# Mass Spectrometry-Based Identification of New Serum Biomarkers in Patients with Latent Infection Pulmonary Tuberculosis

**Yan-Xia Li**
Foshan Forth People's Hospital
**Kang-Di Zheng**
Foshan Longsee Biotechnology Co, Ltd
**Yu Duan**
Foshan Forth People's Hospital
**Hua-Juan Liu**
Foshan Forth People's Hospital
**Yu-Qun Tang**
Foshan Forth People's Hospital
**Jun Wu**
Affiliated Hospital of Guangdong Medical University
**Dong-Zi Lin**
Foshan Forth People's Hospital
**Zhao Zhang** ( zhangzhao0402@163.com )
Peptide and Protein Research and Application Key Laboratory of Guangdong Medical University

**Research Article**

**Keywords:** pulmonary tuberculosis, latent infection, serum biomarker, LC-MS/MS

**Posted Date:** December 20th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-1118703/v1

# Abstract

**Background:** To screen specific metabolic markers serum metabolic biomarkers which can achieve the main monitoring indicators to evaluate the development from latent infection to active tuberculosis infection, and analysis its underlying mechanisms and functions.

**Methods**: Four groups of serum, including healthy control, latent infection, drug sensitivity (DS), and drug resistant tuberculosis, were collected. The metabolites in all serum samples were extracted by oscillatory, deproteinization, and then were detected by LC-MS/MS analysis. Normalization by Pareto-scaling method, the difference analysis was carried out by Metaboanalyst 4.0 software, one-way ANOVA analysis among groups showed that *p*-value ≤0.05 was regarded as a different metabolite. To clarify the dynamic changes and functions of differential metabolites with disease progression, and explore its significance and mechanism as a marker by further cluster analysis, functional enrichment analysis, and relative content change analysis of differential metabolites.

**Results**: There were 565 significantly different metabolites in four groups. Differential metabolites, including Indole-3-acetaldehyde, Theophylline, Inosine and Prostaglandin H2, *etc.*, may be the key serum biomarkers to diagnose the period of latent infection of *Mycobacterium tuberculosis* (*M. tuberculosis*). which was closely related to Amino acid metabolism, Biosynthesis of other secondary metabolites, Nucleotide metabolism, Endocrine system, Immune system, Lipid metabolism, and Nervous system.

**Conclusion:** Indole-3-acetaldehyde, Theophylline, Inosine, and Prostaglandin H2, the 4 metabolites may be potential markers diagnosing the period of latent infection of *M. tuberculosis*. Meanwhile, Inosine and Prostaglandin E1 can become potential biomarkers for the diagnosis of latent infection, and Theophylline and Cotinine 1 can be used as potential markers to monitor disease progression, which established strategy provided promising clinical application prospects for the development of disease assessment by combining small molecule metabolic markers to improve the sensitivity and specificity of disease diagnosis.


# Background

Tuberculosis (TB) is mainly caused by *Mycobacterium tuberculosis* (*M. tuberculosis*) [1]. There are 9.4 million people diagnosed with TB every year in the world, and more than 1.3 million people die of it. TB has become a major health issue of global concern. Since the culture of *M. tuberculosis* is tedious and time-consuming, and the number of people with tuberculosis latent infection is very large. At present, accurate and rapid diagnosis of latent infection is a major limiting factor in TB control. Therefore, it is of great significance for rapid diagnostic methods with high specificity, sensitivity, and low cost.

Small molecule metabolites are the end products of the cell regulation process. The changes of their types and quantities are regarded as the final response of organisms to a gene or environmental changes [2, 3]. GC-MS, LC-MS, NMR, and other metabolomics detection technologies with high flux and high sensitivity are the mainstream methods for metabolite detection [4]. So far, metabonomics technology

has been widely used in clinical research, such as the identification of new biomarkers and how biomarkers help to improve the discovery and diagnosis of diseases [5, 6, 7].

The discovery of TB-related metabolites has also brought new methods for the diagnosis and treatment of tuberculosis. Susanna KP Lau *et al.* [8] reported that the content of 24 metabolites in the culture supernatant of *M. tuberculosis* was significantly higher than that of non-*M. tuberculosis* group, of which 4 metabolites were identified as 1-tuberculin adenosine derivatives, which may be used as a new marker of metabolism of *M. tuberculosis*. Yeware A *et al* [9] focused on the mechanism of rapid transition of *M. tuberculosis* from the active state to the viable but non-culturable state caused by ammonium diphenylene alkylate treatment through metabolomics and compared it with dormant phenotype.

The above studies indicated that new metabolic markers in *M. tuberculosis*, which may be related to the progression of tuberculosis. Simultaneously, the concentration of serum metabolites is a direct reading of human biological processes, which is related to cardiovascular and metabolic diseases. Therefore, it is necessary to study the serum metabolites of tuberculosis patients and the progress of tuberculosis. Weiner J *et al.* [10] studied the metabolic patterns of tuberculosis patients (TB+), healthy uninfected persons (TST-), and latent infection (TST+) individuals. The difference in serum small molecule abundance among the three groups was determined by comparing three groups. There were significant differences in 176 compounds between TB patients and two healthy groups (TST- and TST+). These researches show that the dynamic changes of metabolomics in TB patients during the treatment process are systematically analyzed, and the serum biomarkers of new latent infection and active tuberculosis are screened, and the main molecular mechanisms are explored. To aim to screen specific metabolic markers to achieve the main monitoring indicators to evaluate the development from latent infection to active tuberculosis infection, it is expected to formulate personalized treatment plans for patients.

## Methods

# Subjects and sample collection

The Blood samples were collected from the patients undergoing pulmonary tuberculosis (TB) screening in Foshan Forth People's Hospital and Dongguan Sixth People's Hospital from October 2017 to March 2018. The criteria for enrollment of TB patients included the following: 1) *Mycobacterium tuberculosis* (Mtb) positive for sputum smear and sputum culture. 2) HIV test was negative. 3) The TB patients had never been treated with any anti-tuberculosis (anti-TB) drug. 4) Patients agreed to participate in the study and signed informed consent. The exclusion criteria included the following: patients with negative pnitrobenzoic acid (PNB) test. According to the results of sputum smear and sputum culture, the collected samples were divided into groups of healthy control (HC) with 148 (77.08%) cases, and latent infection (LI) with 44 (22.92%) cases. After collecting the blood samples, the serum was collected by centrifugation and stored at -20 ℃ for LC-MS / MS detection. This study was approved by the ethics committees of the Dongguan Key Laboratory of Medical Bioactive Molecular Developmental and Translational Research, Foshan Fourth People's Hospital, and Dongguan Sixth People's Hospital.

# Preconditioning of plasma

100 μL of plasma sample was mixed 600 μL of methanol (containing 1 mM butylhydroxytoluene) for 5 min ultrasonic processing. Then, add 1.8 mL of Methyl Tertbutyl Ether and oscillate at room temperature for 1 h. Add 500 μL of water and incubate for 10 min at room temperature with oscillating without interruption. After 2-min still standing, samples were centrifuged (12,000 rpm) for 10 mins. To pipette 600 μL of the upper layer lipid and 300 μL of the lower layer (2:1) aqueous layer and transfer to a new EP tube, and dissolved with 200 μL of acetonitrile and water (1:1) after vacuum dry, then centrifuged to take the supernatant, using mixed layer and aqueous layer metabolite conditions.

# Metabolic profiling detection by LC-MS

The experiment used ACE (Aberdeen, Scotland) Excel2C-18PFP (100×2.1 mm, 2 μm) chromatographic column and C18 guard column. Mobile phase A contained aqua with 0.1% formic acid and mobile phase B contained acetonitrile with 0.1% formic acid. The chromatographic gradient started from 2% B for 1 min and entered 98% B in 10 mins. This had kept for 2 min and then it dropped to 2% within 30 s before reaching equilibrium for 3 min. The injection volume was 2 μL and the column was maintained at 35°C. All samples were injected twice. The mass spectrometer was operated in heated electrospray ionization (HESI) mode with a spray voltage of 3.5 kV, a capillary temperature of 300°C, a sheath gas flow of 50, an assist gas of 10.S lens RF level set at 40, and an S lens voltage set at 25 V. The collecting resolution ratio of the full scan was 70,000 and the MS/MS collecting list was 17,500 (New serum biomarker identification and analysis by mass spectrometry in cervical precancerous lesion and acute cervicitis in South China).

# The identification and screening of the metabolites

The results were normalized by the Pareto-scaling method, then the differences analysis and enrichment analysis were performed simultaneously by using MetaboAnalyst 4.0 software. This study applied the Fold Change Analysis (FC Analysis), univariate statistical analysis of metabolites by one-way ANOVA (The one-way ANOVA was used for comparison over two groups), and the drawing of variable importance in the projection (VIP) plot. The differential data were analyzed by Metlin_AMRT_PCDL and Metlin_Lipids_AM_PCDL database to identify the corresponding metabolites. Then using the metabolites with significant differences to conduct Hierarchical Clustering Analysis of the samples among different groups to predict the potential serum biomarkers for diagnosis of latent infection and active tuberculosis.

# Statistical analysis

To analyze differential metabolites by MetaboAnalyst Enrichment Analysis, and analyze the related metabolic pathways by MetaboAnalyst Pathway Analysis. The data were analyzed by SPSS20.0 software. The results of the normal distribution of the measurement data were expressed as mean ± standard deviation (x ± s). The t-test was used for comparison between the two groups. One-way ANOVA analysis of variance was used for comparison over two groups. $p < 0.05$ was considered to be statistically significant.

# Results

## Clinical information of the subjects

This study included 148 (77.08%) healthy control (HC) and 44 (22.92%) latent infection (LI) group, and clinical information of two groups were summarized in Table 1.

**Table 1**

The clinical information between healthy control and latent infection group (x ± s).

| Indicators | healthy control | latent infection |
|---|---|---|
| Number/n | 148 | 44 |
| Age（median±SD, range） | 37.65 ± 8.84（18-60） | 42.34 ± 9.11（22-57） |
| Male | 78 (52.70%) | 29 (65.91%) |
| Female | 70 (47.30%) | 15 (34.09%) |
| γ Interferon / basic level N | 0.10 ± 0.34 | 0.27 ± 0.75 |
| γ Interferon / actual release level（T-N） | -0.02 ± 0.22 | 10.46 ± 14.30 |
| γ Interferon / positive actual level（M-N） | 40.63 ± 7.12 | 36.86 ± 12.51 |
| Leukocyte count | 6.81 ± 1.58 | 6.91 ± 1.65 |
| Lymphocyte ratio (%) | 31.85 ± 6.47 | 30.45 ± 6.91 |
| Monocyte ratio (%) | 7.44 ± 2.04 | 7.69 ± 1.80 |

## Analysis of differential metabolite distribution

The distribution of serum metabolites in healthy control and latent infection group were analyzed by LC-MS. Before the analysis by SMICA-P13.0 software, the data was normalized to ensure more intuitive and reliable results. Principal Component Analysis (PCA) showed that there was no significant difference between healthy control and latent infected groups (Fig. 1A). Partial Least Squares Discriminant Analysis (OPLS-DA) could distinguish the latent infected group from the healthy control group, suggesting that there were differences in metabolite components between the two groups (Fig. 1B). Through metabolomics technology, a series of serum metabolic markers of drug-resistant and drug-sensitive groups were screened in the early stage. Through PCA and OPLS-DA analysis (Fig. 1A and 1B), drug-resistant and drug-sensitive groups could be significantly distinguished from healthy control and latent patients groups, suggesting that drug treatment may significantly affect the composition of serum metabolites.

# Screening differential metabolites in different groups

The serum metabolites of patients of health control, latent infection, drug sensitivity, and drug resistance group were statistically tested by one-way ANOVA test. The results showed that there were 565 significantly different metabolites in the four groups (Table 2).

**Table 2**

Screening results of differential metabolites among groups (partial results)

| Metabolite | Mean | | | | *p* |
|---|---|---|---|---|---|
| | Heal | Lat | Res | Sen | |
| Betaine | 3.94 | 4.03 | 6.61 | 6.91 | 1.38E-25 |
| 4-Hydroxybenzaldehyde | -1.39 | -1.57 | -0.66 | -0.23 | 2.79E-14 |
| (R)-(+)-2-Pyrrolidone-5-carboxylic acid | 2.89 | 3.07 | -0.66 | -0.23 | 3.31E-25 |
| PHENACYLAMINE | 0.14 | 0.00 | -0.66 | -0.23 | 6.29E-12 |
| Hypoxanthine | 0.30 | 0.11 | 0.51 | 0.59 | 0.00378 |
| 4-formyl Indole | -1.80 | -1.94 | -0.66 | -0.23 | 1.10E-19 |
| Coumarin | -2.54 | -2.82 | -0.66 | -0.23 | 1.59E-24 |
| L-Lysine | -2.29 | -2.65 | -0.66 | -0.23 | 1.12E-23 |
| 2-Hydroxycinnamic acid | 0.14 | 0.06 | -0.66 | -0.23 | 3.75E-13 |
| Val Gly | -1.62 | -2.02 | -0.66 | -0.23 | 2.55E-17 |
| Theophylline | -2.76 | -2.57 | -0.66 | -0.23 | 1.29E-18 |
| butamben | 1.33 | 1.21 | -0.66 | -0.23 | 5.09E-26 |
| p-CHLOROPHENYLALANINE | 0.78 | 0.87 | -0.66 | -0.23 | 2.12E-25 |
| Acetylcarnitine | 0.55 | 0.57 | 3.68 | 3.74 | 4.50E-24 |
| DL-Tryptophan | 3.05 | 3.04 | -0.66 | -0.23 | 2.79E-25 |
| Pro Leu | -1.78 | -1.87 | -0.66 | -0.23 | 7.77E-19 |
| Inosine | -3.95 | -6.22 | -0.66 | -0.23 | 1.03E-24 |
| Phe Phe | 1.32 | 1.01 | -0.66 | -0.23 | 4.11E-23 |
| Leu Leu Phe | -1.24 | -1.36 | -0.66 | -0.23 | 2.80E- |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | 11 |
| Palmitoyl-L-carnitine | -1.45 | -1.59 | -0.66 | -0.23 | 1.27E-14 |
| 11β-PGF2α Ethanolamide | -0.11 | -0.21 | -0.66 | -0.23 | 5.84E-09 |
| PC(16:0/0:0)[U] / PC(16:0/0:0)[rac] | 6.46 | 6.38 | -0.66 | -0.23 | 1.13E-20 |
| 1-heptadecanoyl-sn-glycero-3-phosphocholine | -0.14 | -0.36 | -0.66 | -0.23 | 9.57E-05 |
| Tyr Arg Leu Ile Val | -1.25 | -3.35 | -0.66 | -0.23 | 6.60E-19 |
| w/o MS2:δ-Valerolactam | -0.15 | 0.29 | -0.66 | -0.23 | 0.00847 |
| w/o MS2:Hydroxyhydroquinone | -1.98 | -1.09 | -0.66 | -0.23 | 6.42E-15 |
| w/o MS2:CYCLOCREATINE | -2.69 | -3.05 | -0.66 | -0.23 | 1.45E-24 |
| w/o MS2:2-Aminopropiophenone | -2.32 | -2.29 | -0.66 | -0.23 | 2.91E-23 |
| w/o MS2:Pyroglutamic acid | -1.69 | -1.69 | -0.29 | 0.51 | 3.22E-18 |
| w/o MS2:N-HYDROXYMETHYLNICOTINAMIDE | -2.07 | -2.03 | -0.66 | -0.23 | 3.76E-20 |
| w/o MS2:Mechlorethamine | -2.30 | -2.49 | -0.66 | -0.23 | 1.62E-24 |
| w/o MS2:1-Benzylimidazole | -2.06 | -2.34 | -0.66 | -0.23 | 1.50E-23 |
| w/o MS2:3-[Bis(2-hydroxyethyl)amino]propanenitrile | -2.57 | -2.53 | -0.66 | -0.23 | 5.18E-24 |
| w/o MS2:Indoleacetaldehyde | -4.20 | -5.67 | -0.66 | -0.23 | 1.24E-24 |
| w/o MS2:3-thio-Pheneacrylic Acid methyl ester | 1.61 | 1.67 | -0.66 | -0.23 | 5.11E-26 |
| w/o MS2:Cotinine | -2.49 | -2.23 | 2.55 | 2.85 | 2.29E-21 |
| w/o MS2:3-Methylethcathinone | -4.17 | -3.67 | -0.66 | -0.23 | 5.97E-25 |
| w/o MS2:Thiabendazole | -0.84 | -0.83 | -0.66 | -0.23 | 5.37E-10 |
| w/o MS2:Kynurenine | -2.26 | -2.66 | -0.66 | -0.23 | 3.17E- |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | 24 |
| w/o MS2:PROPOXUR | -4.33 | -6.90 | -0.66 | -0.23 | 1.42E-24 |
| w/o MS2:Tuckolide; Decarestrictine D | -2.31 | -2.78 | -0.66 | -0.23 | 2.19E-23 |
| w/o MS2:5-Hydroxy-4-[3-(2-hydroxy-2-propanyl)-2-oxiranyl]-1-methyl-7-oxabicyclo[4.1.0]hept-3-en-2-one | -1.45 | -1.69 | -0.66 | -0.23 | 1.33E-14 |
| w/o MS2:Melatonin | -2.66 | -3.27 | -0.66 | -0.23 | 3.70E-26 |
| w/o MS2:3,4,5-Trimethoxycinnamic acid | -2.35 | -2.94 | -0.66 | -0.23 | 2.13E-22 |
| w/o MS2:Ribothymidine | -0.40 | -0.38 | -0.66 | -0.23 | 0.00216 |
| w/o MS2:Phe Ile | -3.01 | -3.80 | -0.66 | 1.04 | 8.36E-25 |
| w/o MS2:Phe Asp | -2.82 | -3.58 | -0.66 | -0.23 | 1.14E-24 |
| w/o MS2:1-Methyladenosine | -2.37 | -2.78 | -0.66 | -0.23 | 3.16E-25 |
| w/o MS2:AG-17 | -1.65 | -3.18 | -0.66 | -0.23 | 1.71E-23 |
| w/o MS2:(±)-Octanoylcarnitine | -2.73 | -1.95 | -0.66 | -0.23 | 1.20E-20 |
| w/o MS2:4-(1-Acetyloxypropen-2-yl-)-2-methoxyphenylisobutyrat; 4-(1-Acetoxy-2-propen-1-yl)-2-methoxyphenyl 2-methylpropanoate | -3.43 | -4.72 | -0.66 | -0.23 | 2.20E-25 |
| w/o MS2:Ser Lys Ser | -0.38 | -0.54 | -0.66 | -0.23 | 3.23E-06 |
| w/o MS2:PGH2 | -3.93 | -6.90 | -0.66 | -0.23 | 3.52E-24 |
| w/o MS2:Trimethylolpropane trimethacrylate | -1.68 | 0.00 | -0.66 | -0.23 | 1.39E-10 |
| w/o MS2:Lisuride | -4.18 | -5.54 | -0.66 | -0.23 | 1.41E-24 |
| w/o MS2:N-Oleoyl Glycine | -1.60 | -1.46 | 0.84 | -0.23 | 9.58E-18 |
| w/o MS2:Ala Met Lys | -2.63 | -6.90 | -0.66 | -0.23 | 2.14E-23 |
| w/o MS2:Spiromesifen | -3.80 | -5.71 | -0.66 | -0.23 | 1.01E-24 |

| | | | | | |
|---|---|---|---|---|---|
| w/o MS2:13,14-dihydro-19(R)-hydroxyPGE1 | -1.00 | -1.90 | -0.66 | -0.23 | 4.89E-15 |
| w/o MS2:16,16-dimethyl-6-keto Prostaglandin E1 | -4.05 | -6.90 | -0.66 | -0.23 | 2.59E-24 |
| w/o MS2:Pro Arg Ile | -0.46 | -1.20 | -0.66 | -0.23 | 3.97E-08 |
| w/o MS2:Asn Phe Ile | -2.50 | -6.90 | -0.66 | -0.23 | 2.14E-23 |
| w/o MS2:Arg Gly Tyr | -3.70 | -4.83 | -0.66 | -0.23 | 1.27E-24 |
| Unknown | 1.35 | 0.50 | -0.66 | -0.23 | 2.76E-19 |
| w/o MS2:Myriocin | -1.01 | -1.78 | -0.66 | -0.23 | 1.25E-14 |
| w/o MS2:His Lys Met | -4.37 | -6.08 | -0.66 | -0.23 | 1.46E-24 |
| w/o MS2:Ala Asn Val Asp | -3.38 | -4.23 | -0.66 | -0.19 | 1.17E-24 |
| w/o MS2:Phe Leu Arg | -0.58 | -1.07 | -0.66 | -0.23 | 1.44E-08 |
| w/o MS2:Arg Met Met | -3.93 | -5.10 | -0.66 | -0.23 | 7.15E-24 |
| w/o MS2:HC Toxin | -2.40 | -6.90 | -0.66 | -0.23 | 2.14E-23 |
| w/o MS2:PE(16:0/0:0) | -1.33 | -1.75 | -0.66 | -0.23 | 1.64E-16 |
| w/o MS2:Pristimerin | -0.73 | -0.63 | -0.66 | -0.23 | 2.85E-10 |
| w/o MS2:Buprenorphine | -0.22 | -0.43 | -0.66 | -0.23 | 0.00079 |
| w/o MS2:(3E)-7-Hydroxy-3,7-dimethyl-3-octen-1-yl 6-O-(6-deoxy-?-L-mannopyranosyl)-?-D-glucopyranoside | -2.51 | -6.90 | -0.66 | -0.23 | 2.14E-23 |
| w/o MS2:PG(18:1(9Z)/0:0) | -3.30 | -5.33 | -0.66 | -0.23 | 7.09E-25 |
| w/o MS2:Arg Thr Asp Arg | -2.93 | -3.16 | -0.66 | -0.23 | 1.38E-24 |
| w/o MS2:Phe Glu Ser Phe Gly | -4.37 | -4.19 | -0.66 | -0.23 | 1.83E-25 |
| w/o MS2:Leu Leu Asp Leu Leu | 3.21 | 3.38 | -0.66 | -0.23 | 2.35E-26 |

| | | | | | |
|---|---|---|---|---|---|
| w/o MS2:Arjunglucoside II; (2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yl (4aS,6aS,6bR,9R,10R,11R,12aR)-10,11-dihydroxy-9-(hydroxymethyl)-2,2,6a,6b,9,12a-hexamethyl-1,3,4,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-octadecahydropicene-4a(2H)-carboxylate | -3.52 | -4.46 | -0.66 | -0.23 | 2.35E-23 |

Heal means healthy control, Lat means latent infection, Res means resistance, and Sen means sensitivity group.

To further screen the differential metabolites among different groups, variable importance in the project (VIP) was used to screen the key variables in the grouping based on OPLS-DA analysis. As shown in Fig. 2, the VIP scores of pos_879 (Cotinine) compound was the largest, which was a significant difference metabolite in the four groups (Table 2), and was up-regulated in the drug sensitive group and down-regulated in the latent infection group. Pos_6484 compounds increased in the healthy control group, followed by the latent infection, drug sensitivity, and drug resistance group. pos_ 8430□unknown□□pos_ 2688□unknown□□pos_ 1509□Phencyclidine□□pos_2594□Ranitidine□□pos_4682□unknown□□pos_297 □unknown□□pos_874□unknown□□pos_10662□unknown□□pos_1154□C10H16N4O; PlaSMA ID-740□□pos_ 3932 (NCGC00384560-01_C16H24O9_beta-D-Glucopyranoside, 3-hydroxy-2-(4-hydroxy-3-methoxyphenyl)propyl□□pos_3048□13E-Docosenamide□□pos_3416□unknown□□pos_3563 (unknown) and other 13 compounds were up-regulated in drug sensitive and drug resistant group and down-regulated in latent infection and healthy control group. They can be used as potential markers to distinguish drug sensitive patients, including pos_3048 (13e docosenamide) was a significantly different metabolite in four groups (Table 2).

# Cluster analysis of serum metabolic biomarkers

The cluster analysis of the differential metabolites of patients in groups of healthy control, latent infection, drug resistance patients, drug sensitivity group (Fig. 3). We found that in the healthy control and latent infection groups, pos_6484□PC(16:0/0:0)[U] / PC(16:0/0:0)[rac]□, pos_1123□DL-Tryptophan□, pos_8385□Leu Leu Asp Leu Leu□were up-regulated, and pos_2176□(±)-Octanoylcarnitine□, pos_2011□Phe Ile□, pos_906□Theophylline□, pos_879□Cotinine□, pos_4727□Ala Asn Val Asp□0, pos_8368 □Phe Glu Ser Phe Gly□, pos_2488□4-(1-Acetyloxypropen-2-yl-)-2-methoxyphenylisobutyrat, 4-(1-Acetoxy-2-propen-1-yl)-2-methoxyphenyl 2-methylpropanoate□, pos_5241□Arg Met Met□, pos_3272□Ala Met Lys□, pos_5257□HC Toxin□, pos_4158□Asn Phe Ile□, pos_6142□(3E)-7-Hydroxy-3,7-dimethyl-3-octen-1-yl 6-O-(6-deoxy-?-L-mannopyranosyl)-?-D-glucopyranoside□, pos_2962□PGH2□, pos_3878□16,16-dimethyl-6-keto Prostaglandin E1□, pos_1154□PROPOXUR□, pos_4682□His Lys Met□, pos_1800□Inosine□, pos_772 □Indoleacetaldehyde□, pos_7613□Arg Thr Asp Arg□, pos_1010□3-Methylethcathinone□, pos_4201□Arg Gly Tyr□, pos_9608□Arjunglucoside II; (2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yl (4aS,6aS,6bR,9R,10R,11R,12aR)-10,11-dihydroxy-9-(hydroxymethyl)-2,2,6a,6b,9,12a-hexamethyl-

1,3,4,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-octadecahydropicene-4a(2H)-carboxylate, pos_6824 PG(18:1(9Z)/0:0), pos_3043Lisuride, pos_3744Spiromesifenwere down-regulated.

# Metabolic Pathway Analysis

KEGG (Kyoto Encyclopedia of genes and genes) annotation and enrichment analysis were performed on the differential metabolites among four groups with *p* adjust < 0.05. The results are shown in Fig. 4 and Table 3.

The significantly enriched pathway included the Oxytocin signaling pathway, Platelet activation, Retrograde endocannabinoid signaling, Serotonergic synapse, and Caffeine metabolism (Fig. 4A). And the significant compounds were Prostaglandin H2 and Theophylline. Simultaneously, the annotation of level 2 KEGG pathway included Amino acid metabolism, Biosynthesis of other secondary metabolites, Nucleotide metabolism, Endocrine system, Immune system, Lipid metabolism, and Nervous system (Fig. 4B). And the significant compounds were Indole-3-acetaldehyde, Theophylline, Inosine, and Prostaglandin H2 (Table 3). Among these, Prostaglandin H2 and Theophylline were showed by enriched and annotated pathways at the same time.

Combined with the above results, Theophylline (pos_906) and Inosine (pos_1800) in drug resistance patients, the drug sensitivity group was higher than that in the healthy control and latent infection groups (Fig. 3).

**Table 3**

KEGG pathway enrichment results

| ID | Description | Count | *p* value | *p* adjust | Enrichment Fold | Compound Sig |
|---|---|---|---|---|---|---|
| ko00232 | Caffeine metabolism | 1 | 0.01 | 0.03 | 74.77 | Theophylline; |
| ko04611 | Platelet activation | 1 | 0.01 | 0.03 | 104.68 | Prostaglandin H2; |
| ko04723 | Retrograde endocannabinoid signaling | 1 | 0.01 | 0.03 | 82.64 | Prostaglandin H2; |
| ko04921 | Oxytocin signaling pathway | 1 | 0.01 | 0.03 | 120.79 | Prostaglandin H2; |
| ko04726 | Serotonergic synapse | 1 | 0.03 | 0.04 | 37.39 | Prostaglandin H2; |
| ko00380 | Tryptophan metabolism | 1 | 0.05 | 0.06 | 19.39 | Indole-3-acetaldehyde; |
| ko00590 | Arachidonic acid metabolism | 1 | 0.05 | 0.06 | 20.94 | Prostaglandin H2; |
| ko00230 | Purine metabolism | 1 | 0.06 | 0.06 | 16.70 | Inosine |

# Relationship between serum markers and disease progression

Through OPLS-DA analysis, one-way ANOVA analysis, cluster analysis, and pathway analysis of differential metabolites, four potential serum markers were screened. The relative contents of these four serum markers (Inosine, Prostaglandin E1, Theophylline, and Cotinine 1) in the four groups were compared. The results showed that the relative contents of Inosine (Fig. 5A) and Prostaglandin E1 (Fig. 5C) were the lowest in the latent infection group and the highest in the drug sensitive group, which suggested that these two metabolites are expected to be potential biomarkers for the diagnosis of latent infection. In addition, the relative contents of Theophylline (Fig. 5B) and Cotinine 1 (Fig. 5D) increased in healthy control, latent infections, drug resistance, and drug sensitivity group, suggesting that these two metabolites can be used as potential markers to monitor disease progression and provide important ideas for timely control of disease deterioration.

## Discussion And Conclusions

As a strong pathogenic bacterium invading the host, tuberculosis causes a series of immune reactions, eventually causing the death of the host. It is a major public health problem facing the world. However, the diagnosis and treatment methods related to tuberculosis still have many limitations, such as time consuming, high cost, and heavy workload. This is also the main reason for the difficulty of TB conquering. With the development of metabolomics technology, it seems to bring hope to conquer

tuberculosis. Serum metabolites are the specific reflection of human biological processes. Different types of metabolites and their concentrations often suggest different information, and this feedback is timely and accurate. At present, the research on metabolic markers of *M. tuberculosis* and tuberculosis has also attracted more and more attention [11, 12, 13]. However, there is no report on the main monitoring indicators of tuberculosis from latent infection to active TB infection, which seriously affects the treatment cycle and prognosis of patients with tuberculosis.

In our study, LC-MS/MS technology was used to detect serum markers of metabolic markers in latent infection patients and active TB patients in Guangdong. The serum monitoring model of early diagnosis of latent infection and active pulmonary tuberculosis was constructed by machine learning and birth analysis, and the main molecular mechanisms were explored. It is expected to screen suitable markers to evaluate the matching monitoring indicators from latently infected persons to active TB infection, to provide some clinical guidelines for early detection, timely diagnosis, and precise treatment of tuberculosis [14].

In 2012, Weiner *et al.* [10] explored over 400 small molecules metabolite in serum of uninfected individuals, latently infected healthy individuals, and patients with active TB. The results showed histidine, cysteine, glutamine, tryptophan, citrulline, and creatine were at lower levels in the TB active group compared to the two control group, and other identified compounds with markedly differentiated abundance between the latently infected group and TB active groups were sialic acid (N-acetylneuraminate), 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF), inosine, xanthine, hypoxanthine, fibrinopeptide A, glucose, and mannose. This is not entirely consistent with our results. Both our results showed the relative content of Inosine was lowest in the latently infected group, but other metabolites were not consistent. The reason may be the ethnic differences and eating habits of clinical samples. Furthermore, our results showed that the relative contents of Inosine and Prostaglandin E1 were the lowest in the latent infection group, which suggested that these can become potential biomarkers for the diagnosis of latent infection. And the relative contents of Theophylline and Cotinine 1 successively increased in healthy control, latent infections, drug resistance, and drug sensitivity group, suggesting that these two metabolites can be used as potential markers to monitor disease progression and provide important ideas for timely control of disease deterioration. These metabolites were annotated to the following KEGG pathway, including Purine metabolism (Inosine), Arachidonic acid metabolism; Serotonergic synapse; Oxytocin signaling pathway; Platelet activation; Retrograde endocannabinoid signaling (Prostaglandin H2), and Caffeine metabolism (Theophylline).

In conclusions, this study is based on LC-MS technology to build a serum monitoring model for the early diagnosis of latent infection and active pulmonary tuberculosis. It provides a reference for early diagnosis and diagnosis of latent tuberculosis patients and has great clinical application prospects.

# Abbreviations

GC/MS: Gas chromatography-mass spectrometry; GC/MS: Liquid chromatography-tandem mass spectrometry; NMR: Nuclear Magnetic Resonance. TB: Tuberculosis; M. tuberculosis: Mycobacterium tuberculosis; Mtb Mycobacterium tuberculosis; KEGG: Kyoto Encyclopedia of genes and genes; PLS-DA: Partial least squares discriminant analysis; VIP: variable importance in the project; OPLS-DA: Partial Least Squares Discriminant Analysis; PCA: Principal Component Analysis.

## Declarations

# Acknowledgements

We thank all the patients and healthy individuals for their participation.

# Authors' contributions

Y-X.Li and K-D.Zheng analyzed the data and wrote the manuscript. Y.D and H-J.Liu conducted the experiments. Y-Q.Tang and J.Wu evaluated the clinical data and provided the clinical samples. D-Z.Lin and Z.Zhang designed and supervised the studies. All authors read and approved the manuscript.

# Funding

This project was supported by the National Natural Science Foundation of China (No. 82003786), Guangdong Basic and Applied Basic Research Foundation (No. 2019A1515110338), Science and Technology Program of Guangzhou (202002030200), Natural Science Foundation of Guangdong Province (No. 2020A1515010335), Scientific Research Foundation of Guangdong Medical University (GDMUZ2019002), High-level Hospital Construction Summit Plan Project of the Fourth People's Hospital of Foshan, Grant/Award (No. FSSYKF-2020021, FSSYKF-2020019), *Innovation and Entrepreneurship Team* Introduction *Project* of Foshan (2018IT100262).

# Availability of data and materials

The datasets supporting the conclusions of this article are included within the article.

# Ethics approval and consent to participate

Informed written consent was received from all participants who provided serum after detailed explanation of the study. The study was approved by the ethics committees of the Dongguan Key Laboratory of Medical Bioactive Molecular Developmental and Translational Research, Foshan Fourth People's Hospital, and Dongguan Sixth People's Hospital. All experiments were performed in accordance with relevant guidelines and regulations.

# Consent for publication

Not applicable.

# Competing interests

The authors declare that they have no competing interests.

# Author details

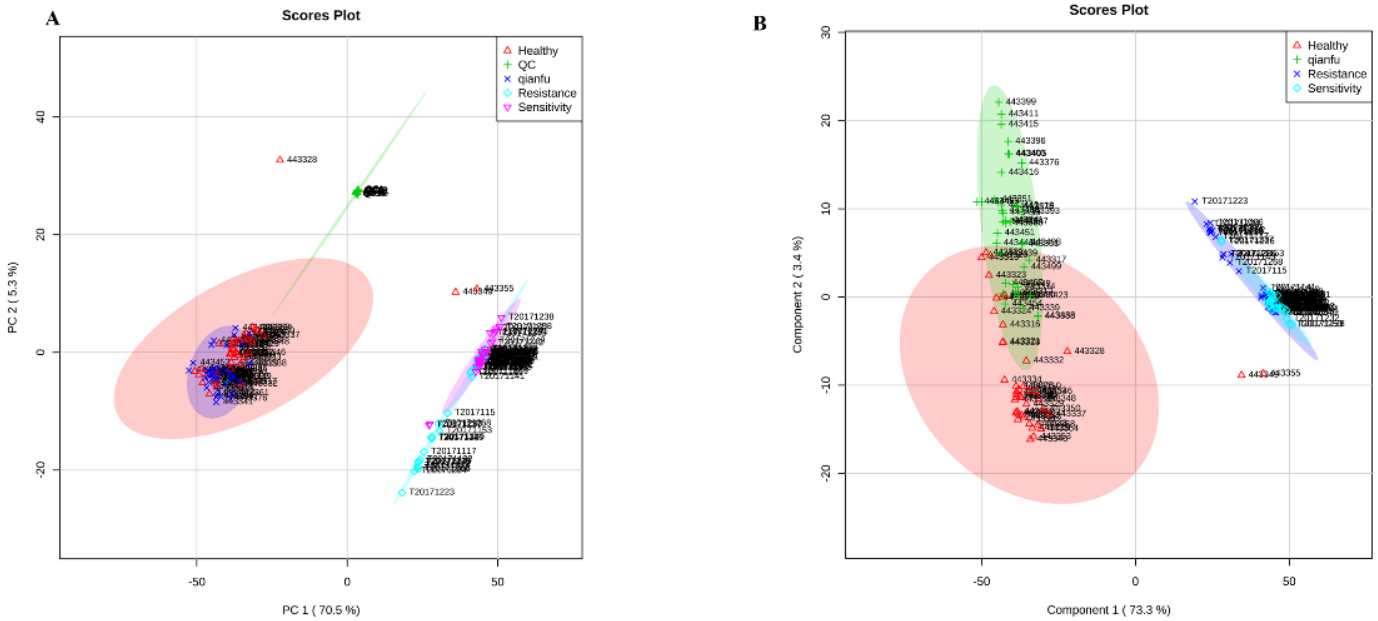
[1]Department of Laboratory Medicine, Foshan Forth People's Hospital, Foshan, Guangdong 528041, China. [2]Research and Development Department, Foshan Longsee Biotechnology Co, Ltd., Foshan, Guangdong 528200, China. [3]Department of Respiratory Medicine, Affiliated Hospital of Guangdong Medical University, Zhanjiang, Guangdong 524000, China. [4]Peptide and Protein Research and Application Key Laboratory of Guangdong Medical University, Zhanjiang, Guangdong 524000, China

# References

1. Cardona PJ. Pathogenesis of tuberculosis and other mycobacteriosis. *Enferm Infecc Microbiol Clin (Engl Ed)* 2018; 36(1):38–46. doi: 10.1016/j.eimc.2017.10.015
2. Brindle JT, Nicholson JK, Schofield PM, et al. Application of chemometrics to 1H NMR spectroscopic data to investigate a relationship between human serum metabolic profiles and hypertension. *The Analyst* 2003; 128(1):32–6. doi: 10.1039/b209155k
3. Yang J, Xu G, Zheng Y, et al. Diagnosis of liver cancer using HPLC-based metabonomics avoiding false-positive result from hepatitis and hepatocirrhosis diseases. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences* 2004;813(1-2):59–65. doi: 10.1016/j.jchromb.2004.09.032
4. Wang C, Kong H, Guan Y, et al. Plasma phospholipid metabolic profiling and biomarkers of type 2 diabetes mellitus based on high-performance liquid chromatography/electrospray mass spectrometry and multivariate statistical analysis. *Analytical chemistry* 2005;77(13):4108–16. doi: 10.1021/ac0481001
5. Dereziński P, Klupczynska A, Sawicki W, et al. Amino Acid Profiles of Serum and Urine in Search for Prostate Cancer Biomarkers: a Pilot Study. *International Journal of Medical Sciences* 2017;14(1):1-12. doi: 10.7150/ijms.15783. eCollection 2017.
6. Tsai CY, Chen CT, Lin CH, et al. Proteomic analysis of Exosomes derived from the Aqueous Humor of Myopia Patients. *International Journal of Medical Sciences* 2021;18(9):2023-2029. doi: 10.7150/ijms.51735. eCollection 2021.

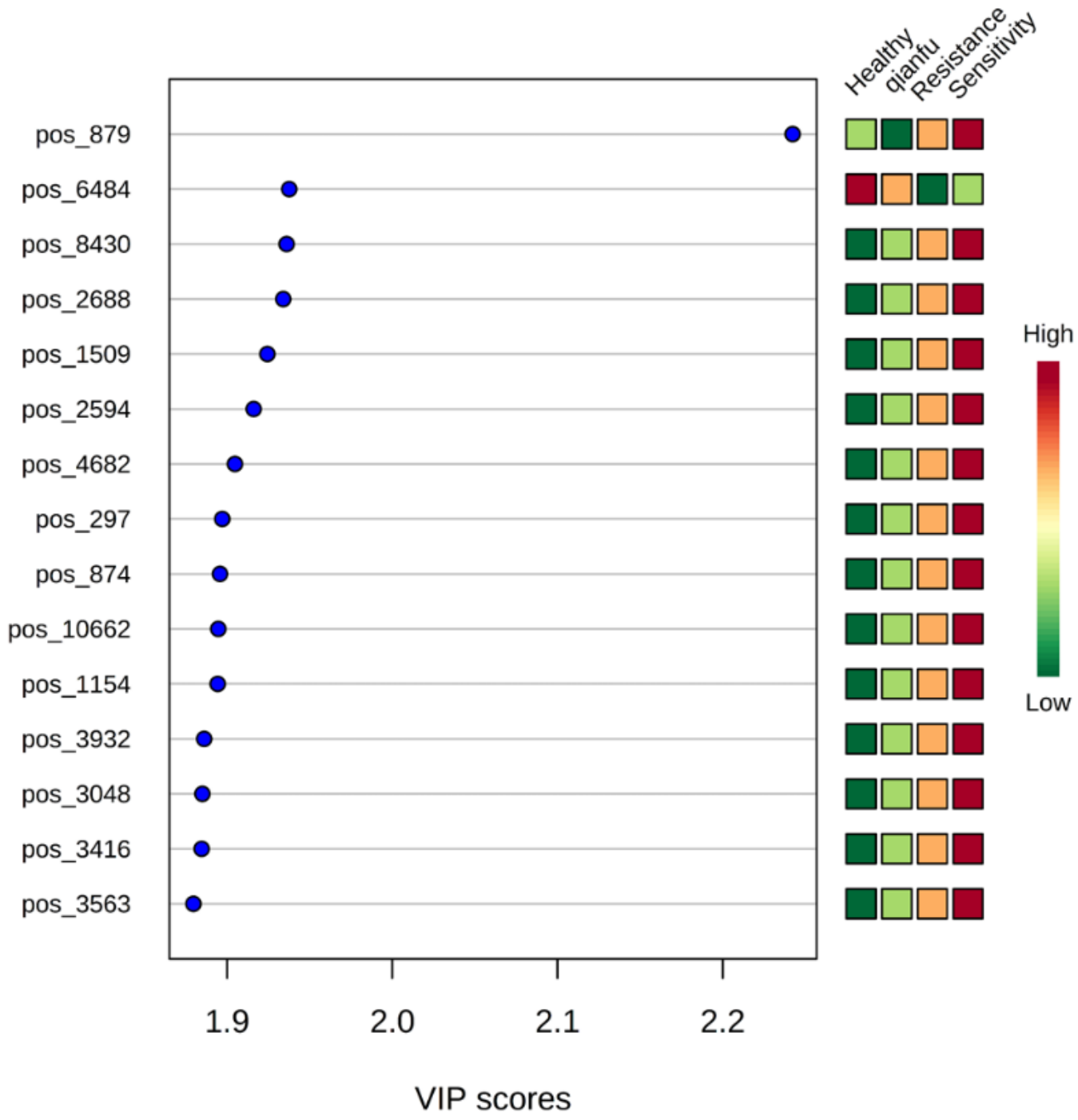
7. Suhre K, Wallaschofski H, Raffler J, et al. A genome-wide association study of metabolic traits in human urine. *Nature genetics* 2011;43(6):565–9. doi: 10.1038/ng.837
8. Lau SK, Lam CW, Curreem SO, et al. Identification of specific metabolites in culture supernatant of Mycobacterium tuberculosis using metabolomics: exploration of potential biomarkers. *Emerging microbes & infections* 2015;4(1):e6. doi: 10.1038/emi.2015.6
9. Yeware A, Gample S, Agrawal S, et al. Using diphenyleneiodonium to induce a viable but non-culturable phenotype in Mycobacterium tuberculosis and its metabolomics analysis. *PloS one* 2019;14(8):e0220628. doi: 10.1371/journal.pone.0220628
10. Weiner J, 3rd, Parida SK, Maertzdorf J, et al. Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients. *PloS one* 2012;7(7):e40221. doi: 10.1371/journal.pone.0040221
11. Chang K, Lu W, Wang J, et al. Rapid and effective diagnosis of tuberculosis and rifampicin resistance with Xpert MTB/RIF assay: a meta-analysis. *The Journal of infection* 2012;64(6):580–8. doi: 10.1016/j.jinf.2012.02.012
12. Bäckhed F, Fraser CM, Ringel Y, et al. Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. *Cell host & microbe* 2012;12(5):611–22. doi: 10.1016/j.chom.2012.10.012
13. Blaser MJ, Falkow S. What are the consequences of the disappearing human microbiota? *Nature reviews Microbiology* 2009;7(12):887–94. doi: 10.1038/nrmicro2245
14. Zhou J, Yin Y. Use of Liquid Chromatography-Mass Spectrometry-Based Metabolomics to Identify Biomarkers of Tuberculosis. *Methods in molecular biology (Clifton, NJ)* 2019;1859:241–51. doi: 10.1007/978-1-4939-8757-3_13

# Figures



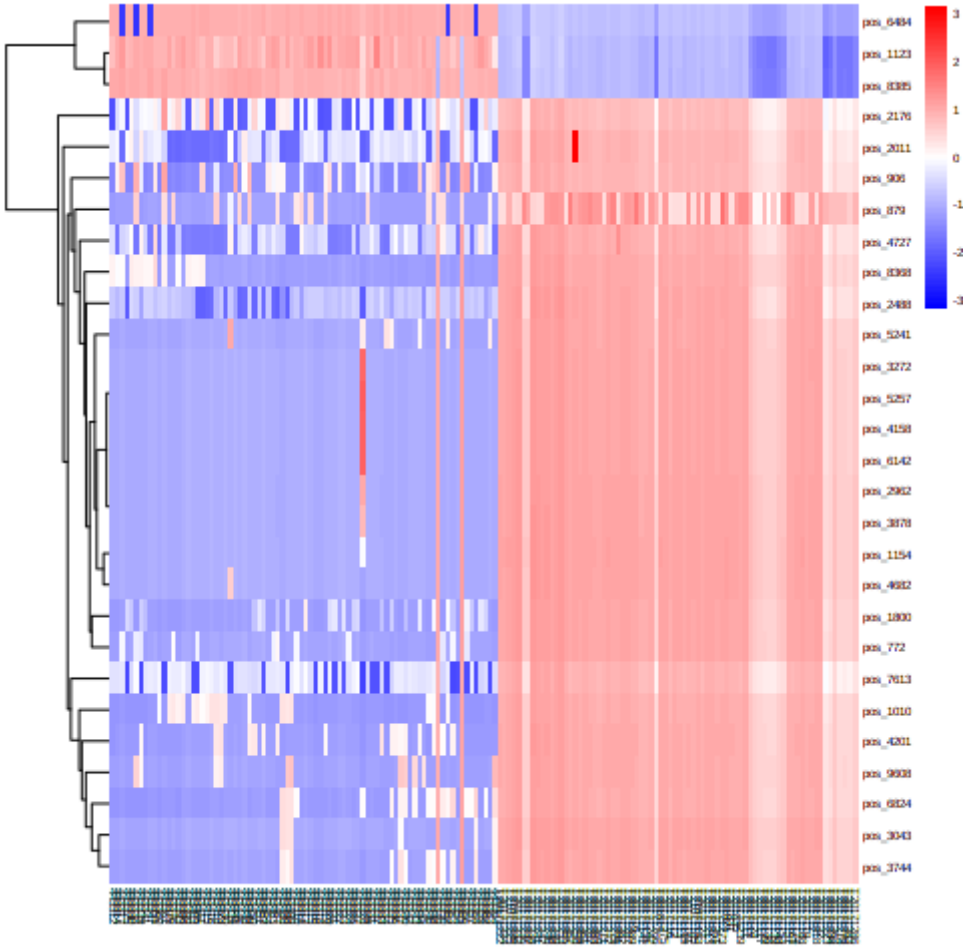

## Figure 1

Analysis of serum metabolites in each group by PCA and OPLS-DA. (**A**) The PCA analysis of healthy control, latent infection (qianfu), drug resistance, and drug sensitivity group; (**B**) The OPLS-DA analysis of healthy control, latent infection (qianfu), drug resistance, and drug sensitivity group.



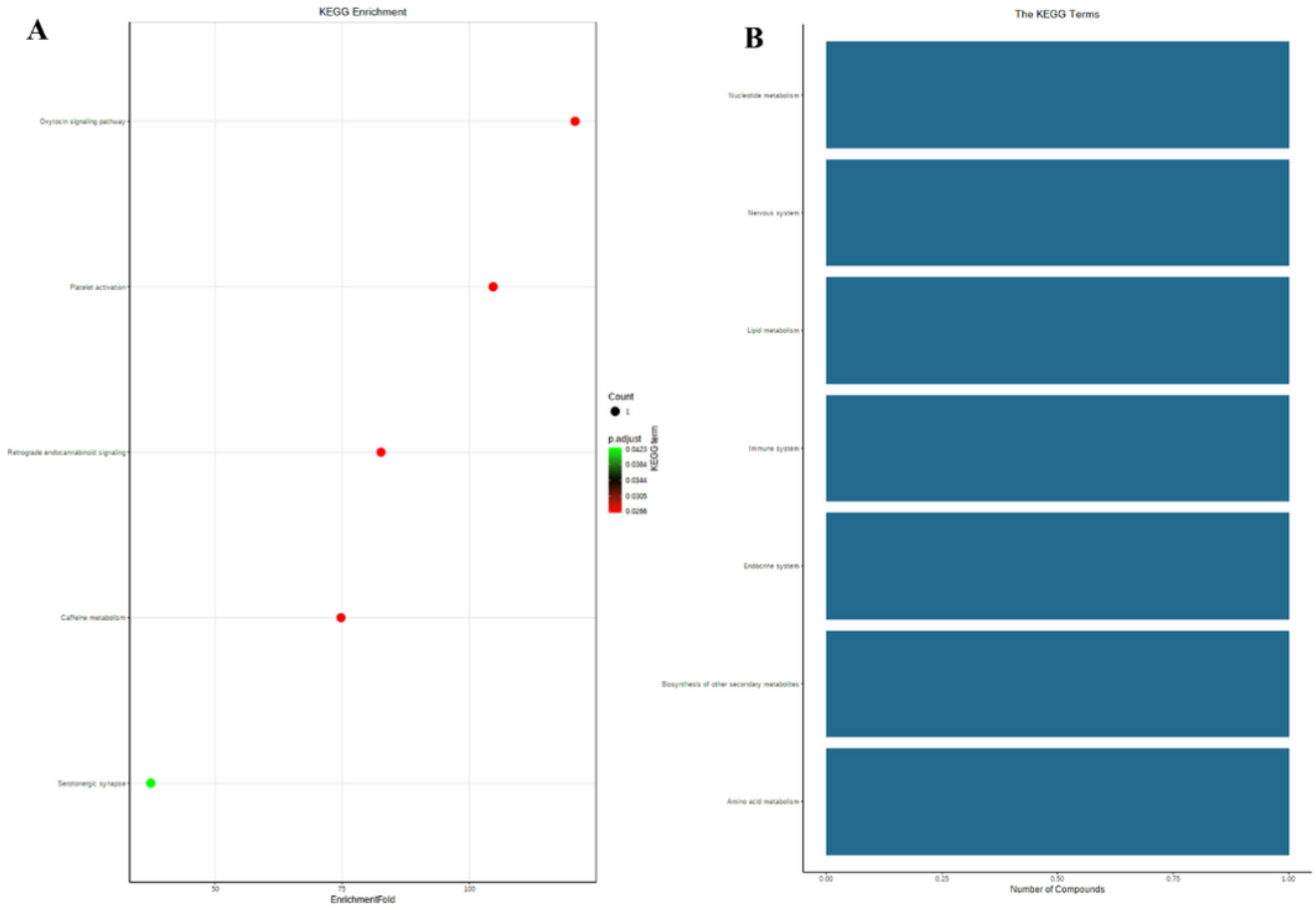

## Figure 2

Partial least squares discriminant analysis (PLS-DA) variable importance in projection (VIP) plot of significantly differential metabolites in healthy control, latent infection, drug resistance patients, drug sensitivity group. The x-axis represented the VIP scores, and the y-axis represented the compounds. Red and green colors represented increased and decreased levels of metabolites, respectively.



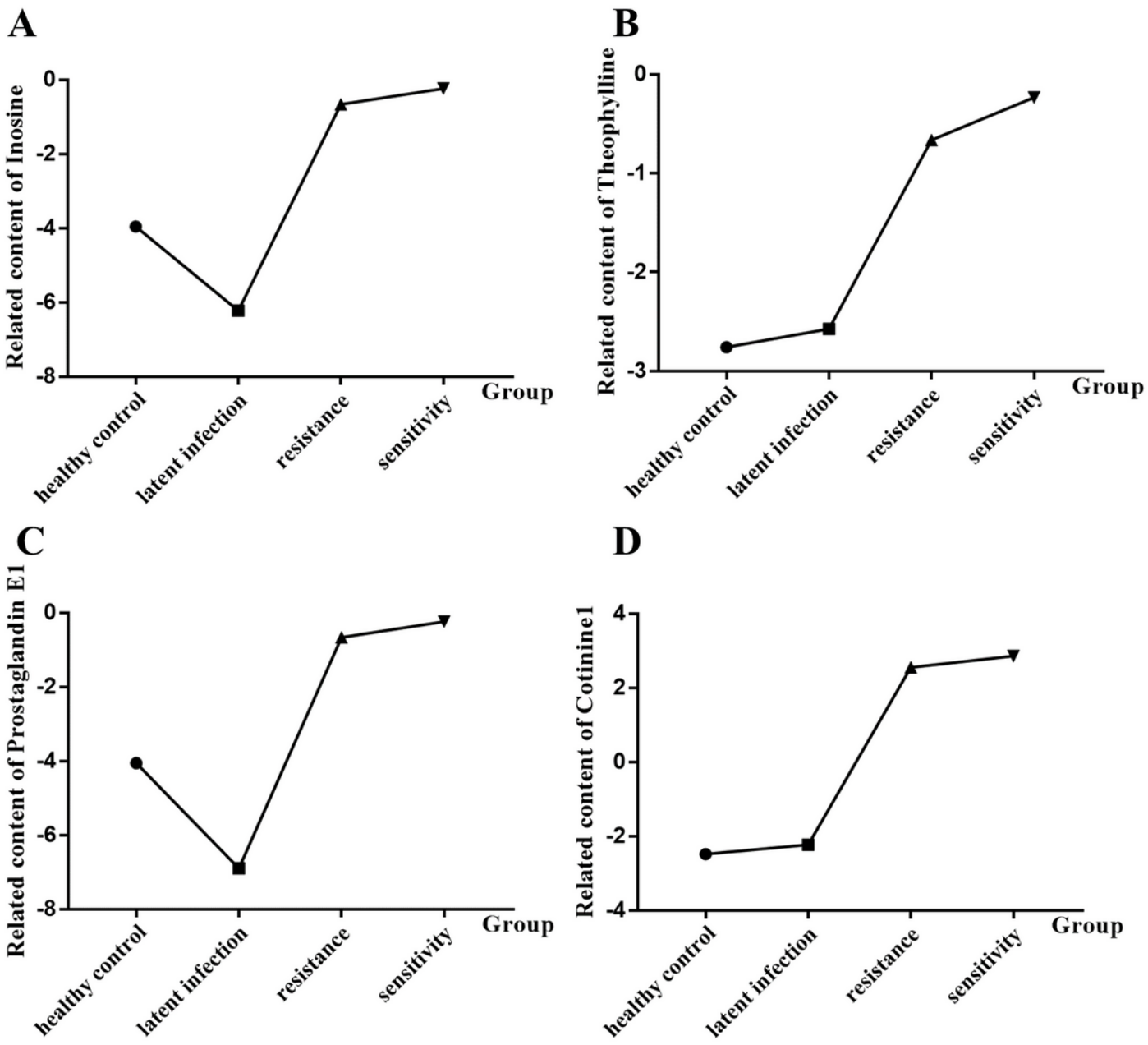

## Figure 3

Clustering results of metabolic biomarkers in healthy control, latent infection, drug resistance, and drug sensitivity group. The x-axis represented samples, and the y-axis represented the M/z of compounds.




## Figure 4

Pathway analysis of differential metabolites in healthy control, latent infection, drug resistance, and drug sensitivity group. (**A**) The significantly enriched pathway (*p* adjust < 0.05); (**B**) The annotation of level 2 KEGG pathway.




## Figure 5

The relative contents of four potential serum markers. (**A**) Inosine; (**B**) Theophylline; (**C**) Prostaglandin E1; (**D**) Cotinine 1.