

## Original Research Article

# Urine metabonomic study on primary liver cancer with spleen deficiency and excess dampness syndrome based on ultra-performance LC-MS

Li Jingjie, Liu Shumei, Wang Dong, Liu Jiang, Lv Shuqin\*

Traditional Chinese Medical Hospital of Xinjiang Uygur Autonomous Region, 116 Huanghe Road, Urumqi, Xinjiang 830000, China

\*For correspondence: **Email:** lvshuqin@126.com

Sent for review: 10 April 2020

Revised accepted: 20 February 2021

### Abstract

**Purpose:** To investigate spleen deficiency and excess dampness syndrome (SDES) in primary liver cancer (PLC) and the underlying mechanism using ultra pressure liquid chromatography-mass spectrometry (UPLC-MS).

**Methods:** Ultra-performance liquid chromatography coupled with mass spectrometry (UPLC-MS) was used to detect urine metabolites from untreated and IPED-treatment PLC-SDES patients. The metabolites were annotated using Kyoto Encyclopedia of Genes and Genomes (KEGG), Human Metabolome Database (HMDB), and Lipidmaps. Principle component analysis (PCA) and partial least squares to latent structure-discriminant analysis (PLS-DA) models were built to reveal the metabolic differences between untreated, IPED-treated patients and healthy controls. The differential metabolites in PLC-SDES patients were screened according to variables important in the project (VIP) and p-value.

**Results:** In urine, 537 metabolites (256 in negative and 281 in positive mode) were considered differential in PLC-SDES patients when compared to healthy controls. In untreated patients, 100 metabolites (38 in negative and 62 in positive mode) were differential when compared to IPED-treatment patients. The urine of PLC-SDES patients showed overlap of 32 metabolites.

**Conclusion:** The results reveal comprehensive urine metabonomic changes in PLC-SDES patients, relative to healthy controls and IPED-treated patients. The identified metabolites may be potential biomarkers for diagnosis and IPED therapy.

**Keywords:** Metabonomics, Liver cancer, Ultra performance liquid chromatography-mass spectrometry, UPLC-MS, Biomarker, Traditional Chinese medicine

This is an Open Access article that uses a fund-ing model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>) and the Budapest Open Access Initiative (<http://www.budapestopenaccessinitiative.org/read>), which permit unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

Tropical Journal of Pharmaceutical Research is indexed by Science Citation Index (SciSearch), Scopus, International Pharmaceutical Abstract, Chemical Abstracts, Embase, Index Copernicus, EBSCO, African Index Medicus, JournalSeek, Journal Citation Reports/Science Edition, Directory of Open Access Journals (DOAJ), African Journal Online, Bioline International, Open-J-Gate and Pharmacy Abstracts

## INTRODUCTION

Primary liver cancer (PLC) is one of the most malignant carcinomas in the world, and its occurrence is strongly associated with eating habits and lifestyles [1]. The mortality of liver

cancer is about 8.2 %, second to that of lung and colorectal cancer [2]. China has the highest liver cancer incidence worldwide and accounts for over 50 % of all newly diagnosed patients and deaths [3]. Populations in the underdeveloped western area of China in particular have the

highest incidence and mortality [4]. Xinjiang, located on the western border of China, shows a huge difference in day and night temperatures and has experienced water shortages for years. The theory of traditional Chinese medicine (TCM) holds that this lifestyle is often damaging to the spleen and stomach and supports the *endogenesis* of dampness syndrome. A survey conducted in 2003 showed that the frequency of "spleen deficiency and excess-dampness syndrome" (SDES) was about 79.07 % among patients with clinically diagnosed liver cancer in China [5]. In TCM theory, spleen deficiency will cause a stagnation of dampness, and the dampness will further aggravate spleen deficiency. The frequency of SDES occurrence in stage II and III liver cancer is often quite high [6]. The TCM treatment "invigorating spleen and eliminating dampness" (IPED) is regarded as an effective therapy for SDES [7]. However, the biological mechanisms of IPED treatment on SDES in liver cancer are still unclear.

Metabonomics is a vital part of systematic biology which focuses on dynamic changes in biological components [8]. The experimental methods used in metabonomics include nuclear magnetic resonance (NMR), gas chromatography-mass spectrometry (GC-MS), and liquid chromatography-mass spectrometry (LC-MS) [9]. Ultra-performance liquid chromatography coupled with mass spectrometry (UPLC-MS) is a powerful technique for metabolomic studies, because it allows metabolites to be tested in liquids such as urine [10]. Typically, urine contains a diverse range of polar metabolites, and LC-MS-based methods can rapidly detect these components in urine samples. Chen reported that hydrophilic interaction chromatography can detect metabolites in urine more efficiently than reverse-phase liquid chromatography [11]. In the present study, untargeted ultra-performance liquid chromatography combined with mass spectrometry (UPLC-MS) was used to study the effects of IPED on urine metabolites in patients with SDES-PLC. Principle component analysis (PCA) and partial least squares to latent structure-discriminant analysis (PLS-DA) were used to identify significant metabolites associated with IPED.

## EXPERIMENTAL

### Patients

Fifty-five patients with SDE hepatocellular carcinoma were randomly recruited from the Affiliated Hospital of Traditional Chinese Medicine of Xinjiang Medical University. Among

them, twenty-four patients who had received IPED treatment served as the IPED treatment group, and the remaining thirty-one untreated patients served as the untreated group. All patients were diagnosed with PLC according to the western medicine diagnosis standard, and with SDES according to the traditional Chinese medicine diagnosis, and all were diagnosed with PLC as the first diagnosis. In addition, twenty-eight healthy persons were recruited as a healthy control group from the physical examination center of Affiliated Hospital of Chinese medicine of Xinjiang Medical University. Each patient and healthy volunteer provided written consent for participation in this study.

### Sample collection and preparation

A 5 mL volume of morning urine was collected from each patient and healthy control and stored in Eppendorf (EP) tube at room temperature after adding 500  $\mu$ L sodium azide (1 mmol/L). The urine samples were cooled on ice and then centrifuged at 5440 g for 10 min at 20 °C. The solid impurities were discarded, and a 2 mL of the urine supernatant was kept in an EP tube. A 100  $\mu$ L sample of urine supernatant was transferred to a new EP tube, and 400  $\mu$ L pure methanol was added to precipitate proteins. The EP tubes were vortexed, placed on an ice bath for 5 min, and then centrifuged at 25000 g for 10 min at 4 °C. The supernatant was diluted using 60 % methanol with MS grade water, and filtered through a 0.22  $\mu$ m filter membrane using a 10 min centrifugation at 15000 g at 4 °C. The filtrate was collected and analyzed using LC-MS. Blank control samples were generated using 60% methanol containing 0.1 % formic acid.

### QC sample preparation

Equal volumes of urine samples from each group were mixed to generate a quality control (QC) sample. The mixed QC sample was treated using the same preparation method described for the test samples. The pooled QC sample was used to validate the study method.

### UPLC-MS analysis

UPLC-MS data was analyzed using a Vanquish UHPLC system (Thermo Fisher) coupled to an Orbitrap Q Exactive HF-X mass spectrometer (Thermo Fisher) operating in the data-dependent acquisition (DDA) mode. Samples were injected onto an Accucore HILIC column (100  $\times$  2.1 mm, 2.6  $\mu$ m) using a 16-min linear gradient at a flow rate of 0.2 mL/min. The mass spectrometer was operated in positive/negative polarity mode with a spray voltage of 3.2 kV, capillary temperature

of 320 °C, sheath gas flow rate of 35 arb, and auxiliary gas flow rate of 10 arb.

### Metabonomic validation method

The precision and repeatability of the experimental instruments were tested and monitored by analyzing 9 QC samples by UPLC-MS/MS. Three QC samples were used to monitor instrument status and to balance the chromatograph-mass spectrum before the injection of test samples. Another 3 QC samples were analyzed to evaluate the systematic stability of the entire experimental procedure during the analysis of urine samples. The last 3 QC samples were used for qualitative analysis of metabolites using secondary mass spectrometry. Pearson correlation coefficient analysis and principle components analysis (PCA) for the QC samples were performed using the peak area values.

### Metabolite identification and classification

The raw data files generated by UHPLC-MS/MS were processed using Compound Discoverer 3.0 (CD3.0, Thermo Fisher) to perform peak alignment, peak picking, and quantitation for each metabolite. The experimental conditions were set as follows: retention time tolerance, 0.2 min; actual mass tolerance, 5 ppm; signal intensity tolerance, 30%; signal/noise ratio, 3; and minimum intensity, 100000. The peak intensities were normalized to the total spectral intensity, and the normalized data were used to predict the molecular formula based on additive ions, molecular ion peaks, and fragment ions. The peaks were then matched with the mzCloud (<https://www.mzcloud.org/>) and ChemSpider (<http://www.chemspider.com/>) databases to obtain accurate qualitative and relative quantitative results. Statistical analyses were performed using the statistical software R (version R-3.4.3), Python (version 2.7.6), and CentOS (CentOS release 6.6). When data were not normally distributed, normal transformations were attempted using the area normalization method. The classifications and pathway annotations of the identified metabolites were derived using the Kyoto Encyclopedia of Genes and Genomes (KEGG), Human Metabolome Database (HMDB) and Lipidmaps databases [12-14].

### Discrimination model building and validating

Differential metabolites were screened using principal component analysis (PCA), and partial least squares discriminant analysis (PLS-DA) models were established using metaX (a flexible

and comprehensive software for processing metabolomics data) [15]. The established model was then assessed by Y-scrambling statistical validation to test the possibility of a chance correlation when the class membership was randomly shuffled 200 times, and the parameters for model fitness ( $R^2$ ) and predictive ability ( $Q^2$ ) were calculated [16]. The  $Q^2$  value was expected to be lower than  $R^2$ , which would suggest that the models were not over-fitted.

### Differential metabolite screening

The univariate analysis (t-test) was applied to calculate the statistical significance (P-value), and the variable importance in the project (VIP) values of metabolites were calculated from the established PLS-DA models. A larger VIP value indicates a more significant metabolite in model [17]. Metabolites with  $VIP > 1$  and  $p < 0.05$  and fold change (FC)  $\geq 2$  or  $FC \leq 0.5$  were considered differential metabolites. Volcano plots were used to demonstrate filtered metabolites of interest according to the  $\text{Log}_2$  (FC) and  $-\text{log}_{10}$  (P value) determined for the metabolites.

### Clustering and KEGG enrichment analysis

For clustering heat maps, the metabonomic data were normalized using z-scores of the intensity areas of differential metabolites and were plotted with the Pheatmap package in the R language. The correlations between differential metabolites were analyzed using the `cor()` function in the R language (method = Pearson). Statistically significant correlations between differential metabolites were calculated with the `cor.mtest()` function in the R language. A  $p < 0.05$  was considered statistically significant, and correlation plots were plotted with the `corrplot` package in the R language.

### Data analysis

The functions of the differential metabolites and metabolic pathways were annotated using the KEGG database. The metabolic pathway enrichment of differential metabolites were determined. When the ratio satisfies the relation  $x/n > y/N$ , a metabolic pathway was considered enriched. The KEGG pathways with  $p < 0.05$  were considered as statistically significantly enriched.

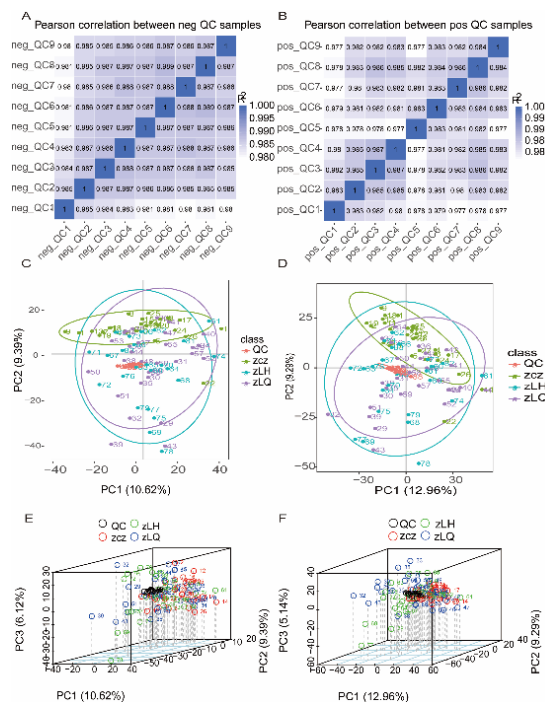
## RESULTS

### Validated metabonomic method

Metabolic components can undergo interferences and changes due to external

factors. Therefore, a QC procedure is necessary to validate the repeatability and precision of metabolomic results. Figure 1 shows the correlations between different QC samples, reflected by the R2 value. A larger R2 (ranging from 0 to 1) means a stronger correlation between two samples, as well as better data quality. In this study, the R2 in negative mode ranged from 0.98 to 1.00, while the R2 in positive mode ranged from 0.977 to 1.00.

The principle components analysis (PCA) was performed using the peak calls of all tested and QC samples. The PCA result of all samples showed that the QC samples were gathered intently, confirming the stability of the metabolomic method based on UPLC-MS (Figure 1).

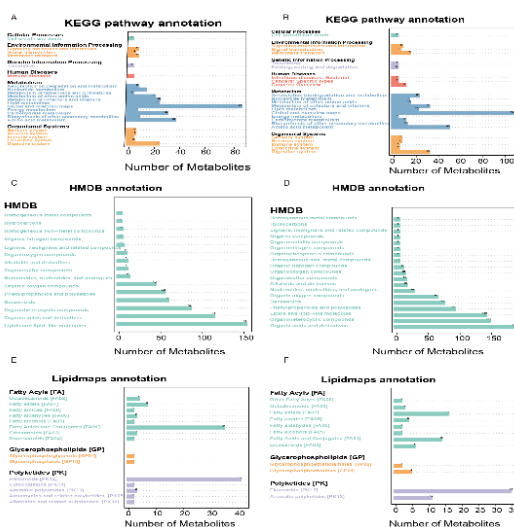


**Figure 1:** Correlation plots of QC samples in negative mode (A), and in positive mode (B). PCA score plots of urine in 2D negative mode (C), in 3D negative mode (E), in 2D positive mode (D), and in 3D positive mode (F). In negative mode, PC1 represents 10.62 % variance, PC2 represents 9.39 % and PC3 represents 6.12 %; in positive mode, PC1 represents 12.96 %, PC2 represents 9.29 % and PC3 represents 5.14 %. QC: quality control, zcz: group of healthy controls, zlh: group of IPED-treatment patients, zlq: group of untreated patients

### Metabolite classification and annotation

The metabolites identified from the urine metabolomic profile were classified and annotated using KEGG, HMDB, and Lipidmaps

analysis (Figure 2). The most important categories of metabolites are listed in Table 1. The KEGG analysis for the negative mode identified the top three metabolic pathways, annotated by the largest number of metabolites, as “Global and overview maps,” “Amino acid metabolism,” and “Carbohydrate metabolism”; for the positive mode, these top pathways were “Global and overview maps,” “Amino acid metabolism,” and “Metabolism of cofactors and vitamins.” The HMDB results for the negative mode indicated that the largest number of metabolites, at 151, belonged to “Lipids and lipid-like molecules”; for the positive mode, the largest number of metabolites, at 186, belonged to “Organic acids and derivatives.” The Lipidmaps analysis revealed that the largest number of lipid metabolites belonged to flavonoids for both negative and positive modes.



**Figure 2:** KEGG annotations of metabolites in negative mode (A) and in positive mode (B). HMDB annotations of metabolites in negative mode (C) and in positive mode (D). Lipidmaps annotations of metabolites in negative mode (E) and in positive mode (F).

### Establishment discrimination model and validation

The spatial distribution and its two-dimensional projection of PCA analysis between the untreated PLC-SDES and the healthy control groups, and between the untreated and the IPED-treatment PLC-SDES group, are shown in Figure S1 and Figure S2, respectively. Figure S1 shows the 95 % confidence intervals between the untreated group and healthy control group and indicates partially overlapped samples in the two-dimensional projection.

**Table 1:** Top three categories of metabonomic profile annotated by KEGG, HMDB and Lipidmaps

Database	Category	Number of metabolites in negative	Number of metabolites in positive
KEGG	Global and overview maps	84	106
	Amino acid metabolism	35	48
	Metabolism of cofactors and vitamins	23	30
HMDB	Lipids and lipid-like molecules	151	138
	Organo-heterocyclic compounds	84	145
	Organic acids and derivatives	112	186
Lipidmaps	Flavonoids	40	34
	Fatty acids and Conjugates	34	13
	Fatty esters	6	15

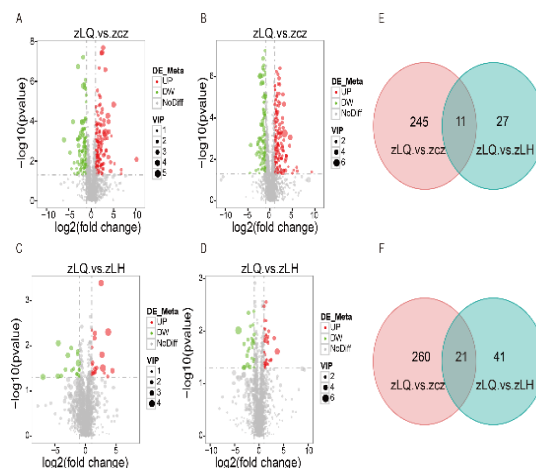
However, the 3D distribution plot shows differences in the metabolic components between the two groups, for metabolites analyzes in both the negative and positive modes. Figure S2 shows that the 95 % confidence intervals between the untreated and IPED-treatment groups were totally overlapped, suggesting a lesser significance of the metabonomic difference between these two groups than between the untreated and healthy control groups.

The sample scores in the PLS-DA models plotted in Figure S3 show a good separation of the confidence intervals of three groups, suggesting that the PLS-DA models built for this study are effective for screening differential metabolites. The results of Y-scrambling validation tests for the PLS-DA models are shown in Figure S4. The model fitness ( $R^2$ ) was higher than the predictive ability ( $Q^2$ ) for all comparisons between groups.

### Differential metabolite screening

Differential metabolites were screened according to the VIP values of the first principle component in the PLS-DA models and the fold changes combined with the P-values. The volcano plots in Figure 3 show the numbers and regulation directions of differential metabolites from each comparison between groups. Compared with healthy controls, 256 of the 2067 negative mode metabolites and 281 of the 2839 positive mode metabolites in the untreated patients were considered differential (See Table S1). Compared with IPED-treatment patients, 38 of the 2067 negative mode metabolites and 62 of the 2839 positive mode metabolites in the untreated patients were considered differential (See Table S2). An overlap was detected for 11 negative mode and 21 positive mode differential metabolites between the two pairwise comparisons, while 27 negative mode and 41 positive mode differential metabolites were unique in the discrimination of untreated and IPED-treatment patients. The overlapped differential metabolites and their regulation

directions are shown in Table 2. Norethisterone acetate, which was up-regulated, had the highest VIP value among the unique differential metabolites in the comparison between the untreated and healthy control groups. Sulfosalicylic acid, which was also up-regulated, had the highest VIP in a comparison of the IPED-treatment and untreated groups.



**Figure 3:** Volcano plots of differential metabolites identified from comparison between untreated group and healthy control group in negative mode (A), and in positive mode (B). Volcano plots of differential metabolites identified from comparison between IPED-treatment group and untreated group in negative mode (C), and in positive mode (D). Venn plots for pairwise comparisons of untreated patients vs healthy controls and untreated patients vs IPED-treatment patients in negative mode (E), and in positive mode (F). Note: The abscissa represents the log of fold change ( $\log_2(\text{fold change})$ ) of metabolites, the ordinate represents the difference significance level ( $-\log_{10}(P\text{-value})$ ). Each point in the volcanic map represents a metabolite; significantly up-regulated metabolites are colored in red, the significantly down-regulated metabolites are colored in green, and the size of the dot represents VIP value. zcz: group of healthy controls, zLH: group of IPED-treatment patients, zLQ: group of untreated patients, UP: up-regulated, DW: down regulated, NoDiff: no difference

**Table 2:** Overlapped differential metabolites identified in untreated patients vs healthy controls and vs IPED-treatment patients

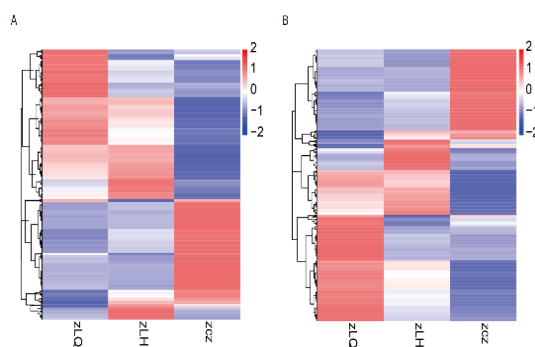
Polarity	Name	MW	VIP	Up/Down
Negative	(Z)-desulfoglucotropeolin	329.09	2.53	Down
Negative	Cefcapene	453.08	3.91	Up
Negative	MFCD24849356	299.11	1.89	Up
Negative	indole-3-propanol phosphate	255.07	1.55	Up
Negative	MFCD00210294	374.12	1.90	Down
Negative	epsilon-(gamma-Glutamyl)-lysine	275.15	3.60	Down
Negative	2H-1,4-Benzoxazin-3(4H)-one	149.05	2.09	Down
Negative	Pyridoxine phosphate	249.04	3.60	Down
Negative	Narcotoline	399.13	3.79	Down
Negative	2-(2-Amino-2-carboxyethyl)-5-oxotetrahydro-2-furancarboxylic acid	217.06	2.48	Down
Negative	Cycloxazoline	546.32	2.45	Up
Positive	Vernadigin	564.29	2.44	Up
Positive	Tenofovir	287.08	4.90	Up
Positive	QK4247000	307.10	2.39	Down
Positive	Cefpodoxime	427.06	3.17	Down
Positive	Rosmarinine	353.18	2.33	Down
Positive	(1S,6S)-6-Aminooctahydro-1-indoliziny acetate	198.14	2.96	Up
Positive	Sulfurol	143.04	1.52	Up
Positive	12,22-Dihydro-2H,3H-porphine	312.14	1.70	Up
Positive	Amylose	370.15	1.99	Down
Positive	Aloin B	418.13	3.05	Down
Positive	Roxatidine	306.19	6.39	Down
Positive	Tributyl citrate acetate	402.22	1.67	Up
Positive	Phenanthrene	178.08	1.99	Up
Positive	MFCD00059633	244.20	2.46	Up
Positive	17(S)-HETE	320.23	2.35	Up
Positive	gitogenin	432.32	1.58	Up
Positive	Cefdinir	395.04	3.92	Down
Positive	1,3-Di-tert-butylbenzene	190.17	2.05	Up
Positive	(-)-4'-Demethylepipodophyllotoxin	400.12	2.37	Down
Positive	1,5-Diisocyanatonaphthalene	210.04	1.69	Up
Positive	Cyclopentolate	291.18	2.28	Down

**MW:** molecular weight, **VIP:** variable importance in the project

### Clustering and KEGG enrichment analysis

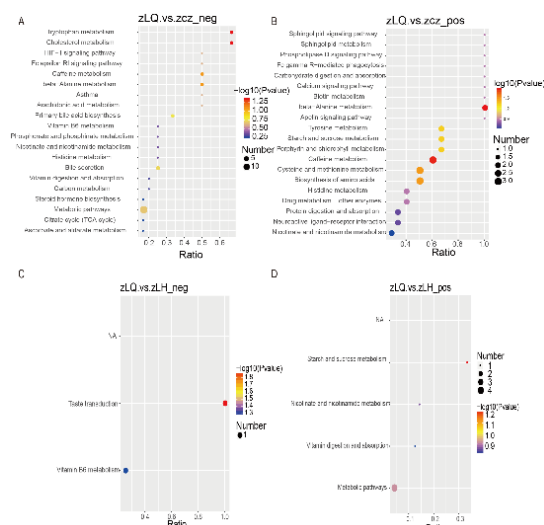
The clustering analysis unambiguously showed that the metabolic pattern of urine components in PLC-SDES patients differed from that of healthy controls (Figure 4). The details of tendencies in metabolic changes for each metabolite are shown in Figure S5. The significant KEGG pathways showing metabolite enrichment are plotted in Figure 5. The P-values identify the significantly enriched metabolic pathways in untreated patients vs. healthy controls as “Caffeine metabolism,” “beta-alanine metabolism,” “Tryptophan metabolism,” “Cholesterol metabolism,” “Cysteine and methionine metabolism,” “Biosynthesis of amino acids,” “Tyrosine metabolism,” “Starch and sucrose metabolism,” and “Porphyrin and chlorophyll metabolism.” “Caffeine metabolism” and “beta-Alanine metabolism” were the only two pathways that showed simultaneous involvement of metabolites in both the negative and positive modes. The significant pathway enriched with differential metabolites between the untreated

patients and IPED-treatment patients was “Taste transduction” (Table 3).



**Figure 4:** Heatmap for metabolite clustering in negative mode (A), in positive mode (B). Each grid in the heatmap denotes a up-regulated (red color) or down-regulated (blue color) metabolite. The left vertical axis represents the metabolite clusters while the horizontal axis represents group names. zCZ: group of healthy controls, zLH: group of IPED-treatment patients, zLQ: group of untreated patients





**Figure 5:** KEGG enrichment dot plots for differential metabolites from untreated patients vs healthy controls in negative mode (A), in positive mode (B), and from untreated patients vs IPED-treatment patients in negative mode (C), in positive mode (D). Note: Each dot denotes a KEGG pathway, of which the size denotes number of enriched metabolites. zcz: group of healthy controls, zLH: group of IPED-treatment patients, zLQ: group of untreated patients

## DISCUSSION

Metabonomics studies are increasingly improving research into the pathenogenesis of various diseases by identifying and summarizing internal rules for metabolic component changes in organisms [18]. The syndrome theory used in TCM has a similar aim, namely to reveal the correlations between abnormalities of metabolic networks and factors like diet, sleeping, season changes, and life circumstances [19]. However, the syndrome traits of TCM only reflect the nature of disease development in a certain period. Thus, the results suffer from timeliness and fuzziness. In addition, due to the limited

levels of knowledge and clinical experience, each TCM clinician has a different understanding about the same patient and makes different judgments of syndrome types. Thus, in the present study, common metabonomic methods, including UPLC-MS coupled with PCA analysis and the use of PLS-DA models, were used to improve the understanding of *spleen deficiency* and *excess-dampness syndrome* in primary liver cancer patients. The results of PCA and PLS-DA models, based on the UPLC-MS data shown in Figure S1, S2, and S3, indicated that the PLC-SDES patients were well separated from the healthy controls. This finding suggests that the use of UPLC-MS to detect urine metabolic component changes can help in clinical diagnosis for further validation of SDES.

In the present study, the metabolites in urine of a Xinjiang population were annotated and classified using the KEGG, HMDB, and Lipidmaps databases. The KEGG pathway indicated a strong enrichment of several metabolism-related pathways. The HMDB annotation showed that most metabolites that changed frequently in urine were organic acids and their derivatives, lipids and lipid-like molecules, and organoheterocyclic compounds. The Lipidmaps results suggested that flavonoids were the most enriched urine metabolic compounds.

Flavonoids cannot be absorbed by the human body due to the presence of glycosides [20]. In addition, the results showed that the urinary excretion of flavonoid metabolites was higher in Xinjiang population. Dietary flavonoids are metabolized by both primary and secondary metabolism. The flavonoids in blood are transported to the liver through the portal veins, followed by excretion via the urinary system [21].

**Table 3:** Significant KEGG pathways enriched by differential metabolites in untreated patients vs healthy controls and in untreated patients vs IPED-treatment patients

Group	Map Title	P-value	Polarity
Untreated vs healthy controls	Caffeine metabolism	0.082	Negative and positive
Untreated vs healthy controls	beta-Alanine metabolism	0.082	Negative and positive
Untreated vs healthy controls	Tryptophan metabolism	0.045	Negative only
Untreated vs healthy controls	Cholesterol metabolism	0.045	Negative only
Untreated vs healthy controls	Cysteine and methionine metabolism	0.033	Positive only
Untreated vs healthy controls	Biosynthesis of amino acids	0.033	Positive only
Untreated vs healthy controls	Tyrosine metabolism	0.048	Positive only
Untreated vs healthy controls	Starch and sucrose metabolism	0.048	Positive only
Untreated vs healthy controls	Porphyrin and chlorophyll metabolism	0.048	Positive only
Untreated vs IPED-treatment	Taste transduction	0.013	Negative only

Enzymes of secondary metabolism may catalyze the modification of flavonoids with glucuronide, sulfate and methyl moieties, resulting in the accumulation of flavonoids in the liver [22]. Thus, an extremely high content of flavonoids may increase the metabolic burden of the liver. Thus, patients should be advised to decrease their intake of dietary flavonoids.

Many metabolites in the urine of PLC-SDES patients showed abnormal expression according to the VIP values and fold changes. The criteria of  $VIP > 1$ ,  $|\log_2(FC)| > 2$ , and  $P\text{-value} < 0.05$  were used to compare the differential metabolites in untreated patients with those in healthy people or in IPED-treatment patients. The finding of norethisterone acetate, a synthetic derivative of progestogen commonly used in therapy for contraception and ovulation inhibition, and sulfosalicylic acid, a derivative of salicylic acid that acts as a chelating agent for iron ions [23], suggests that patients with PLC-SDES may have a risk of exposure to drug residues, since neither of these two compounds can be synthesized by the human body.

The further identification of 32 overlapped differential metabolites (21 positive and 11 negative mode) reveals their potential for use as biomarkers for diagnosis of PLC-SDES and for IPED treatment because they showed stable differences in patients before and after IPED treatment.

These differential metabolites indicated significant effects on several metabolic pathways related to chemical molecules including caffeine,  $\beta$ -alanine, and tyrosine, in PLC-SDES patients. The primary metabolites of caffeine, which include paraxanthine, theobromine, and theophylline, are biologically active and are metabolized in the liver. Damage to the caffeine metabolic pathway to varying degrees has been reported in patients with liver diseases such as cirrhosis and hepatitis B or C [24].  $\beta$ -Alanine is spontaneously produced by the liver and serves an important component of vitamin B5 and carnosine.

Tyrosine is related to signaling pathways which may play a vital role in liver oncogenesis. For example, the PI3K/AKT signaling pathway, which uses a tyrosine kinase receptor as a factor for cascade amplification, is impaired in hepatocellular carcinoma [25]. Generally, metabolic disorders involving alkaloids and amino acids occur in PLC-SDES patients. After IPED treatment, the number of differential metabolites decreased, indicating a reduction in the extent of the disorder.

## CONCLUSION

This is the first study to reveal systematic changes in urine metabonomic components in primary liver cancer patients with SDES. Qualitative and quantitative methods based on UPLC-MS were confirmed effective and sufficiently stable for the detection of urine metabolites in patients. Multivariate analysis, including PCA and PLS-DA, were able to discriminate PLC-SDES patients from healthy controls. In all, 537 metabolites in urine were identified as significantly different between PLC-SDES patients and healthy controls. After treatment with IPED, the number of differential metabolites decreased to 100. The VIP value, fold change, and P-value revealed 21 differential metabolites that could serve as potential biomarkers for PLC-SDES diagnosis and indicators for IPED treatment in a Xinjiang population.

## DECLARATIONS

### Acknowledgement

This study was supported by National Natural Science Foundation of China (grant no. 81460689).

### Conflict of interest

No conflict of interest is associated with this work.

### Contribution of authors

This study was done by the authors named in this article and all liabilities pertaining to claims relating to the content of this article will be borne by the authors. Li Jingjie should be considered as the first author. Lv Shuqin is the corresponding author. Liu Shumei, Wang Dong and Liu Jiang are the second, third and fourth authors respectively.

### Open Access

This is an Open Access article that uses a funding model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>) and the Budapest Open Access Initiative (<http://www.budapestopenaccessinitiative.org/read>), which permit unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.



## REFERENCES

- Bosch FX, Ribes J, Mireia Díaz, Cléries R. Primary liver cancer: Worldwide incidence and trends. *Gastroenterol* 2004, 127(5-suppl-S1): S5-S16.
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. *Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. CA: A Cancer J Clinician 2018, 68(6): 394-424.
- Allemani C, Weir HK, Carreira H, Harewood R, Spika D, Wang XS, Bannon F, Ahn JV, Johnson CJ, Bonaventure A, et al. Global surveillance of cancer survival 1995–2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). *The Lancet* 2015, 385(9972): 977-1010.
- Zheng R, Qu C, Zhang S, Zeng H, Sun K, Gu X, Xia C, Yang Z, Li H, Wei W, et al. Liver cancer incidence and mortality in China: Temporal trends and projections to 2030. *Chinese J Cancer Res* 30(6): 571-579.
- Pan M, Zeng P, Pan B. Study on the rule of Chinese medicine in the treatment of advanced primary liver cancer. *Study J Trad Chinese Med* 2003, 21(10): 1641-1642. (in Chinese)
- Yarong L, Jingjie L, Shuqin L. Plasma Metabonomics Analysis of Liver Cancer Patients of Damp Abundance Due to Spleen Asthenia Syndrome Based on Nuclear Magnetic Resonance, 2016. (in Chinese)
- Su X. Modern research on dampness. *Jiangxi Trad Chinese Med* 2004; (4): 63-64. (in Chinese)
- Nicholson JK, Lindon JC, Holmes E. "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 1999, 29(11): 1181-1189.
- Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, Brown M, Knowles JD, Halsall A, Haselden JN. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols* 2011; 6(7): 1060-1083.
- Want EJ, Masson P, Michopoulos F, Wilson ID, Theodoridis G, Plumb RS, Shockcor J, Loftus N, Holmes E, Nicholson JK. Global metabolic profiling of animal and human tissues via UPLC-MS. *Nature Protocols* 2012; 8(1): 17-32.
- Chen J, Wang W, Lv S, Yin P, Zhao X, Lu X, Zhang F, Xu G. Metabonomics study of liver cancer based on ultra-performance liquid chromatography coupled to mass spectrometry with HILIC and RPLC separations. *Analytica Chimica Acta* 2009; 650(1): 3-9.
- Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 1999, (1): 29-34.
- Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, et al. HMDB: the Human Metabolome Database. *Nucleic Acids Res*. 2007, 35: D521-D526.
- Fahy E, Subramaniam S, Murphy RC, Nishijima M, Raetz CR, Shimizu T, Spener F, van Meer G, Wakelam MJ, Dennis EA. Update of the LIPID MAPS comprehensive classification system for lipids. *Journal of Lipid Research*, 2009, 50: S9-14.
- Wen B, Mei Z, Zeng C, Liu S. metaX: a flexible and comprehensive software for processing metabolomics data. *BMC Bioinformatics*, 2017, 18(1): 183-196.
- Chang D, Weljie A, Newton J. Leveraging latent information in NMR spectra for robust predictive models. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing*, 2007, 12(4): 115-126.
- Lappalainen M, Jokkala J, Juutilainen A, Hämäläinen S, Koivula I, Jantunen E, Hanhineva K, Pulkki K. Novel Biomarker Candidates for Febrile Neutropenia in Hematological Patients Using Nontargeted Metabolomics. *Disease Markers*, 2018, 2018: 6964529.
- Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 2004, 5(2): 101-113.
- Tang JL, Liu BY, Ma KW. Traditional Chinese Medicine. *The Lancet*, 2008, 372(9654): 1938-1940.
- Erlund I. Review of the flavonoids quercetin, k hesperetin, and naringenin. Dietary sources, bioactivities, bioavailability, and epidemiology. *Nutrition Research*, 2004, 24(10): 851-874.
- Aedin Cassidy, Minihane AM. The role of metabolism (and the microbiome) in defining the clinical efficacy of dietary flavonoids. *American Journal of Clinical Nutrition*, 2016, 105(1): 10-22.
- Bolling BW, Court MH, Blumberg JB, Chen CY. The kinetic basis for age-associated changes in quercetin and genistein glucuronidation by rat liver microsomes. *Journal of Nutritional Biochemistry*, 2010, 21(6): 498-503.
- Liu Q, Du K, Liu M, Lv R, Sun B, Cao D, He N, Wang Z. Sulfosalicylic acid/Fe<sup>3+</sup> based nanoscale coordination polymers for effective cancer therapy by the Fenton reaction: an inspiration for understanding the role of aspirin in the prevention of cancer. *Biomater Sci*. 2019, 7(12): 5482-5491.
- Nehlig A, Alexander SPH. Interindividual Differences in Caffeine Metabolism and Factors Driving Caffeine Consumption. *Pharmacological Reviews*, 2018, 70(2): 384-411.
- Liao J, Jin H, Li S, Xu L, Peng Z, Wei G, Long J, Guo Y, Kuang M, Zhou Q, et al. Apatinib potentiates irradiation effect via suppressing PI3K/AKT signaling pathway in hepatocellular carcinoma. *J Exp Clin Cancer Res*. 2019, 38(1): 454-466.