Int J Clin Exp Pathol 2016;9(3):3256-3266
www.ijcep.com /ISSN:1936-2625/IJCEP0021658

Original Article

Serum metabolomic study for the detection of candidate biomarkers of tuberculosis

Lijun Zhong[1], Juntuo Zhou[2], Xi Chen[1], Yuxin Yin[1,2,3]

[1]Medical and Health Analytical Center, Peking University Health Science Center, Beijing, China; [2]Department of Pathology, Institute of Systems Biomedicine, School of Basic Medical Sciences, Peking University Health Science Center, Beijing, China; [3]Beijing Key Laboratory of Tumor Systems Biology, Beijing, China

Received December 11, 2015; Accepted February 18, 2016; Epub March 1, 2016; Published March 15, 2016

Abstract: Tuberculosis (TB) is a widespread infectious disease caused by various strains of mycobacteria. Early diagnosis of this disease is important. However, only a few biochemical tests or markers are available for diagnosis. In this study, serum metabolites from normal and TB patients were analyzed by mass spectrometry coupled with high-performance liquid chromatography to identify serum markers that could provide clues for TB diagnosis and altered metabolic pathways in the disease state. Multivariate statistical analysis showed a distinction between the normal and TB patients based on 1949 identified metabolites. Further examination combined with t-test analysis and fold-change determination revealed significant perturbations in 99 serum metabolites in the TB group. Finally, pathway analysis enriched the perturbed metabolic pathways in TB (i.e., glycerophospholipid and arachidonic acid metabolism). Overall, our study demonstrated the potential in developing biomarkers from serum for TB diagnosis. Furthermore, the implementation of the metabolomics approach and bioinformatics helped us gain an understanding of the numerous metabolic pathways affected in pathologic conditions.

Keywords: Tuberculosis, biomarkers, metabolomics, multivariate statistical analyses, pathway analysis

Introduction

Tuberculosis (TB) results in an estimated 1.7 million deaths each year; the global number of new cases (more than 9 million) continues to increase. Twenty-two low- and middle-income countries account for more than 80% of the active cases worldwide [1]. The common medical evaluation for TB includes physical examination, chest X-ray, and microbiological examination. The culture of *Mycobacterium tuberculosis* remains the gold standard for the laboratory diagnosis of pulmonary TB [2]. Reviewing data reported over 20 years, scholars established a state-of-the-art procedure for the routine culture of *M. tuberculosis* in both developed and developing countries [3]. Meanwhile, biomarkers are indispensable to disease diagnosis and the development of new TB therapeutics and vaccines. However, in the development of new anti-TB drugs, especially with the multiple permutations of drug combinations, doses, and dosing intervals needing evaluation in phase 2 clinical trials, the use of a two-month culture conversion as the principal endpoint will significantly prolong the time to introduce these new drugs into clinical practice [4-7]. Therefore, early detection can be instrumental in treating TB.

Metabolic profiling is an important tool for biomarker studies [8, 9]. Metabolomics can be used to determine the biochemical mechanisms underlying a given pathologic condition. Metabolism is a complex system comprised of thousands of individual molecules [10]. These molecules play different roles in cellular functions, ranging from membrane structural components to secondary messengers [11]. Any perturbation in a biological system is expected to give rise to changes in the composition of the metabolite pool. Several techniques have been implemented for metabolite profiling; such techniques include nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry coupled with high-performance liquid chromatography (LC-MS) or gas chromatography (GC-MS) [12]. The LC-MS-based diagnosis from biofluids has shown growing importance in the last few years. This approach requires a simpler

Table 1. Demographic information on the study subjects

Sample	Number	Mean age (± SD)	Gender	
			Male	Female
TB	20	42.1 (±15.1)	11	9
Normal	18	42.6 (±16)	10	8

sample preprocessing procedure compared with GC-MS techniques and has been proven to be a powerful strategy for obtaining essential information from complex and intact biological samples [13]. Thus, the present study aims to investigate the serum metabolic profile of TB patients by LC-MS to gain insights into the pathophysiologic processes involved in this disease and search for disease biomarkers.

Materials and methods

Ethics statement

The study was performed and conducted at Peking University Health Science Center. All samples were collected at the 309th Hospital of The Chinese People's Liberation Army. The research participants were approached using approved ethical guidelines and informed about the study both verbally and in written form. The informed consent forms were signed and acquired from the participants prior to sample collection.

Material and reagents

HPLC grade ammonium acetate and formic acid were purchased from Sigma, and HPLC grade acetonitrile, methanol, isopropanol and water were purchased from Fisher Scientific. All other reagents were from commercial suppliers and of standard biochemical quality.

Sample collection

All samples involved in this study were collected from a Chinese population. The demographic information on the study subjects is shown in **Table 1**. Samples of venous blood were collected from TB patients who underwent medical evaluation for TB at the 309th hospital of the Chinese people's liberation army. The control samples were selected from age- and gender-matched healthy donors who exhibited no sign of neurological and psychiatric illness. Samples were harvested from 18 normal and 20 diseased individuals by standard venipuncture. Once extracted, 5 mL of blood were immediately transferred to a sterile universal container without anticoagulant. The sample was centrifuged at 3,500 rpm for 15 min. The supernatant was transferred into 1.5 mL centrifuge tubes, which were immediately transferred into a freezer at -80°C for storage.

Preparation of samples

Serum (50 μL) was extracted using fourfold cold chloroform: methanol (2:1) solution. The samples were vortexed for 30 s and then shaken at 1,400 rpm for 15 min at room temperature. The mixture was centrifuged at 13,000 rpm for 15 min, and then the lower organic phase was collected and evaporated at room temperature under vacuum. The resulting residue was dissolved in 50 μL chloroform:methanol (1:1) and then diluted with isopropanol: acetonitrile: H_2O (2:1:1) prior to Liquid Chromatography-Mass Spectrometer (LC-MS) analysis.

LC-MS analysis

Samples were subjected to LC-MS using a Water Acquity CSH 3.5 μm C18 column (4.6 mm × 100 mm) under the following conditions. For the LC, A-acetonitrile/water (60/40) and B-isopropanol/acetonitrile (90/10) were employed. Both A and B contain 10 mM ammonium acetate and 0.1% formic acid. The following gradient was applied: initial 60% A to 57% A at 2 min, 50% A at 2.1 min*, 46% A at 12 min, 30% A at 12.1 min*, and 1% A at 18 min. The conditions were returned to the initial values at 18.5 min with equilibration for an additional 2 min (*indicates ballistic gradient). Flow rate was 0.7 mL/min, and column temperature was maintained at 50°C. MS analysis was performed in positive and negative modes, which were operated in full-scan mode at m/z 150-950 with a resolution of 70 k. The maximum injection time was 100 ms, and the number of microscans per scan was 1. Each full scan was followed by 10 higher energy collisional dissociation scans by tandem MS. The source temperature was set to 350°C, whereas the capillary voltages were 3.2 kV for the positive mode and 2.8 kV for the negative mode.

Peak extraction and metabolite identification

SIEVE 2.0 software (Thermo Fisher Scientific) was used for LC-MS raw data analysis. SIEVE aligns the MS spectra over time for different

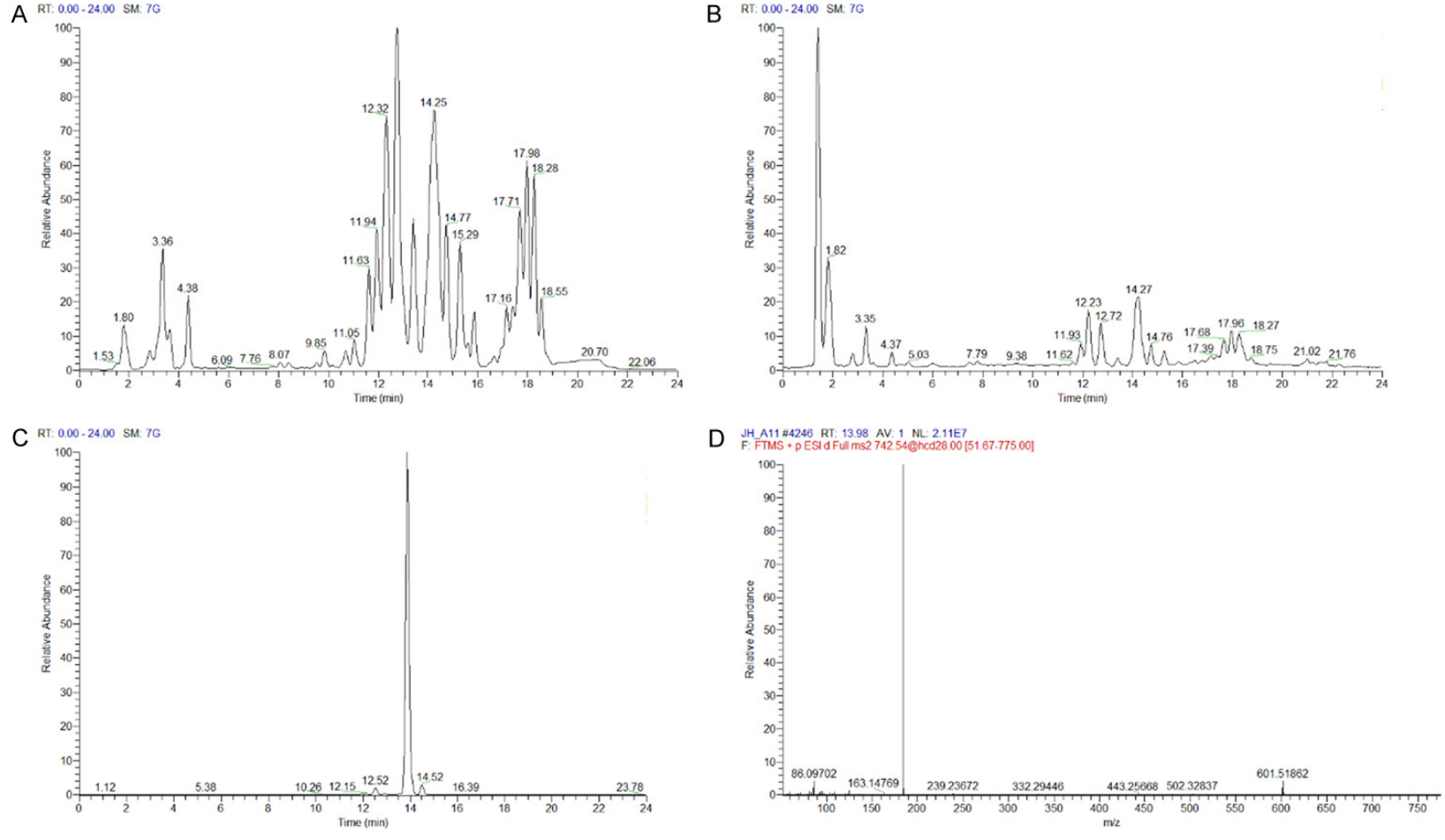
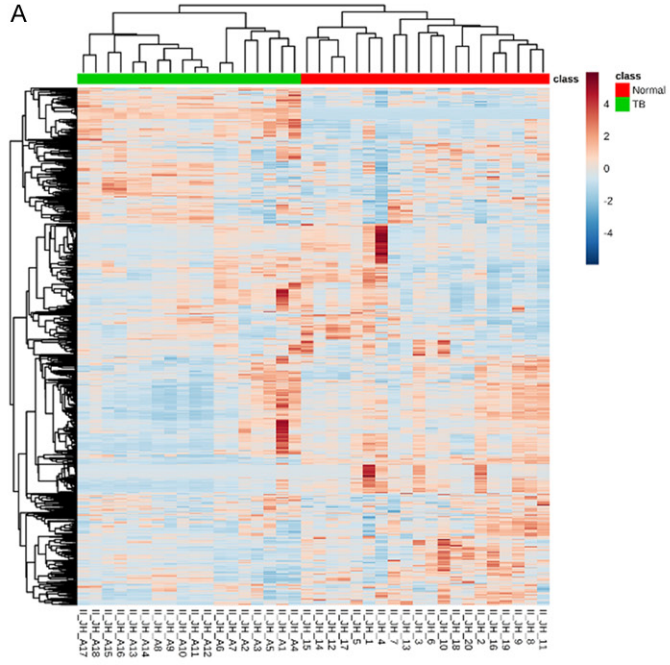
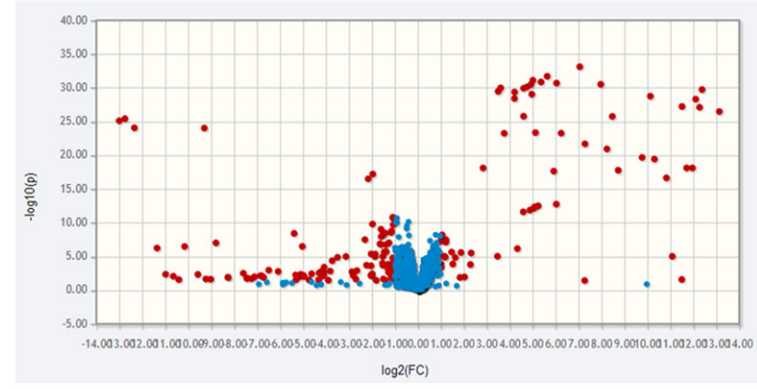
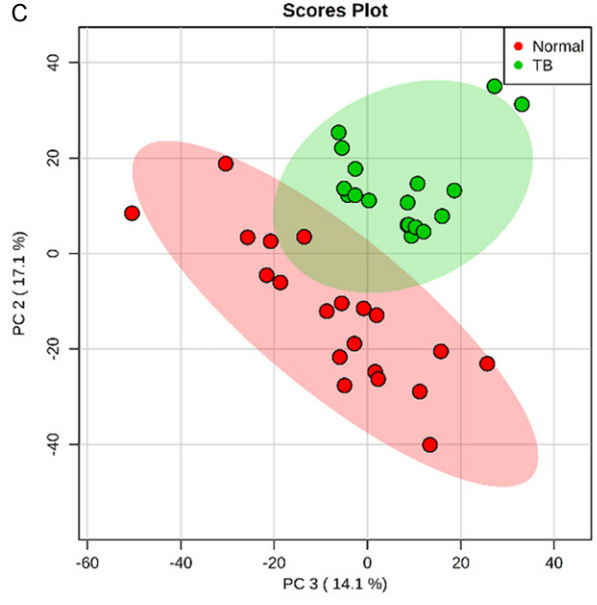


Figure 1. Metabolic profiling results of TB patients and normal controls. A. A typical total ion chromatogram obtained by LC-MS for metabolomic profiling (positive ion mode). B. A typical total ion chromatogram obtained by LC-MS for metabolomic profiling (negative ion mode). C. Extracted ion chromatogram for PC (16:1(9Z)/P-18:1(9Z)) (m/z 742.57642). D. Tandem MS spectra for PC (16:1(9Z)/P-18:1(9Z)).


A
class
Normal
TB
4
2
0
-2
-4


B
-log10(p)
log2(FC)


C
Scores Plot
Normal
TB
PC 2 (17.1 %)
PC 3 (14.1 %)

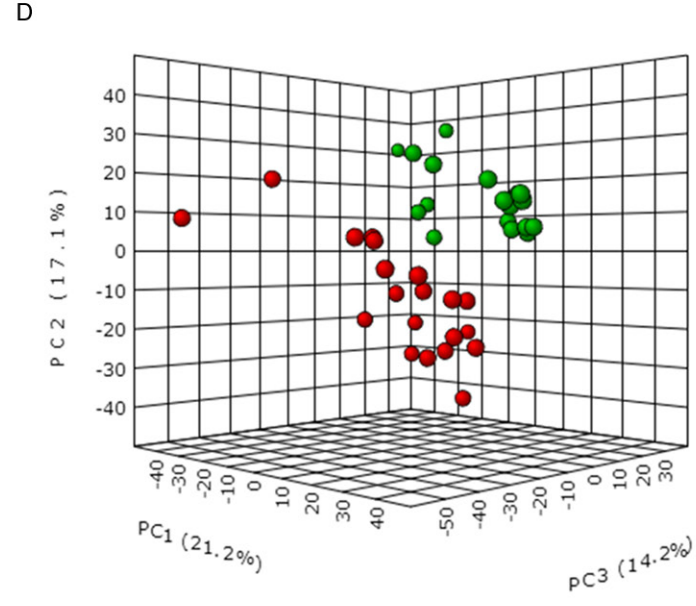

D
PC 2 (17.1 %)
PC1 (21.2%)
PC3 (14.2%)

Figure 2. Representations of the 1949 extracted metabolites. A. Heat map of the 1949 extracted metabolites. The columns represent 38 serum samples, and the rows represent different metabolites. B. Scatter plots of all the 1949 metabolites with a log 2 fold change (X) and -lg *p*-value (Y). C. PCA analysis of the metabolomic profiles in 2D mode. D. PCA analysis of the metabolomic profiles in 3D mode.

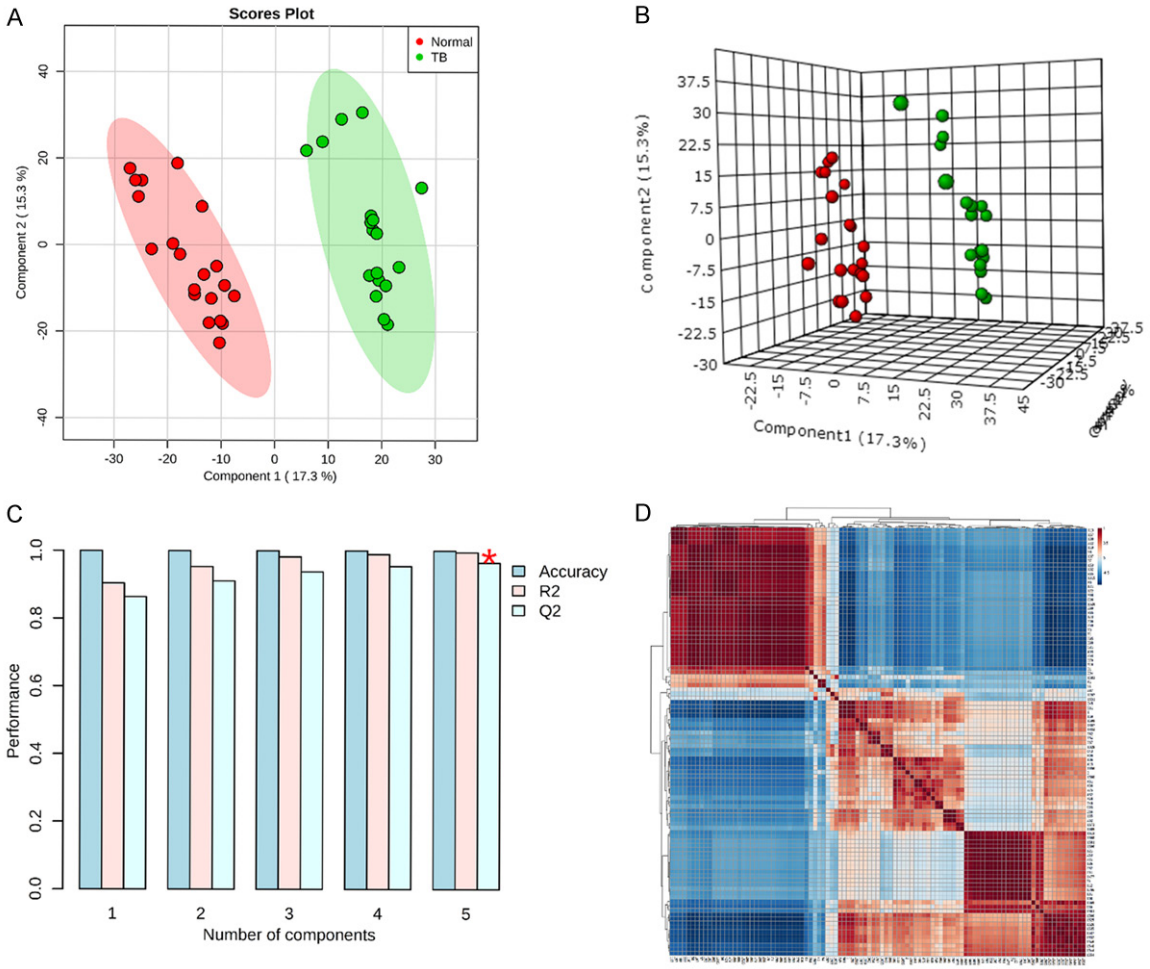


Figure 3. Multivariate statistical analysis of the metabolic profiling result. A. PLS-DA analysis of the metabolomic profiles in 2D mode. B. PLS-DA analysis of the metabolomic profiles in 3D mode. C. Cross validation of the PLS-DA models. D. Correlation matrices of the 99 significantly changed metabolites between the TB patients and controls.

experimental conditions and determines features in the dataset [9]. Parameters for the SIEVE analysis were set as follows: frame time width, 0.5 min; minimum scans across a peak, 9; signal-to-noise ratio, 10; m/z step width, 5; peak intensity threshold, 1,000,000; maximum number of frames, 5000; alignment bypass, false; m/z min, 150; m/z max, 900; starting retention time, 0 min; and end retention time, 23 min. To identify the extracted frames, we searched public metabolomics databases, including the Human Metabolome Database [14] (HMDB; http://www.hmdb.ca/) and LIPID MAPS (http://www.lipidmaps.org/), to find potential candidates.

Multivariate statistical and pathway analyses

After raw data analysis, a data matrix consisting of peak areas corresponding to a unique m/z and retention time was generated for further multivariate statistical analysis using MetaboAnalyst 3.0 web service [15]. The multivariate statistical analyses adopted in the study include principal component analysis (PCA) and partial least-squares discriminant analysis (PLS-DA). The importance of each metabolite in the PLS-DA was evaluated by variable importance in the projection (VIP) score. Metabolites with scores greater than 1.0 were considered important in this model. The Pearson correla-

Table 2. Details of the 99 significantly distinct metabolites in the TB and normal groups

Comp ID	Comp Name	VIP	HMDB ID	Ratio (Normal/TB)	*p* value
1	(±)-3-Hydroxynonanoic acid	2.4062	HMDB31513	3.621161089	4.397E-16
2	L-Acetylcarnitine	1.3485	HMDB00201	2.279195055	0.0004435
31	Hexaethylene glycol	2.4245	HMDB61822	0.128649868	2.843E-14
37	Heptaethylene glycol	2.4956	HMDB61835	0.068570557	4.534E-16
41	Levofloxacin	1.6736	HMDB01929	0.068686003	1.043E-05
54	Fenofibrate	1.8175	HMDB15173	0.036775754	1.932E-05
58	6-Keto-prostaglandin F1a	2.5318	HMDB02886	0.050349974	1.346E-16
70	Resolvin D2	2.5247	HMDB02294	0.000179044	4.26E-14
71	11-Hydroxyeicosatetraenoate glycerylester	2.5161	HMDB12530	0.002568657	1.498E-13
74	Succinyladenosine	1.3093	HMDB00912	99.88549854	0.0013304
86	Ramipril Diketopiperazine	2.3935	HMDB60869	0.000529187	2.162E-13
101	N-desalkyl delavirdine	1.2782	HMDB61021	16.68478936	0.0014649
107	alpha-hydroxysalmeterol	2.5362	HMDB61045	0.050098939	3.564E-16
108	Melledonal A	1.2698	HMDB35865	22.11523995	0.0016574
110	Latanoprost	2.5393	HMDB14792	0.034254302	1.375E-15
112	Artocarpin	1.1998	HMDB30849	247.9017308	0.003426
113	Flurandrenolide	2.2232	HMDB14984	0.039797729	1.509E-09
121	Aripiprazole	1.1548	HMDB05042	40.99139432	0.0048159
136	Sofalcone	1.1226	HMDB42013	37.90894258	0.0061657
153	Netilmicin	2.5387	HMDB15090	0.076062199	2.401E-16
157	Citronellyl beta-sophoroside	2.2353	HMDB32839	0.032680582	5.749E-10
159	26-Methyl nigranoate	2.5026	HMDB33409	663.8719319	7.435E-17
165	Polyporusterone C	1.8523	HMDB38497	1204.916467	2.88E-06
172	(24E)-15alpha-Acetoxy-3alpha-hydroxy-23-oxo-7,9(11),24-lanostatrien-26-oic acid	2.4246	HMDB35387	0.000281467	8.912E-15
173	alpha-Ionol O-[arabinosyl-(1->6)-glucoside]	2.5334	HMDB38732	0.000763991	1.272E-13
182	LysoPC(18:2(9Z,12Z))	2.5366	HMDB10386	0.082219984	3.426E-16
188	VPGPR Enterostatin	2.2565	HMDB03577	0.028223483	3.46E-10
202	3alpha-Acetomethoxy-11alpha-oxo-12-ursen-24-oic acid	1.826	HMDB36673	2678.743318	3.797E-06
204	endo-1,4-beta-Xylanase	1.6722	HMDB40436	0.000439066	2.343E-05
206	Ganoderic acid eta	2.5313	HMDB36309	0.00015983	1.285E-13
222	Protoporphyrinogen IX	2.2632	HMDB01097	0.02538496	1.894E-10
224	2'',4'',6''-Triacetylglycitin	1.0784	HMDB34694	99999.9	0.00001
245	Lutein 5,6-epoxide	2.503	HMDB41590	7689.746387	6.429E-17
251	Physalin E acetate	1.18	HMDB31878	29.64570628	0.0037804
252	6-Hydroxysandoricin	1.1306	HMDB37556	48.70927741	0.0058461
253	(24R)-5b,8b-Epidioxyergosta-6,22E-dien-3b-ol 3-glucoside	2.5407	HMDB37956	0.030479355	5.634E-16
259	Sucrose monopalmitate	2.5209	HMDB39579	7.86816E-05	3.653E-13
263	Calendoflaside	1.0988	HMDB38819	40.99824573	0.0076113
266	PE(14:1(9Z)/P-16:0)	1.9061	HMDB08883	457.6340236	1.451E-06
270	2alpha-Hydroxypyracrenic acid	2.5421	HMDB29780	0.022470759	3.098E-16
287	CE(20:5(5Z,8Z,11Z,14Z,17Z)	1.049	HMDB06731	0.248603406	0.0056261
299	Medicagenic acid 3-O-beta-Dglucoside	2.5377	HMDB34552	0.000148361	6.517E-14
316	Gingerglycolipid B	2.5429	HMDB41094	0.028799226	1.351E-16
348	PC(14:0/P-18:1(11Z))	1.1076	HMDB07897	6.522745477	0.0085438
362	SM(d18:0/16:0)	1.3975	HMDB10168	0.480498162	1.469E-05
365	Hovenidulcioside B2	2.5255	HMDB41549	0.000276174	4.847E-13
368	28-Glucosyl-19(29)-dehydroursolic acid 3-arabinoside	2.4247	HMDB38019	0.000239781	5.323E-15
375	PE(18:1(9Z)/P-18:0)	1.4571	HMDB09082	4.512853455	0.0003776
398	Lycoperoside D	2.5415	HMDB34484	0.014102242	7.146E-16
403	Fasciculol E	2.5347	HMDB35189	0.029748663	1.566E-15
406	PC(16:1(9Z)/P-18:1(9Z))	1.4478	HMDB08030	2.291767638	0.0006419
413	PE(P-18:1(9Z)/20:5(5Z,8Z,11Z,14Z,17Z))	1.4584	HMDB11453	2.575243618	0.0007711
421	28-Glucosyloleanolic acid 3-arabinoside	2.4154	HMDB36963	0.015742479	2.092E-15
432	PE(20:1(11Z)/P-18:1(11Z))	1.441	HMDB09281	3.797835694	0.0002299

461	2-O-(5,8,11,14,17-Eicosapentaenoyl)-1-Ohexadecylglycero-3-phosphocholine	1.1408	HMDB39528	2.099227919	0.0101288
466	PE(20:4(5Z,8Z,11Z,14Z)/P-18:1(11Z))	1.1336	HMDB09413	2.087552128	0.0106042
531	PE(18:4(6Z,9Z,12Z,15Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	2.5449	HMDB09210	0.018683592	4.969E-16
561	PC(20:5(5Z,8Z,11Z,14Z,17Z)/P-18:1(11Z))	2.3902	HMDB08523	4.191945566	1.074E-13
643	PE(16:0/24:0)	1.929	HMDB08947	0.458365581	8.518E-07
748	PE(24:1(15Z)/18:4(6Z,9Z,12Z,15Z))	1.2547	HMDB09756	6.756889444	0.0024461
754	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/16:0)	1.6645	HMDB08725	0.438238516	9.443E-06
767	PE(18:3(9Z,12Z,15Z)/24:1(15Z))	1.665	HMDB09179	0.422132939	4.33E-06
795	Goyaglycoside g	2.5491	HMDB38350	0.007071605	4.992E-16
806	Rhizonin A	2.5163	HMDB30525	0.038039027	2.388E-14
809	PC(16:1(9Z)/22:2(13Z,16Z))	1.9669	HMDB08020	0.456246595	1.527E-07
865	Vinaginsenoside R1	2.4476	HMDB34624	0.00074502	7.785E-16
900	Sucrose acetate isobutyrate	1.8534	HMDB37136	30.45298929	2.326E-06
993	PC(MonoMe(9,5)/DiMe(9,5))	1.073	HMDB61466	31.10412059	0.0090453
1028	PI(20:3(8Z,11Z,14Z)/16:0)	1.9512	HMDB09886	0.403340811	3.044E-07
1297	Oleanolic acid 3-[rhamnosyl-(1->4)-glucosyl-(1->6)-glucoside]	1.756	HMDB40860	0.334914765	1.803E-05
1307	TG(20:2n6/20:5(5Z,8Z,11Z,14Z,17Z)/20:2n6)	1.0272	HMDB53602	2.039168819	0.0194611
1315	PIP(18:1(9Z)/20:1(11Z))	2.4694	HMDB09977	0.003180384	6.028E-18
1334	TG(18:3(6Z,9Z,12Z)/20:3n6/22:5(7Z,10Z,13Z,16Z,19Z))	1.4853	HMDB53108	2.204821012	0.0005838
1477	PIP3(16:0/16:0)	1.0922	HMDB10147	21.74853096	0.0070475
1520	Cavipetin C	2.4515	HMDB30366	0	2.775E-11
1525	Bis(methylsulfonylmethyl)disulfide	2.0397	HMDB41188	2.169630537	5.376E-10
1526	Succimer	2.0335	HMDB14706	2.624525638	7.794E-10
1532	1,4-Benzodioxin-2(3H)-one	2.0123	HMDB40528	2.781658266	3.101E-08
1535	Succimer	1.8919	HMDB14706	3.090619514	8.269E-08
1537	(Z)-Methyl 3-(methylsulfinyl)-1-propenyl disulfide	1.8834	HMDB32124	2.974452881	1.074E-07
1544	Dantron	1.7192	HMDB29752	3.732566749	4.903E-07
1546	3,4-Methylenedioxybenzaldehyde	1.8809	HMDB32612	2.504129467	6.593E-07
1548	Xanthopurpurin	1.7207	HMDB36006	4.238297851	8.824E-07
1554	Maritimetin	1.6661	HMDB29711	3.831818265	1.16E-06
1566	Colupulone	1.5088	HMDB30065	2.072756106	6.577E-06
1568	Isoamyl p-anisate	1.6679	HMDB38894	8.253368449	1.145E-05
1569	3-Galloylgallocatechin	1.6516	HMDB41598	11.82229424	1.97E-05
1581	Epigallocatechin gallate	1.5734	HMDB03153	14.721378	7.684E-05
1583	Chatenaytrienin 2	1.7391	HMDB32647	0.17768732	8.71E-05
1599	PC(22:5(7Z,10Z,13Z,16Z,19Z)/14:0)	1.512	HMDB08689	2.644207344	0.0002312
1631	Muridienin 1	1.4913	HMDB32648	0.306691067	0.0010024
1643	L-N-(1H-Indol-3-ylacetyl)glutamic acid	1.2654	HMDB38665	79.06382599	0.0018716
1656	PE(P-18:1(11Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	1.3807	HMDB11427	2.048096951	0.0026974
1668	Glycyrin	1.14	HMDB33712	40.05351494	0.0048644
1673	Cohibin B	1.1951	HMDB31169	2.160966146	0.0054013
1681	Indole-3-acetic-acid-Oglucuronide	1.1027	HMDB60001	160.0533848	0.0068926
1689	Glycerol triundecanoate	1.1031	HMDB31089	3.878060725	0.0086305
1696	Tazarotenic acid	1.056	HMDB60619	345.835058	0.0096407
1706	Ascorbigen	1.027	HMDB29839	170.7488883	0.0123314

tion of the significantly perturbed serum metabolites and the two-tailed Student's *t*-test were also employed using MetaboAnalyst. For the Student's *t*-test $P < 0.05$ and fold change of 2 were set as cutoff values.

The metabolites which were significantly different in the TB and normal samples were subjected to the MetaboAnalyst pathway analysis to investigate the pathways correlated with TB pathologic progress. The HMDB ID of each metabolite was used as an identifier.

Results

Metabolic profiling

Serum samples were obtained from 18 normal and 20 individuals with TB (**Table 1**) matched for age (normal = 42.6; TB = 42.1 years) and gender. A typical total ion chromatogram

Table 3. Six metabolic pathways relevant to TB

Pathway Name	Total	Hits	Details
Glycerophospholipid metabolism	39	3	KEGG
Arachidonic acid metabolism	62	2	KEGG SMP
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	14	1	KEGG
Linoleic acid metabolism	15	1	KEGG SMP
Sphingolipid metabolism	25	1	KEGG SMP
alpha-Linolenic acid metabolism	29	1	KEGG SMP
Galactose metabolism	41	1	KEGG SMP
Glyoxylate and dicarboxylate metabolism	50	1	KEGG
Tyrosine metabolism	76	1	KEGG SMP
Porphyrin and chlorophyll metabolism	104	1	KEGG

obtained by LC-MS for metabolomic profiling is shown in **Figure 1A** (positive ion mode) and **1B** (negative ion mode). After peak extraction, 1949 components were obtained from raw data. **Figure 1C** shows an extracted ion chromatogram for PC (16:1(9Z)/P-18:1(9Z)) (m/z 742.57642), and its tandem MS spectra are shown in **Figure 1D**. For metabolite identification, the exact m/z of the extracted peaks was utilized for web searching public databases (HMDB and LIPID MAPS) that contain information on known metabolites. Furthermore, the tandem MS spectra of metabolites (**Figure 1D**) can provide additional evidence for identification.

The intensities of the 1949 extracted metabolites were inspected using clustering algorithms (**Figure 2A**). For the heat map representation, the columns represent 38 serum samples, both from TB (green) and normal (red) subjects, whereas the rows represent different metabolites. Clustering was conducted using linkages corresponding to Pearson correlations. The heat map colors represent the intensity of the corresponding metabolite from low (blue) to high (red). Moreover, statistical analysis of the metabolomics profiling results revealed 175 significantly changed metabolites (red points in **Figure 2B**) using a fold change of 2 and *t*-test *p*-value of 0.05 as cutoffs. **Figure 2B** shows scatter plots of all the 1949 metabolites with a log 2 fold change (X) and -lg *p*-value (Y).

Multivariate statistical analysis

PCA analysis of the metabolomic profiles from the serum demonstrated a clear distinction between TB patients and controls subjects in both 2D and 3D modes (**Figure 2C** and **2D**). For an extensive investigation of the difference between the two groups, the data were subjected to PLS-DA to classify the samples (controls vs. TB) on the basis of metabolite concentration. A clear distinction between the two groups was also observed in the latter analysis (**Figure 3A** and **3B**), and cross validation showed no overfitting of the classification models (**Figure 3C**). To identify which variables were responsible for this difference, the VIP parameter was used to select variables that most significantly contributed in discriminating between the metabolomics profiles of the TB patients and controls in the PLS-DA model. The VIP score positively reflects the metabolite's influence on the classification, and metabolites with scores greater than 1.0 were considered important in this study.

After combining the statistical analyses and PLS-DA model VIP scores, the results were narrowed down to 99 significantly distinct metabolites between the TB and normal groups. Details of these metabolites are shown in **Table 2**. Furthermore, correlation matrices were generated from the 99 metabolites to create a compendium of metabolic interdependency for both the TB patients and controls (**Figure 3D**). The correlation plots reveal a wide range of correlation coefficients among all the 99 metabolites, ranging from 1.0 (positive correlation, red) to -1.0 (negative correlation, blue), with 0 indicating the absence of correlation.

Pathway analysis

The 99 significantly changed metabolites between the TB and control groups were subjected for pathway analysis using the of KEGG pathways database. In total, six metabolic pathways were obtained (**Table 3**; **Figure 4A**). **Figure 4B** shows information on one pathway, glycerophospholipid metabolism, in which three metabolites were noted to be changed in the TB group when compared with the control group.

Discussion

Metabolomics (also known as metabolic profiling) is a technique used to carry out qualitative

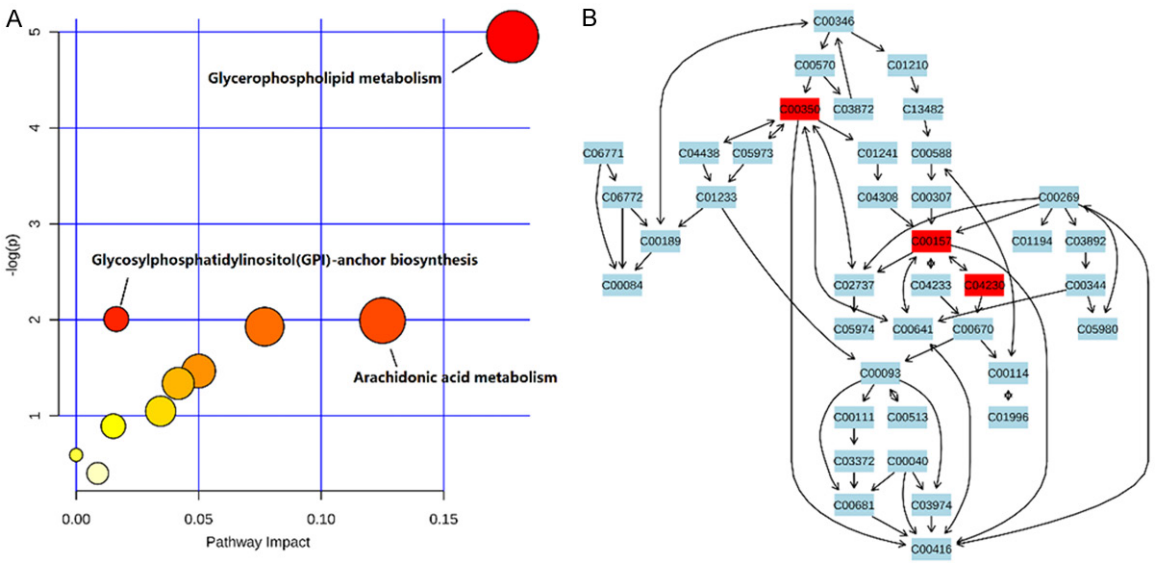


Figure 4. Pathway analysis using the database of KEGG pathways. A. The six metabolic pathways obtained. B. Information on the glycerophospholipid metabolism pathway.

and quantitative analyses of a large number of small molecular compounds in a sample. In recent years, the study of metabolic groups has been increasingly widely employed to search for biomarkers and has proven to be an effective method [16]. Potential biomarkers can be found for disease diagnosis and prognosis estimation but also can provide valuable information for the elucidation of disease mechanisms [17]. By using NMR, GC-MS, and LC-MS, metabolomics has been successfully used for biomarker investigation in coronary heart disease and hypertension [18], liver and epithelial ovarian cancer [19], type 2 diabetes [20], Huntington's disease [21], and schizophrenia [22]. In the current study, we used LC-MS to perform research on TB metabolomics. A total of 1949 compounds were detected, including 99 that showed remarkable differences in normal and TB samples, which may be explored as candidate biomarkers for future clinical validation.

Biomarkers are indicators of the diseased or normal state of the body as well as drug efficacy [23]. They can reflect disease status, evaluate the risk of deterioration, or provide information on the effectiveness and side effects of drug therapy [24]. Over the past few decades, studies have been carried out to discover biomarkers for *M. tuberculosis* and patient immune response [6]. The candidate biomarkers studied include *M. tuberculosis* DNA and RNA and antiphospholipid antibodies [25]. However, correctly judging disease condition has been difficult to achieve with the above biomarkers, often yields poor accuracy. Hence, the continued exploration for new TB biomarkers remains essential. In our metabolomics studies, 99 serum metabolites were significantly different between the two groups, as determined by statistical and multivariate statistical analyses. These metabolites are potential biomarkers, but further validation is needed for their progression to clinical practice.

Metabolomics is not only applied for studying biological markers but also for obtaining valuable information on disease pathologic and molecular mechanisms [26]. The 99 differential metabolites were subjected to metabolic pathway analysis, and the related metabolic pathways were obtained. As indicated in **Table 3** and **Figure 4A**, these different metabolites were involved in the metabolic pathways of glycerophospholipid and arachidonic acid. One of the main functions of glycerophospholipid is to serve as a structural component of the cell membrane. Notably, the metabolism of glycerophospholipids differs between eukaryotes and prokaryotes [27]. In prokaryotes, the synthesis of glycerophospholipids involves the synthesis of phosphatidic acid and polar head groups [28]. By contrast, the phosphatidic acid synthesis in eukaryotes involves two routes: one toward phosphatidylcholine synthesis and the

other toward phosphatidylethanolamine production [29]. Hence, a detailed study of the biosynthetic process and metabolic changes of glycerophospholipid metabolism in TB patients may be valuable and necessary for next-generation studies.

Conclusion

Serum serves as a liquid medium that interacts with every tissue and organ. It carries small molecules secreted and excreted by different tissues in response to pathologic progress. This study aimed to distinguish sera from TB and normal patients using LC-MS data in the context of metabolic profiling. Our approach demonstrated the importance of implementing multivariate statistical analysis and bioinformatics in determining the significantly changed metabolites and candidate disease biomarkers. Our results showed the broad-spectrum metabolite variation in TB that can be used for the diagnosis of this disease. Moreover, our findings suggest that the changes in serum glycerophospholipid levels serve as a potential molecular indicator for TB. However, this outcome must be further verified and validated in a larger number of clinical samples prior to application in the clinical setting.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Yuxin Yin, Department of Pathology, Institute of Systems Biomedicine, School of Basic Medical Sciences, Peking University Health Science Center, 38 Xueyuan Road, Haidian District, Beijing 100191, China. Tel: +86-10-82802676; E-mail: yinyuxin@bjmu.edu.cn

References

[1] Gomez JE and McKinney JD. M. tuberculosis persistence, latency, and drug tolerance. Tuberculosis 2004; 84: 29-44.

[2] Davis JL, Cattamanchi A, Cuevas LE, Hopewell PC, Steingart KR. Diagnostic accuracy of same-day microscopy versus standard microscopy for pulmonary tuberculosis: a systematic review and meta-analysis. Lancet Infect Dis 2013; 13: 147-54.

[3] Oxlade O, Huang CC, Murray M. Estimating the impact of reducing under-nutrition on the tuberculosis epidemic in the central eastern states of India: a dynamic modeling study. PLoS One 2015; 10: e0128187.

[4] Nahid P, Saukkonen J, Mac Kenzie WR, Johnson JL, Phillips PP, Andersen J, Bliven-Sizemore E, Belisle JT, Boom WH, Luetkemeyer A, Campbell TB, Eisenach KD, Hafner R, Lennox JL, Makhene M, Swindells S, Villarino ME, Weiner M, Benson C, Burman W; National Institutes of Health; Centers for Disease Control and Prevention. CDC/NIH Workshop. Tuberculosis biomarker and surrogate endpoint research roadmap. Am J Respir Crit Care Med 2011; 184: 972-9.

[5] Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, Wiegand R, Berriz GF, Roth FP, Gerszten RE. Metabolomic identification of novel biomarkers of myocardial ischemia. Circulation 2005; 112: 3868-3875.

[6] Wallis RS, Kim P, Cole S, Hanna D, Andrade BB, Maeurer M, Schito M, Zumla A. Tuberculosis biomarkers discovery: developments, needs, and challenges. Lancet Infect Dis 2013; 13: 362-72.

[7] Whittaker E, Gordon A, and Kampmann B. Is IP-10 a better biomarker for active and latent tuberculosis in children than IFNgamma? PLoS One 2008; 3: e3901.

[8] Bogdanov M, Matson WR, Wang L, Matson T, Saunders-Pullman R, Bressman SS, Flint Beal M. Metabolomic profiling to develop blood biomarkers for Parkinson's disease. Brain 2008; 131: 389-96.

[9] Malysheva SV, Arroyo-Manzanares N, Cary JW, Ehrlich KC, Vanden Bussche J, Vanhaecke L, Bhatnagar D, Di Mavungu JD, De Saeger S. Identification of novel metabolites from Aspergillus flavus by high resolution and multiple stage mass spectrometry. Food Addit Contam Part A Chem Anal Control Expo Risk Assess 2014; 31: 111-20.

[10] Zhang A, Sun H, Wang X. Serum metabolomics as a novel diagnostic approach for disease: a systematic review. Anal Bioanal Chem 2012; 404: 1239-45.

[11] Madsen R, Lundstedt T, Trygg J. Chemometrics in metabolomics--a review in human disease diagnosis. Anal Chim Acta 2010; 659: 23-33.

[12] Lin CY, Wu H, Tjeerdema RS and Viant MR. Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. Metabolomics 2007; 3: 55-67.

[13] Alder L, Greulich K, Kempe G, Vieth B. Residue analysis of 500 high priority pesticides: better by GC-MS or LC-MS/MS? Mass Spectrom Rev 2006; 25: 838-65.

[14] Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, Fung C, Nikolai L, Lewis M, Coutouly MA, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau DD, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan GE, Macinnis GD, Weljie

AM, Dowlatabadi R, Bamforth F, Clive D, Greiner R, Li L, Marrie T, Sykes BD, Vogel HJ, Querengesser L. HMDB: the human metabolome database. Nucleic Acids Res 2007; 35: D521-6.

[15] Xia J, Sinelnikov IV, Han B, Wishart DS. MetaboAnalyst 3.0--making metabolomics more meaningful. Nucleic Acids Res 2015; 43: W251-7.

[16] Kobayashi T, Nishiumi S, Ikeda A, Yoshie T, Sakai A, Matsubara A, Izumi Y, Tsumura H, Tsuda M, Nishisaki H, Hayashi N, Kawano S, Fujiwara Y, Minami H, Takenawa T, Azuma T, Yoshida M. A novel serum metabolomics-based diagnostic approach to pancreatic cancer. Cancer Epidemiol Biomarkers Prev 2013; 22: 571-9.

[17] Kaddurah-Daouk R, Kristal BS, Weinshilboum RM. Metabolomics: a global biochemical approach to drug response and disease. Annu Rev Pharmacol Toxicol 2008; 48: 653-83.

[18] Dunn WB, et al. Serum metabolomics reveals many novel metabolic markers of heart failure, including pseudouridine and 2-oxoglutarate. Metabolomics 2007; 3: 413-426.

[19] Zhang T, Wu X, Ke C, Yin M, Li Z, Fan L, Zhang W, Zhang H, Zhao F, Zhou X, Lou G, Li K. Identification of potential biomarkers for ovarian cancer by urinary metabolomic profiling. J Proteome Res 2013; 12: 505-12.

[20] Wang-Sattler R, Yu Z, Herder C, Messias AC, Floegel A, He Y, Heim K, Campillos M, Holzapfel C, Thorand B, Grallert H, Xu T, Bader E, Huth C, Mittelstrass K, Döring A, Meisinger C, Gieger C, Prehn C, Roemisch-Margl W, Carstensen M, Xie L, Yamanaka-Okumura H, Xing G, Ceglarek U, Thiery J, Giani G, Lickert H, Lin X, Li Y, Boeing H, Joost HG, de Angelis MH, Rathmann W, Suhre K, Prokisch H, Peters A, Meitinger T, Roden M, Wichmann HE, Pischon T, Adamski J, Illig T. Novel biomarkers for pre-diabetes identified by metabolomics. Mol Syst Biol 2012; 8: 615.

[21] Kaddurah-Daouk R, Krishnan KR. Metabolomics: a global biochemical approach to the study of central nervous system diseases. Neuropsychopharmacology 2009; 34: 173-186.

[22] Kaddurah-Daouk R, McEvoy J, Baillie RA, Lee D, Yao JK, Doraiswamy PM, Krishnan KR. Metabolomic mapping of atypical antipsychotic effects in schizophrenia. Mol Psychiatry 2007; 12: 934-945.

[23] Holly MK, Dear JW, Hu X, Schechter AN, Gladwin MT, Hewitt SM, Yuen PS, Star RA. Biomarker and drug-target discovery using proteomics in a new rat model of sepsis-induced acute renal failure. Kidney Int 2006; 70: 496-506.

[24] Taube SE, Clark GM, Dancey JE, McShane LM, Sigman CC, Gutman SI. A perspective on challenges and issues in biomarker development and drug and biomarker codevelopment. J Natl Cancer Inst 2009; 101: 1453-63.

[25] Walzl G, Ronacher K, Hanekom W, Scriba TJ, Zumla A. Immunological biomarkers of tuberculosis. Nat Rev Immunol 2011; 11: 343-54.

[26] Zhang GF, Sadhukhan S, Tochtrop GP, Brunengraber H. Metabolomics, pathway regulation, and pathway discovery. J Biol Chem 2011; 286: 23631-5.

[27] Bleijerveld OB, Houweling M, Thomas MJ, Cui Z. Metabolipidomics: profiling metabolism of glycerophospholipid species by stable isotopic precursors and tandem mass spectrometry. Anal Biochem 2006; 352: 1-14.

[28] Daleke DL. Phospholipid flippases. J Biol Chem 2007; 282: 821-5.

[29] Déchamps S, Shastri S, Wengelnik K, Vial HJ. Glycerophospholipid acquisition in Plasmodium - a puzzling assembly of biosynthetic pathways. Int J Parasitol 2010; 40: 1347-65.