat Rev Endocrinol.* 2018;14(5):270–284.
3. Zawadzki J, Duniaf A. Diagnostic criteria for polycystic ovary syndrome: Towards a rational approach. Boston: Blackwell Scientific Publications. 1992.
4. Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum. Reprod.* 2004;19 (1):41–47.
5. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, et al. Positions statement: Criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: An androgen excess society guideline. *J Clin Endocrinol Metab.* 2006;91:4237–45.
6. Cobin RH. Cardiovascular and metabolic risks associated with PCOS. *Intern Emerg Med.* 2013 Apr;8 Suppl 1:S61-4.
7. Kakoly NS, Moran LJ, Teede HJ, Joham AE. Cardiometabolic risks in PCOS: a review of the current state of knowledge. *Expert Rev Endocrinol Metab.* 2019;14(1):23–33.
8. Peigné M, Dewailly D. Long term complications of polycystic ovary syndrome (PCOS). *Ann Endocrinol (Paris).* 2014 Sep;75(4):194–9.
9. Chang AY, Lalia AZ, Jenkins GD, et al. Combining a Nontargeted and Targeted Metabolomics Approach to Identify Metabolic Pathways Significantly Altered in Polycystic Ovary Syndrome. *Metabolism.* 2017;71:52–63.
10. Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov.* 2016;15(7):473–84.
11. Buszewska-Forajta M, Rachoń D, Stefaniak A, et al. Identification of the metabolic fingerprints in women with polycystic ovary syndrome using the multiplatform metabolomics technique. *J Steroid Biochem Mol Biol.* 2019;186:176–184.
12. Houten SM. Metabolomics: unraveling the chemical individuality of common human diseases. *Ann Med.* 2009;41:402–407.
13. Zhang Z, Hong Y, Chen M, et al. Serum metabolomics reveals metabolic profiling for women with hyperandrogenism and insulin resistance in polycystic ovary syndrome. *Metabolomics.* 2020;16(2):20.
14. Jia C, Xu H, Xu Y, Shi Q. Serum metabolomics analysis of patients with polycystic ovary syndrome by mass spectrometry. *Mol Reprod Dev.* 2019;86(3):292–297.
15. Zhao X, Xu F, Qi B, et al. Serum Metabolomics Study of Polycystic Ovary Syndrome Based on Liquid Chromatography – Mass Spectrometry. *J Proteome Res.* 2014;13(2):1101–11.
16. Murri M, Insenser M, Escobar-Morreale HF. Metabolomics in polycystic ovary syndrome. *Clin Chim Acta.* 2014;429:181–8.
17. Diamanti-Kandarakis E, Papavassiliou AG, Kandarakis SA, Chrousos GP. Pathophysiology and types of dyslipidemia in PCOS. *Trends Endocrinol Metab.* 2007;18(7):280–5.
18. Pirwany IR, Fleming R, Greer IA, Packard CJ Sattar N. Lipids and lipoprotein subfractions in women with PCOS: relationship to metabolic and endocrine parameters. *Clin Endocrinol (Oxf).* 2001 Apr;54(4):447–53.
19. Irani M, Seifer DB, Grazi RV, Irani S, Rosenwaks Z, Tal R. Vitamin D Decreases Serum VEGF Correlating with Clinical Improvement in Vitamin D-Deficient Women with PCOS: A Randomized Placebo-Controlled Trial. *Nutrients.* 2017;9(4).
20. Emami N, Alizadeh A, Moini A, Yaghmaei P, Shahhosseini M. Differences in fatty acid profiles and desaturation indices of abdominal subcutaneous adipose tissue between pregnant women with and without PCOS. *Adipocyte.* 2020;12;9(1):16–23.
21. Hu X, Weng X, Tian Y, et al. Effects of omega-3 polyunsaturated fatty acids on steroidogenesis and cellular development in PCOS rats. *Food Funct.* 2019 May;22;10(5):2504–2514.
22. Komal F, Khan MK, Imran M, et al. Impact of different omega-3 fatty acid sources on lipid, hormonal, blood glucose, weight gain and histopathological damages profile in PCOS rat model. *J Transl Med.* 2020;14;18(1):349.
23. Liu L, Yin TL, Chen Y, et al. Follicular dynamics of glycerophospholipid and sphingolipid metabolisms in polycystic ovary syndrome patients. *J Steroid Biochem Mol Biol.* 2019;185:142–149.
24. Chen YX, Zhang XJ, Huang J, et al. UHPLC/Q-TOFMS-based plasma metabolomics of polycystic ovary syndrome patients with and without insulin resistance. *J Pharm Biomed Anal.* 2016 Mar;121:141–150.
25. Zhao Y, Fu L, Li R, et al. Metabolic profiles characterizing different phenotypes of polycystic ovary syndrome: plasma metabolomics analysis. *BMC Med.* 2012 Nov;30;10:153.

## Figures

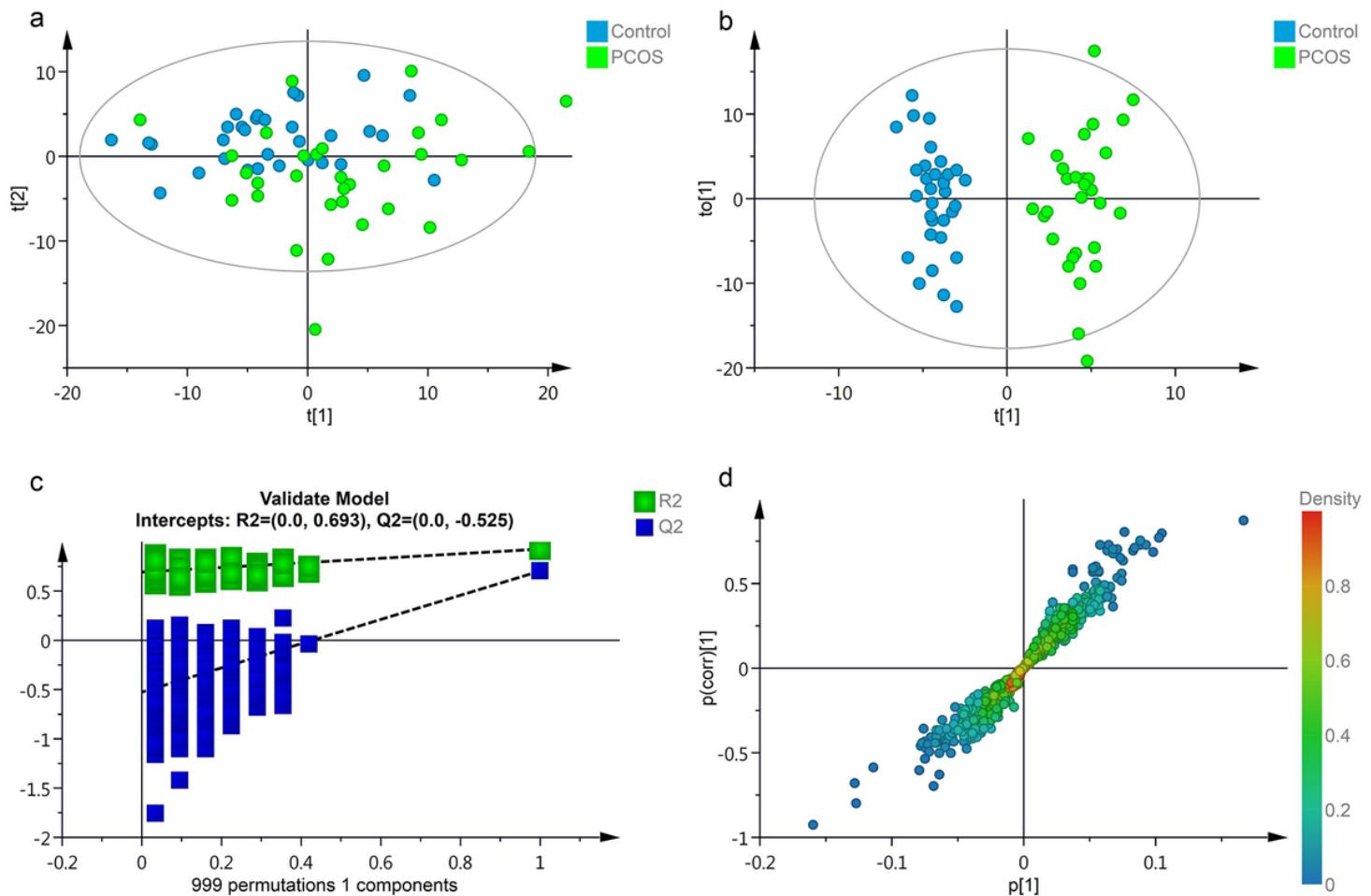


Figure 1

Multivariate statistical analysis of the serum metabolites in the study subjects. a: unsupervised PCA score plots of metabolic phenotypes between PCOS and CON groups. Metabolomics data was log transformed and scaled to unit variance for modeling. Model parameter:  $R^2_X=0.86$  (cumulative variance proportion of 9 principal components). b: Score plot of OPLS-DA modeling to maximize inter-group differentiation of metabolomic data between PCOS and CON groups. Model parameter: 1 predictive component + 2 orthogonal component,  $R^2_Y=0.93$ ,  $Q^2=0.70$ . c: 999 times permutation test result of OPLS-DA modeling. d: S-plot of OPLS-DA modeling.

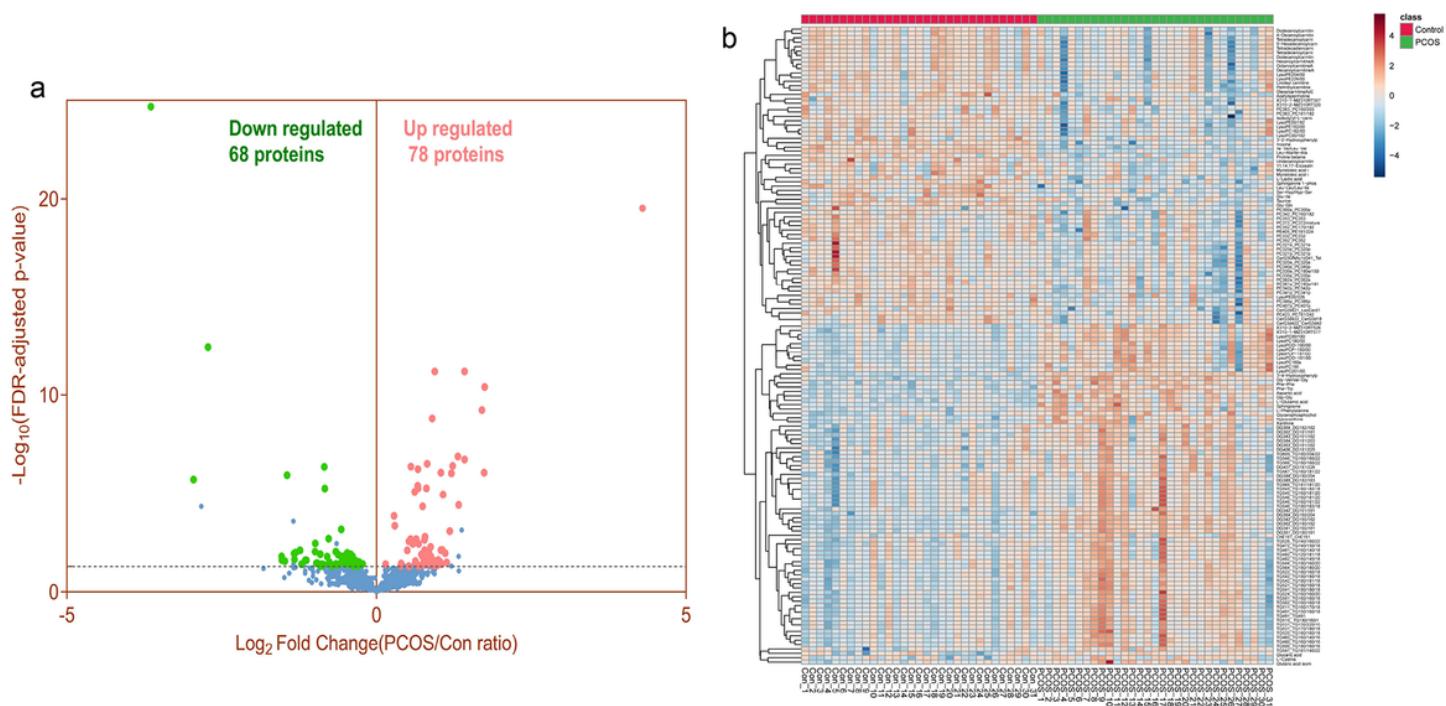


Figure 2

Identification of significant metabolites by UPLC-HRMS. a: Volcano plot to visualize differential metabolites of significance between PCOS and CON group. Metabolites with FDR adjusted p-value  $\leq 0.05$  were highlighted with red (up-regulated) and green color (down-regulated), respectively. b: Clustering analysis result using differential metabolites of significance in Student's t-test analysis (FDR adjusted p-value  $< 0.05$ ).

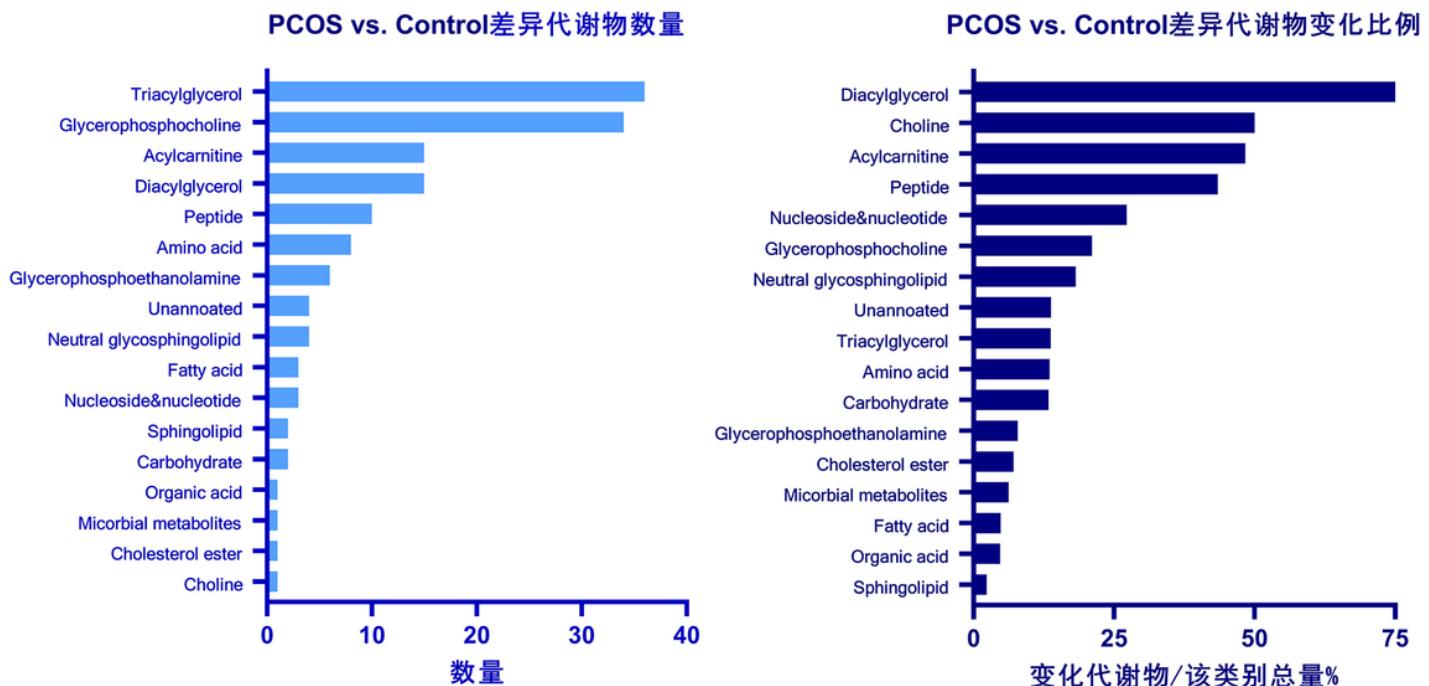
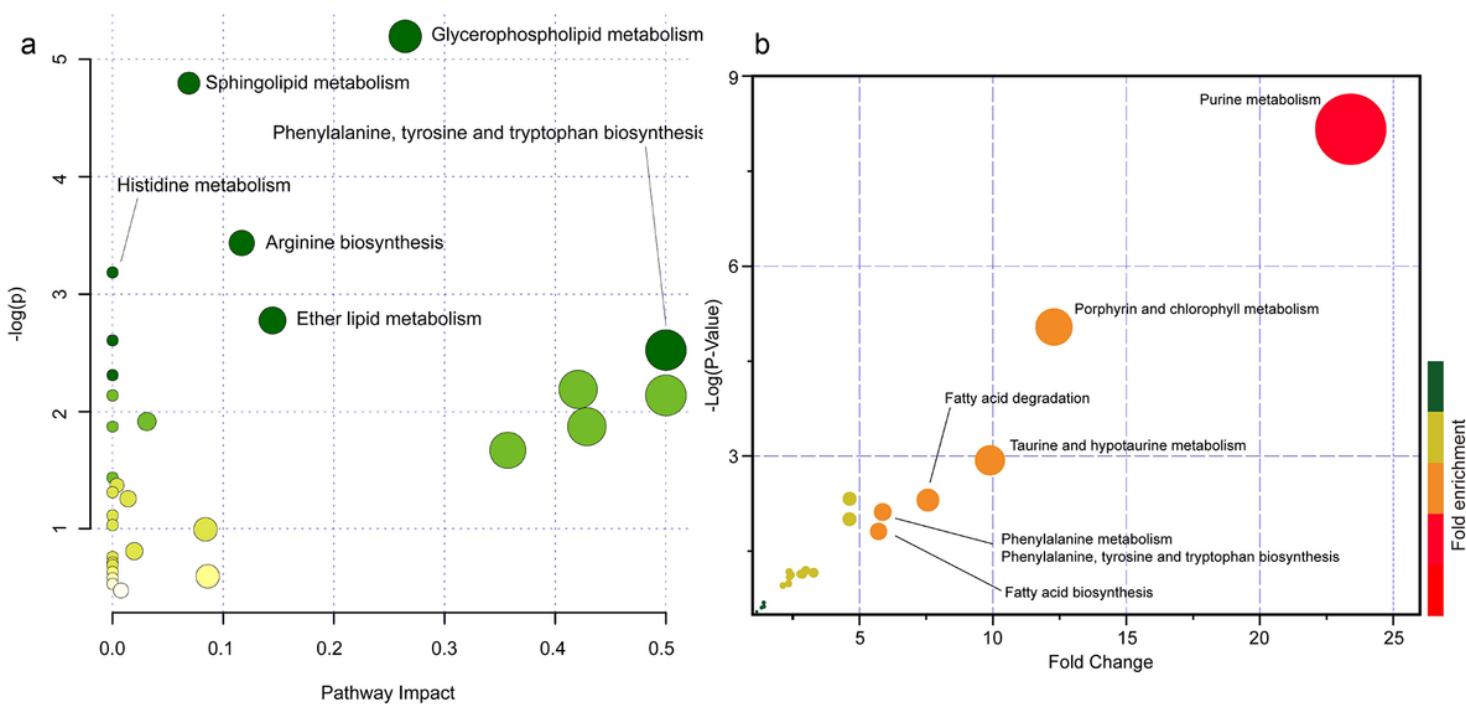


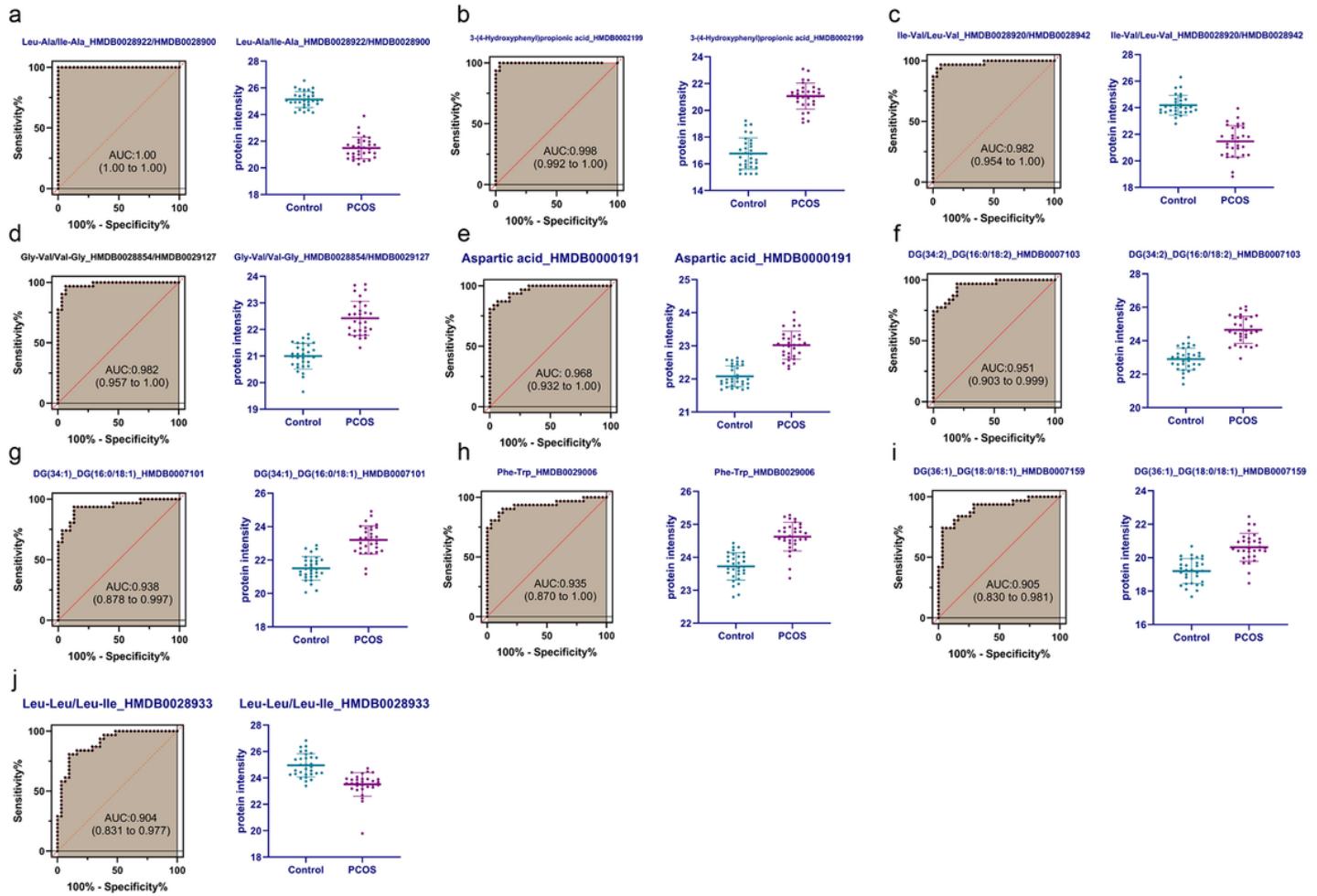
Figure 3

Chemical structure classification of differential metabolites between PCOS and control group.



**Figure 4**

Pathway analysis of the differential metabolites between PCOS vs.CON group. a: Pathway analysis result of differential metabolites between PCOS vs. control group using over-representation method in MetaboAnalyst website ( $p\text{value}<0.05$  of t-test after FDR adjusting). Hypergeometric test and Relativebetweenness centrality algorithm was used for pathway topology analysis, human KEGG pathway library was used. b: Metabolites set enrichment analysis of all metabolites with HMBD identifier using quantitative enrichment analysis method. Pathway-associated metabolite sets (KEGG) containing 84 metabolite sets based on normal human metabolic pathways was used for this MSEA.



**Figure 5**

ROC-curve analysis of top 10 metabolites ranked AUC values.