International Journal of Hygiene and Environmental Health 222 (2019) 434–445

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

journal homepage: www.elsevier.com/locate/ijheh

Comprehensive analysis of the metabolomic characteristics on the health lesions induced by chronic arsenic exposure: A metabolomics study☆

Chaonan Jia[a,b,c,1], Yaping Wei[a,c,1], Yuan Lan[d], Xiangqing Hou[a,c], Jingjing Zuo[d], Tao Wang[a,c], Jushuang Li[a,c], Xiaoju Guan[e], Hui Yang[f], Guangyun Mao[a,c,e,*]

[a] Department of Preventive Medicine, School of Public Health & Management, Wenzhou Medical University, Wenzhou, Zhejiang, China
[b] Taizhou Hospital of Zhejiang Province, Taizhou, Zhejiang, China
[c] Center on Evidence-Based Medicine & Clinical Epidemiological Research, School of Public Health and Management, Wenzhou Medical University, Wenzhou, Zhejiang, China
[d] Center on Clinical Research, The Eye Hospital of Wenzhou Medical University, Wenzhou, Zhejiang, China
[e] Center for Scientific Research, The Second Affiliated Hospital, Wenzhou Medical University, Wenzhou, Zhejiang, China
[f] Department of Medical Affairs, No.118 Hospital of PLA (People's Liberation Army), Wenzhou, Zhejiang, China

ARTICLE INFO

Keywords:
Health lesion
Arsenic
Metabolomics
Biomarker

ABSTRACT

Early detection of the health lesions induced by chronic arsenic exposure (HLICAE) are crucial to prevent permanent arsenic-induced damage. If HLICAE can be identified in time, appropriate preventive and therapeutic measures may be provided without various avoidable lesions. The present study aims to assess the probability of HLICAE early recognition with metabolomics. Applying a case-control study, 94 participants with HLICAE (cases) and other 94 subjects without HLICAE (controls) were matched with gender and age (± 1 year), coming from a previous chronic arsenic exposure cohort. Serum metabolomic profiles were assessed by ultra-high-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC-QTOF-MS) and analyzed with univariate and multivariate statistics. A total of 210 and 364 features were detected in positive and negative ion modes (ESI^+/ESI^-), respectively. The altered metabolic pathways included lipid and amino acid metabolisms. 28 metabolomics-based biomarkers were significantly associated with HLICAE and provided areas under the curve (AUC, 95% confidence interval) of 0.898 (0.836, 0.960) and 0.908 (0.855, 0.960) in the discovery phase, 78.6% and 86.4% of positive predictive values in the validation phase, in distinguishing HLICAE from controls in ESI^+/ESI^-, respectively. This study provides novel insights on mechanisms of health effects probably induced by chronic arsenic exposure, and these biomarkers may be applied in HLICAE early detection.

1. Introduction

Inorganic arsenic (iAs) exposure is a globally public concern due to its health damage and wide distribution in the natural environment, such as drinking water, food, soil, minerals, applying to agricultural and industrial production (Hughes et al., 2011; Matschullat, 2000). In accordance with a previous study, chronic arsenic exposure has already affected more than 200 million people in at least 70 countries worldwide, including China (Minatel et al., 2018). Abundant epidemiologic evidence indicates that iAs exposure is associated with increased risks for multiple health lesions, comprising cancers of skin, lung and others, as well as cardiovascular diseases and metabolic syndrome (Chen et al., 2012a; Dodson and Zhang, 2016; Lin et al., 2002; Wu et al., 2004).

The deleterious effects induced by environmental arsenic exposure have been extensively studied. Most of them are chronic diseases with a quite long course between arsenic exposure and the occurrence of its typical symptoms and signs. At present, most of HLICAE are diagnosed by kinds of skin lesions. This can't accurately reflect the occurrence and development of HLICAE in a timely manner because it is asymptomatic in its early stage and can progress unnoticed by the patient. Once diagnosed, patients will lose the best treatment opportunity, making it difficult for them to obtain better therapeutic effects. Therefore, making early identification, diagnosis and providing timely appropriate treatments are crucial to effectively prevent and delay the process of the permanent health damage (Drenkard et al., 2013; Nickel et al., 2004). However, the diagnoses of conventional diseases have always been

☆ HLICAE: The health lesions induced by chronic arsenic exposure.

* Corresponding author. Room 7A321, Wenzhou Medical University, Chashan University Town, Wenzhou, Zhejiang, 325035, China.
E-mail address: mgy@wmu.edu.cn (G. Mao).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ijheh.2018.12.010
Received 1 June 2018; Received in revised form 18 November 2018; Accepted 20 December 2018

dependent on typical symptoms, signs and some clinical examinations, such as routine biochemical tests and X-ray examinations, having low ability for early identification about diseases. In the past several decades, some specific molecular biomarkers were reported that they could distinguish patients from general population and were widely applied in clinical practices due to their convenient data measurement (Miyata et al., 2015). Nevertheless, many of these individual biomarkers had some limitations, such as low coverage rate, high false-positive rate or high false-negative rate, which would largely reduce the accuracy of the early detection efficiently and seriously restrict further clinical application (Liu et al., 2014).

Fortunately, with the development of modern technologies, the "OMIC" techniques including metabolomics are believed to assist with early identification of many diseases based on abundant reliable and specific biomarkers (Vlaanderen et al., 2010). A proteomic study reported that the altered expression of desmoglein 1 (DSG1), keratin 6c (KRT6C) and fatty acid binding protein 5 (FABP5) were significantly associated with arsenic-induced skin keratosis, which might serve as available biomarkers of early diagnosis in high-risk population (Guo et al., 2016). Nowadays, genomics, transcriptomics and proteomics are widely applied in a lot of biological studies and provide massive information on genotypes, while transmitting inadequate information on phenotypes. As the final downstream products of gene expression, metabolites are largely linked to biological functions and phenotypes, and provide directly readable metabolic status of cells, tissues or organisms, amplifying changes of biological systems in response to pathophysiological stimuli or genetic variation (Sreekumar et al., 2009).

Unlike other "OMIC" manners, metabolomics is mainly focused on the assessment of small molecular metabolites, which reveal crucial information relevant to the individual's health status (Wang et al., 2015). Detecting these metabolic changes may offer important insights into disease mechanisms (Nicholson and Wilson, 2003; Sreekumar et al., 2009; Wang et al., 2011). Currently, multiple platforms, including nuclear magnetic resonance spectroscopy (NMR), liquid chromatography-mass spectrometry (LC-MS) and gas chromatography-mass spectrometry (GC-MS), are used in lots of metabolomics studies (Psychogios et al., 2011). Among them, LC-MS has been widely utilized in recent years, owing to its high sensitivity, specificity and simple sample processing.

Although metabolomics has been widely used in screening some specific diagnostic biomarkers (Chen et al., 2008), its application in HLICAE early recognition in humans is limited. In this study, we aim to further understand the mechanisms of HLICAE with metabolomics in a community-based Chinese chronic arsenic exposure population via drinking water. We also want to know that altered metabolites may be served as useful biomarkers for the early detection and contribute to further mechanistic studies of HLICAE.

2. Materials and methods

2.1. Study participants

The current study was originally from a randomized, double-blind, placebo-controlled clinical trial (RCT, NCT02235948). The study population included 188 subjects (94 HLICAE cases and 94 matched controls) chronically exposed to iAs via drinking water, participating in a community-based chronic arsenic exposure cohort established in 2010 at Wuyuan County of Hetao Plain, Inner Mongolia, China (Fig. 1). This county is one of the most serious chronic arsenic exposure areas via drinking water in China. Detailed data collection and measurement methods for clinical and sociodemographic variables have been published previously (Guo et al., 2015). Briefly, the purpose and procedures of the study were thoroughly explained to all participants and written informed consent was obtained from each subject before he or she was enrolled in the cohort and prior to commencing any study-related procedures. Data on clinical and social variables were obtained via physical examination and face-to-face investigation using a structured questionnaire developed for the cohort. The study protocol was approved by institutional review boards at Wenzhou Medical University and has been registered with ClinicalTrials.gov, number NCT02235948 (https://clinicaltrials.gov/ct2/show/NCT02235948).

A case-control study was performed to assess the early recognition of the HLICAE based on the data of serum metabolomics combined with clinical and social variables. A total of 94 HLICAE cases in the present study were strictly determined according to "the standard of diagnosis for endemic arsenism of China (WS/T211-2001 (2001), established by the Ministry of Health of the People's Republic of China on Nov. 14, 2001, combined with hypercholesterolemia. The detailed criteria are as follows: (1) people were diagnosed as arsenicalism; (2) participants were chronically exposed to arsenic with unexplained skin palmoplantar hyperkeratosis over grade two; (3) participants were chronically exposed to arsenic with unexplained hyperpigmentation and depigmentation over grade two; (4) participants were chronically exposed to arsenic with unexplained peripheral nerve damage; (5) participants were chronically exposed to arsenic as well as with hypercholesterolemia (the serum total cholesterol > 5.72 micro mol/L and triglyceride < 1.7 micro mol/L). In the above-mentioned criteria, "unexplained" means that the skin lesions weren't induced by some other specific diseases with clear causes rather than arsenic exposure. Otherwise, they would be eliminated and not enrolled as cases in the current study. Meanwhile, another 94 subjects without HLICAE in the same cohort were matched for gender and age (± 1 year) with the ratio one by one and defined as the controls. Among them, 50 pairs were randomly chosen as the discovery phase and other 44 pairs as the validation phase. As the results of the present study were only based on the baseline data of the above-mentioned RCT, which indicated that no subjects received any intervention in the present study, the treatment applied in the trial will have no effects on our results.

2.2. Reagents and instrumentation

LC-MS grade formic acid, water, methanol and acetonitrile were purchased from Fisher Scientific (Waltham, MA, USA). High speed refrigerated centrifuge (MIKRO220R) (Hettich, Landkreis Tuttlingen, Baden-Württemberg, Germany) and vacuum concentrator centrifuge (SPD121P) (Thermo fisher Scientific, Waltham, MA, USA) were used to centrifuge and concentrate blood samples. Metabolomics experiments were performed in ultra-high-performance liquid chromatograph (Waters Acquity system, Waters) and hybrid quadrupole-time-of-flight mass spectrometer (Xevo G2 Q-TOF MS, Waters Corp., Milford, MA, USA), equipped with an electro spray ion (ESI) source (UPLC-QTOF-MS). All instruments were controlled through a single software package (Waters MassLynx V4.1).

2.3. Sample collection

After 8–10 h of fasting, an 8 mL venous blood sample was obtained from each subject and placed in tubes without ethylenediaminetetraacetic acid (EDTA) for serum preparation at the enrollment. The serum was acquired at less than 0 °C in Wuyuan county, stored at −86 °C in a freezer and waited for the further metabolomics study. Detailed information on other covariates including urinary arsenic species profile assessment could be found in our previous study (Guo et al., 2015; Chen et al., 2017) and supplementary materials.

2.4. Sample preparation and UPLC-QTOF-MS assay

All serum samples were thawed at a temperature of 4 °C. An aliquot of 200 μL serum was transferred into a labeled 2.0 mL microcentrifuge tube with 600 μL mixture (90% acetonitrile, 10% water), thoroughly mixed in a vortex mixer for 20 s (mixed three times), all of which pelleted the protein precipitate in a centrifuge operating at 4 °C

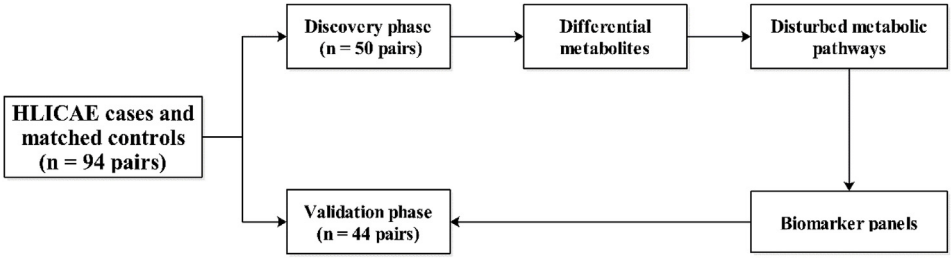

Fig. 1. Study Design of the metabolomics study. This study, involving 188 participants, included both discovery and validation phases. HLICAE: the health lesions induced by chronic arsenic exposure.

(12000 rpm for 5 min). Each sample of transferred 400 μL supernatant was dried (lyophilized) in a vacuum concentrator centrifuge. Furthermore, 130 μL water (including 15% acetonitrile) was added for further dissolution. Later, the process of mixing and protein precipitating was repeated.

To monitor the UPLC-QTOF-MS system's stability and performance along with the reproducibility of the sample treatment procedure, quality control (QC) samples were prepared by mixing an equal volume of every serum sample. The pretreatment of serum QC samples was in accord with that of serum samples. Every QC sample was injected only once, nine QC samples were continuously injected at the start of the run and one QC sample was injected at every eighth sample injection throughout the analytical run.

An aliquot of 2 μL of sample solution was injected into an ACQUITY UPLC HSS T_3 column (2.1 mm × 100 mm, 1.7 μm) held at 50 °C. And the autosampler was held at 4 °C. The total run time was 13 min, including equilibration time. Water and methanol were mobile phase A and B, respectively (both A and B contained 0.1% formic acid). The gradient elution program was described in Table S1 (Supplementary material: Table S1).

The full-scan data in both positive-ion (ESI^+) mode and negative-ion (ESI^-) mode were acquired with MassLynx V4.1 software in MS^E data collection mode, scanning from 50 to 1200 Da with 0.2 s scan time, with capillary voltage of 3.0 kV for ESI^+ mode or 2.5 kV for ESI^- mode, desolvation temperature of 350 °C, sample cone voltage of 20 V, extraction cone voltage of 4 V, source temperature of 120 °C, cone gas flow of 50 L/h and desolvation gas flow of 800 L/h. The mass spectrometer was calibrated across the mass range of 50–1200 Da, using solution of sodium formate. The collision energy was 6 eV of Function 1 (low energy) and 15–45 eV of Function 2 (high energy). Data were centroided and mass was corrected during acquisition using an external reference (Lock-Spray TM) consisting of 0.2 ng/mL solution of leucine enkephalin infused at a flow rate of 5 μL/min via a lock spray interface, generating a reference ion at 556.2771 Da ($[M+H]^+$) or 554.2615 Da ($[M\text{-}H]^-$). The lock spray scan time was set at 0.5 s with the interval of 15 s, and data were averaged over 3 scans.

2.5. Data preprocessing and statistical analysis

UPLC-QTOF-MS data preprocessing was conducted by Progenesis QI 2.0 (Waters Corp., Milford, MA, USA) software, performing background noise elimination, peak alignment, data reduction, etc. Next, the intensity of every ion was normalized to the total ion count and a data matrix was generated, consisting of retention time, mass-to-charge ratio (m/z) value and peak area for each feature. The coefficient of variation (CV) of QCs was applied to assess the stability of the data and those with CV > 30% were deleted from the final dataset in both positive and negative ion modes. Later, data were normalized, Pareto scaled and imported into the Umetrics Ezinfo 3.0 software for Waters (Umetrics, Sweden) for pattern recognition analysis (multivariate statistical analysis). To increase the accuracy and reliability of the results, we performed the following comprehensive analyses of the metabolomic characteristics on HLICAE in the discovery and validation phases, respectively. Firstly, principal components analysis (PCA) was adopted for preliminary analysis. Outliers ($T^2 > T^2$ Crit (99%)) were checked with the Hotelling T^2 test and removed from the final dataset. Later, orthogonal partial least squares discriminant analysis (OPLS-DA) was further conducted. The differential metabolites were chosen according to the variable importance in the project (VIP) value in OPLS-DA model and fold change (FC) (VIP > 1 and FC > 1). R^2 (goodness of fit) and Q^2 (predictability) values were used to assess the quality of the OPLS-DA model. A 999-times permutation test was performed to further validate and avoid overfitting of the OPLS-DA model. Moreover, univariate statistical analysis methods like paired *t*-test or Wilcoxon signed-rank test was also conducted to ensure the significant difference of each metabolite between the cases and the controls. To reduce the probability of false positive results, the method of the false discovery rate (FDR) adjustment was also performed in this study. The identification of metabolites was achieved by comparing retention time and m/z value of metabolites with the Human Metabolome Database (http://www.hmdb.ca). The molecular weight tolerance between the measured m/z values and the exact mass of metabolites was set to within 5 ppm. Additionally, further identification of the metabolites was combined with MS/MS data, using Progenesis QI 2.0 software. Heatmap and pathway analysis were performed by MetaboAnalyst 3.0 online software (http://www.metaboanalyst.ca).

Cubic smoothing spline regression models were firstly used to estimate the "real" relationship between the level of differential metabolite and the risk to develop HLICAE. Multivariable generalized linear regression models (GLMs) were used to examine the association between detected specific small molecular metabolites and the odds of HLICAE. As inappropriate covariates enrolled in the model will greatly affect the credibility of conclusion, what cofactors should be included in each model has been well considered during the data analysis. For the screening of differential metabolites, they were determined by PCA and OPLS-DA models with no cofactors included, which were consistent with many other metabolomics studies. Meanwhile, for the assessment of the association of HLICAE risk with detected metabolites, covariates such as body mass index, smoking habits and others were determined by the stepwise regression model. In addition, receiver operating characteristic (ROC) analysis was applied to assess the value of the early detection on HLICAE with the combination of the detected metabolomic-based biomarkers based on a multiple logistic regression model with adjusting for the above-mentioned covariates in the two phases. $P \leq 0.05$ was set as the significant level. R version 3.4.0 (Copyright (C) 2017 The R Foundation for Statistical Computing Platform) software was performed to finish the data management, the majority of analyses and figures drawing.

3. Results

3.1. Characteristics of the study participants

The discovery and validation phases had 50 pairs (50 HLICAE cases and 50 matched controls) and 44 pairs (44 cases and 44 controls) of participants chronically exposed to iAs via drinking water, respectively.

Table 1
Epidemiological and clinical characteristics of study subjects.[a]

Variables	Discovery phase			Validation phase		
	Control	Case	*p-value*	Control	Case	*p-value*
Number of Subjects	50	50		44	44	
Age, years	48.80 (44.70,57.60)	49.05 (44.90,57.50)	0.956	51.26 ± 9.89	51.31 ± 9.92	0.979
Body mass index, Kg/m^2	24.59 ± 3.26	23.90 ± 2.77	0.260	24.38 ± 3.14	24.76 ± 3.57	0.600
Period of Arsenic Exposure, years	42.35 (38.25,54.34)	47.00 (40.53,54.16)	0.173	46.58 ± 9.63	47.89 ± 10.99	0.554
Fasting Plasma Glucose, mmol/L	5.13 (4.57,5.47)	4.80 (4.47,5.20)	0.134	5.08 (4.59,5.46)	4.97 (4.67,5.31)	0.599
Creatinine (Cr), mg/dl	1195.9 (818.74,1571.70)	1098.9 (667.69,1555.90)	0.430	1345.9 (747.69,1749.60)	1212.3 (968.25,1491.40)	0.946
High Density Lipoprotein, mmol/L	1.09 ± 0.27	1.28 ± 0.29	0.001	1.17 (0.93,1.37)	1.16 (0.96,1.38)	0.937
Low Density Lipoprotein, mmol/L	3.05 ± 0.82	3.21 ± 0.91	0.349	3.19 ± 0.85	3.23 ± 0.78	0.821
Blood Urea Nitrogen, mmol/L	6.16 ± 1.81	6.95 ± 2.11	0.047	6.56 (5.06,8.91)	6.66 (6.17,8.24)	0.438
Trivalent Arsenic/Cr, ‰	1.26 (0.48,2.00)	1.20 (0.66,1.96)	0.771	1.22 (0.80,1.85)	0.93 (0.67,1.99)	0.973
Pentavalent Arsenic/Cr, ‰	0.55 (0.24,1.77)	0.34 (0.16,0.71)	0.048	0.50 (0.18,1.77)	0.54 (0.18,1.31)	0.806
Monomethylarsonous acid/Cr (MMA/Cr), ‰	3.27 (1.30,5.96)	3.57 (2.55,6.39)	0.527	4.06 (1.67,6.57)	3.66 (2.32,6.48)	0.734
Dimethylarsenate/Cr (DMA/Cr), ‰	7.78 (4.02,10.59)	9.08 (5.96,14.02)	0.101	6.87 (4.68,11.68)	9.44 (7.34,14.80)	0.029
MMA/Cr/(DMA/Cr), %	41.81 (30.37,64.51)	41.81 (30.36,57.26)	0.909	41.79 (30.37,54.77)	38.64 (29.27,46.10)	0.564
Han, #(%)	50 (100.00)	49 (98.00)	1.000	43 (97.73)	44 (100.00)	1.000
Illiteracy, #(%)	15 (30.61)	38 (76.00)	0.460	9 (20.45)	12 (27.27)	0.453
Married, #(%)	47 (95.92)	49 (98.00)	0.986	42 (97.67)	43 (97.73)	1.000
Farmer, #(%)	48 (97.96)	47 (94.00)	0.624	40 (90.91)	43 (97.73)	0.357
Smoking[b], #(%)	17 (34.00)	20 (40.00)	0.534	19 (43.18)	16 (36.36)	0.513
Alcohol Drinking[c], #(%)	12 (24.49)	18 (36.73)	0.188	18 (40.91)	15 (34.09)	0.509
Animal Oil[d], #(%)	41 (82.00)	42 (85.71)	0.616	37 (86.05)	34 (79.07)	0.394
High Vegetable Intake[e], #(%)	8 (16.67)	9 (18.00)	0.862	13 (30.95)	9 (20.45)	0.265
High Blood Pressure, #(%)	28 (56.00)	28 (56.00)	1.000	22 (50.00)	21 (47.73)	0.831
Menopause, #(%)	14 (46.67)	14 (45.16)	0.906	11 (50.00)	10 (43.48)	0.661

[a] Continuous data obeying normal distribution were described as Mean ± standard deviation (SD) and the paired *t*-test was applied to compare the differences between the case and control groups, otherwise, median (1st quartile, 3rd quartile) and the Wilcoxon signed-rank sum test were used. Categorical data were described as number of cases (%) and χ^2 test or Fisher's exact test was selected to compare the differences.
[b] Smoking was defined as continuous or cumulative smoking for 6 months or more in lifetime.
[c] Alcohol drinking was defined as more than one time per week or current drinker.
[d] Animal Oil was defined as the source of oil intake per week was mainly from animal oil.
[e] High Vegetable Intake was defined as vegetable intake was more than 1.5 km or above per week.

Of 188 participants, 110 were women (64 in the discovery phase and 46 in the validation phase). The demographic, clinical and biological characteristics of the study participants were presented in Table 1. For the discovery phase, the Han nationality was the largest subgroup (100% for cases and 98% for controls). The duration of arsenic exposure among the study population varied from 23.59 to 72.95 years. More than half of them had no history of smoking (63%) or alcohol drinking (68%), and 83% of the participants normally cooked with animal oil. Moreover, the majority of the participants (83.94%) had inadequate consumption of vegetables. In addition, high density lipoprotein (HDL), blood urea nitrogen (BUN) and urinary pentavalent arsenic/creatinine (Cr) were higher among the cases as compared to the controls (all $P < 0.05$).

In other ways, there were no statistically significant differences across the two groups in the anthropometric and biochemical variables, such as body mass index (BMI), education, high blood pressure, fasting plasma glucose (FPG), low density lipoprotein (LDL), serum folic acid, urinary trivalent arsenic/Creatinine, monomethylarsonous acid/Creatinine (MMA/Cr) and dimethylarsenate/Creatinine (DMA/Cr) (all $P > 0.05$).

3.2. Serum metabolome profiles

All the serum samples, collected in the present study, were separately assayed by UPLC-QTOF-MS in positive and negative ion modes, respectively. The original base peak chromatograms of all the QC samples were superimposed (Supplementary material: Fig. S1). In both ESI^+ and ESI^- modes, the peak time and shape of all quality control samples were very similar, which indicated the good stability and reproducibility of the UPLC-QTOF-MS platform.

The typical metabolome profiles of serum samples could be seen in Fig. 2. Comparing the base peak chromatograms of the case and control groups in ESI^+ and ESI^- modes, we could find that the ion peaks of the two groups were roughly similar under the same retention time. However, the height and peak area of some ions were different in the two groups, indicating the contents of some metabolites in serum samples differed in the cases and their counterparts.

3.3. Pattern recognition analysis

By applying the untargeted metabolomics method, we extracted 210 features in ESI^+ mode and 364 features in ESI^- mode (CV ≤ 30%). Based on the discovery phase, the differences of the peak area of 91 features in ESI^+ mode and 87 features in ESI^- mode between the two groups reached significant level ($P \leq 0.05$). Searched in the Human Metabolome Database (HMDB) and further analyzed the MS/MS fragmentation, 46 and 35 metabolites were obtained in ESI^+ and ESI^- modes, respectively. To further explore metabolic changes between the cases and controls, multivariate statistical analyses were carried out to investigate the complex data. Firstly, PCA was performed to provide an overview of the obtained datasets after Pareto-scaling. In ESI^- (Fig. 3A) and ESI^+ modes (Fig. 3B), tightly clustering of QC samples was observed, indicating the system's good stability and performance as well as the good reproducibility of the sample treatment procedure. The PCA score plots (Fig. 3A and B) showed preliminary separation of the two groups in both ESI^- (R_X^2(cum) = 0.850, Q^2 (cum) = 0.455) and ESI^+ modes (R_X^2(cum) = 0.874, Q^2 (cum) = 0.634), with strong predictive capabilities of the two models. Next, OPLS-DA was applied to discover some metabolites potentially associated with HLICAE. The goodness of fit and predictability of the OPLS-DA models could be observed in both the cumulative of R_Y^2 (0.521) and Q^2 (0.378) for negative ion mode and R_Y^2 (0.372) and Q^2 (0.273) for positive ion mode (Fig. 3C and D). To further validate and avoid overfitting of the OPLS-DA model in either positive or negative ion mode, 999-times permutation tests were

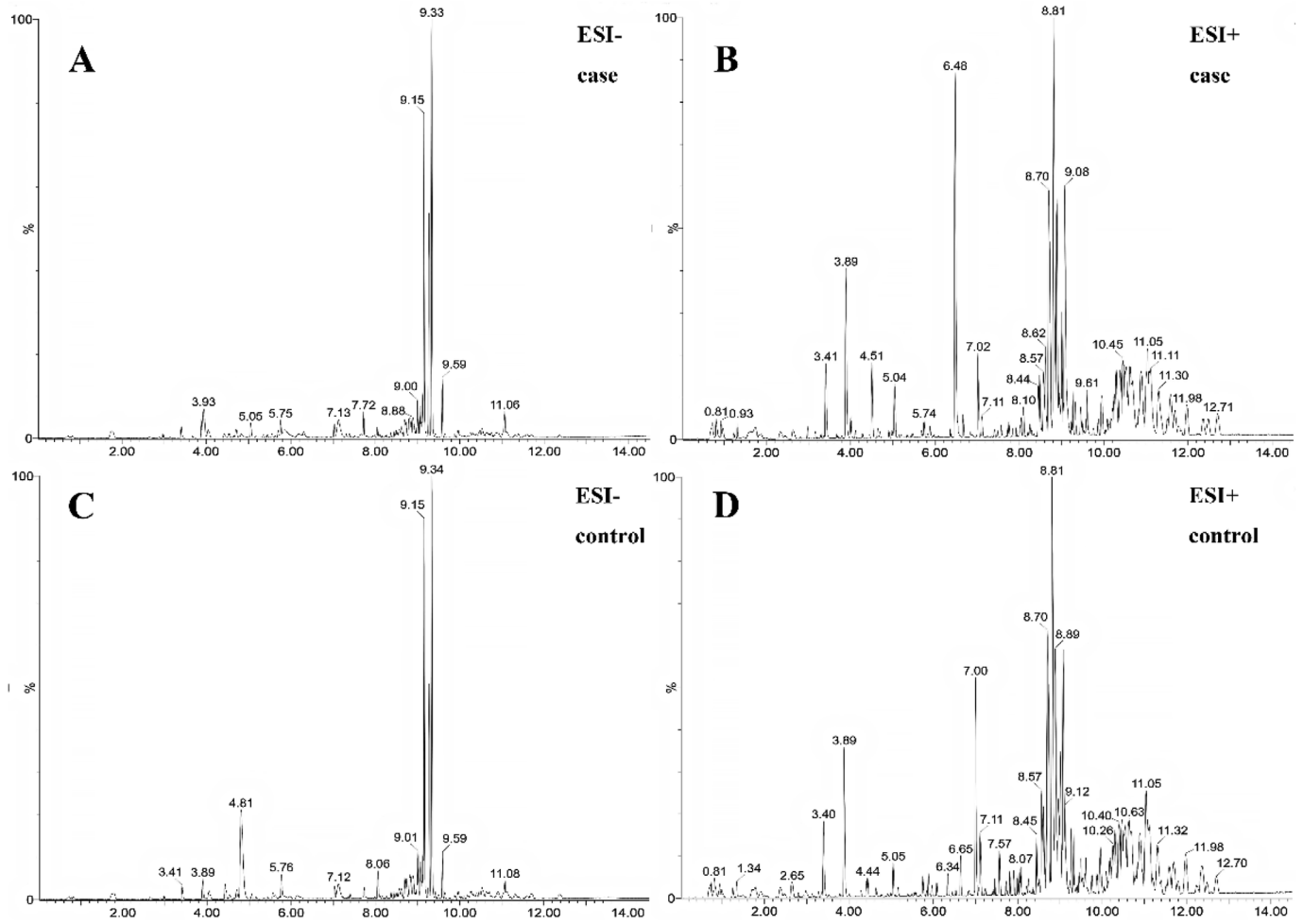

Fig. 2. Base peak chromatogram (BPC) of serum obtained from UPLC-QTOF-MS analysis in negative (A, C) and positive (B, D) ion modes. The horizon axis indicates retention time (minutes) and the vertical axis represents the relative abundance of metabolites (%).

performed separately (Supplementary material: Fig. S2). The two plots both strongly indicated that the OPLS-DA models of the two modes were valid because all Q^2 and R^2 values to the left were lower than original points to the right. Furthermore, the intersections of the regression line of Q^2 and the vertical axis were all under the zero point. In short, the cases and controls were well separated from each other in the OPLS-DA model. As compared to the counterparts, negative ion mode revealed better separation in the two groups.

3.4. Screening of differential metabolites

Differential metabolites, the important indicators of arsenic action, were selected based on both univariate and multivariate statistical analysis methods. Here we mainly presented the results of the discovery phase. In accordance with the combination of the variable importance in the project (VIP) $>$ 1, fold change (FC) $>$ 1, *p-value* $\leq$ 0.05 and the false discovery rate (FDR)-adjusted *p-value* $\leq$ 0.05, 19 metabolites in ESI^+ mode and 9 metabolites in ESI^- mode were screened in this study (Table 2). Among them, 16 metabolites in the cases were significantly higher than those of the controls. Meanwhile, the other 12 metabolites in the cases were much lower than those of the counterparts. To test the reproducibility of the metabolites screened in the discovery phase, we repeated the above-mentioned analysis steps in the validation set and obtained similar results (Supplementary material: Table S3).

3.5. Heatmap and pathway analysis

The distribution of the relative peak area of the screened metabolites in the two groups were presented with a heatmap (Discovery phase, Fig. 4). Compared to the counterparts, Lysophosphatidylcholines, L-Tryptophan, L-Leucine and L-Phenylalanine were down-regulated in the cases. However, oleic acid, linoleic acid, arachidonic acid and phosphatidylcholines were up-regulated. Furthermore, pathway analysis was also conducted to further explore the perturbed pathways based on discovery phase, which mainly covered the changes of the metabolisms in linoleic acid, glycerolphospholipid, phenylalanine, tryptophan and so on (Fig. 5).

3.6. Association between serum metabolite levels and the risk of HLICAE occurrence

In the current study, cubic smoothing spline regression models were applied to fit the curves between the levels of these 28 small molecule metabolites and the risk to get HLICAE, respectively. As could be seen in Fig. S3 (Supplementary material: Fig. S3), there existed obvious dose-response relationships between the screened serum metabolites and the risk of HLICAE development. We also observed that the probability of HLICAE occurrence significantly associated with above-mentioned 28 serum metabolites based on multivariable conditional logistic regression models after adjusting for some potential confounding factors (Supplementary material: Table S4).

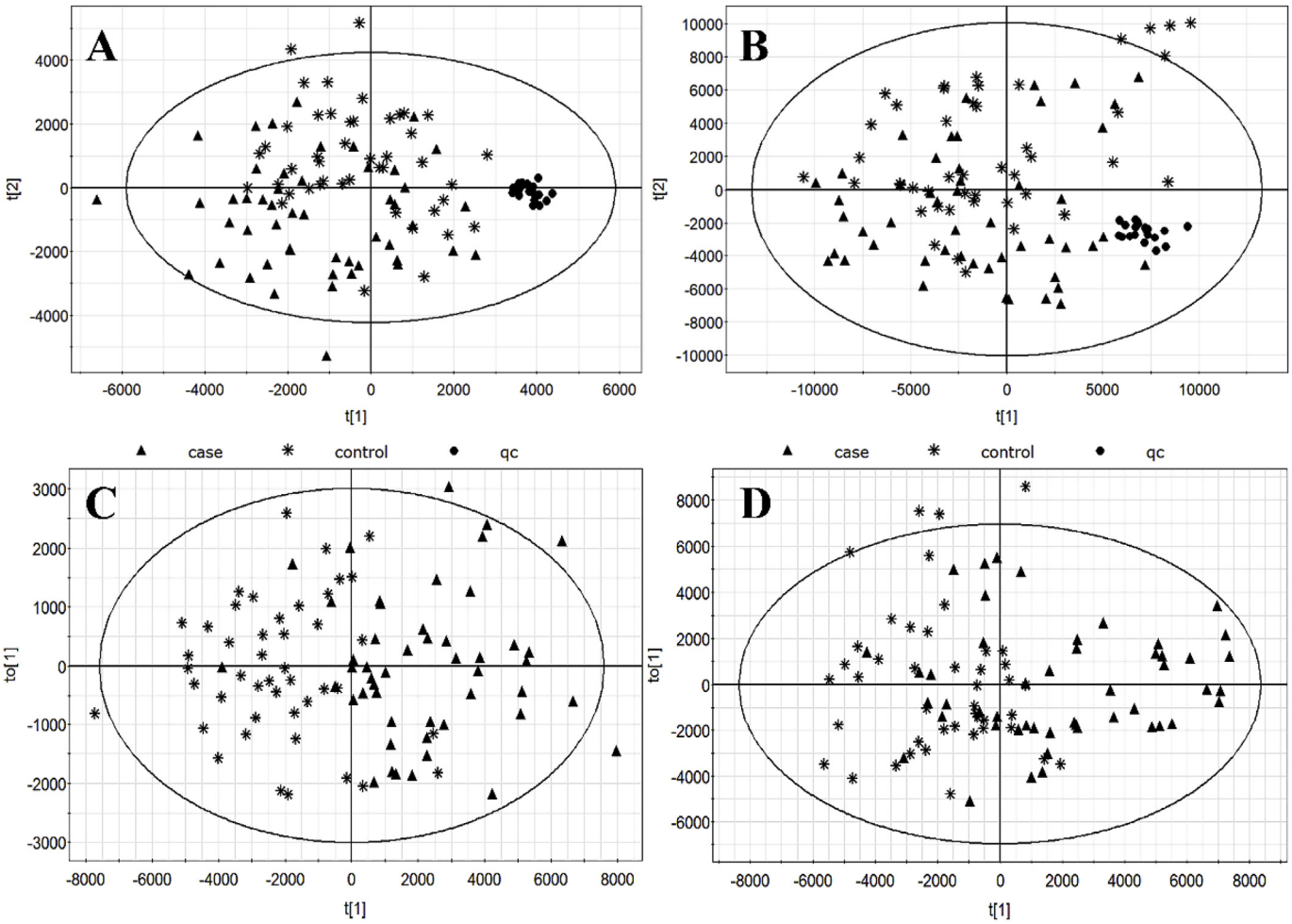

Fig. 3. Multivariate statistical analyses of serum samples in the discovery phase analyzed by fingerprinting method in positive and negative ion modes. A and B indicate the PCA models built for serum samples which were analyzed in ESI^- mode and ESI^+ mode with prediction of quality control (QC) samples. C and D mean the classification of serum samples of the OPLS-DA models built for filtered data generated in ESI^- mode and ESI^+ mode. ▲ - case, ✻ - control, ● - QC.

3.7. The early detection value of the metabolomic-based biomarkers

The areas under the curve (AUC) and its related 95% confidence interval (CI) of each metabolite in the discovery phase were presented in Table 2 and AUC (95% CI) of metabolites in the validation phase can be found in Table S3 (Supplementary material). Furthermore, to improve the sensitivity and specificity, we used the multivariable conditional logistic regression model to evaluate the ability of the combination of all metabolomic-based biomarkers for distinguishing HLICAE from the normal participants in the discovery phase. The AUC (95% CI), sensitivity and specificity were 0.908 (0.855, 0.960), 96.0% and 66.0% in ESI^- mode (Fig. 6A), and 0.898 (0.836, 0.960), 96.0% and 72.0% in ESI^+ mode (Fig. 6B). Meanwhile, for the validation set, the AUC (95% CI), sensitivity and specificity were 0.934 (0.887, 0.982), 86.4% and 86.4% in ESI^- mode (Supplementary material: Fig. S4A), and 0.821 (0.732, 0.909), 75.0% and 79.6% in ESI^+ mode (Supplementary material: Fig. S4B). The positive predictive value of distinguishing HLICAE from the controls were 78.6% in ESI^+ mode and 86.4% in ESI^- mode in the validation phase. These results suggested that the above-mentioned metabolites could be utilized as potential sensitive and specific biomarkers for HLICAE early recognition in a chronic arsenic exposure population.

4. Discussion

Accumulating evidence showed that chronic arsenic exposure was significantly associated with metabolic syndrome (Chen et al., 2012a; Dodson and Zhang, 2016). However, the exact effects of chronic arsenic exposure on the metabolome in humans were not well understood. In this study, our findings clearly showed that HLICAE risk was significantly associated with some specific metabolites including lipids and amino acids, which strongly indicated that metabolomics might be a good option in the early identification of HLICAE in a population chronically exposed to arsenic. Our findings were highly in line with the results of some recent published studies (Wu et al., 2018; Zhou et al., 2017; Wang et al., 2015).

In the present study, both cases and controls came from a single chronic arsenic exposure area and had comparable life-style. Furthermore, participants in either the discovery or validation phase were randomly selected by pairs from the 188 subjects. This perhaps was the reason why few significant differences of arsenic level were observed in either the discovery or validation phase (Table 1). Moreover, there exists a long-term duration, covering normal, pre-disease and disease states, between the exposures or causes and clinical manifestations in most complex diseases. From the pre-disease to disease state, many internal bio-signals including small molecular metabolites will altered dramatically and remain unchanged in the future (Chen et al., 2012b; Liu et al., 2014). Obvious alteration of some bio-signals will be observed in both the patients and undiagnosed participants in its early-disease state. So, comprehensively investigation of these bio-signals can be helpful on disease early detection. It is also believed that most non-communicable chronic diseases (NCDs) including HLICAE are

Table 2
Differential metabolites in chronic arsenic exposure population of discovery phase.[a]

No	HMDB ID	Compound name	Mass-to-charge ratio (m/z)	Retention time (min)	*p-value*	FDR-adjusted *p-value*	VIP	Fold change	Trend	AUC (95%CI)
Negative Ion Mode										
1	00207	Oleic acid	281.249	9.333	< 0.001	< 0.001	13.112	1.634	↑	0.834 (0.753, 0.914)
2	00673	Linoleic acid	279.233	9.154	< 0.001	< 0.001	9.590	1.760	↑	0.853 (0.781, 0.925)
3	01388	Alpha-Linolenic acid	277.217	9.003	< 0.001	< 0.001	3.702	1.489	↑	0.758 (0.665, 0.851)
4	10378	5,8,11-Eicosatrienoic acid	305.248	9.261	< 0.001	< 0.001	2.972	1.625	↑	0.846 (0.772, 0.919)
5	01043	Arachidonic acid	303.233	9.128	< 0.001	< 0.001	2.914	1.262	↑	0.747 (0.649, 0.845)
6	00929	L-Tryptophan	203.082	3.886	< 0.001	< 0.001	2.174	1.153	↓	0.742 (0.839, 0.645)
7	06528	Docosapentaenoic acid	329.248	9.257	< 0.001	< 0.001	2.015	1.623	↑	0.801 (0.716, 0.886)
8	05060	Eicosadienoic acid	307.264	9.422	< 0.001	< 0.001	1.784	1.530	↑	0.803 (0.717, 0.888)
9	10386	LysoPC (18:2)	564.330	8.702	0.004	0.021	1.540	1.108	↓	0.678 (0.784, 0.573)
Positive Ion Mode										
1	02815	LysoPC (18:1)	522.356	8.888	0.003	0.016	4.133	1.143	↓	0.672 (0.777, 0.566)
2	10384	LysoPC (18:0)	524.372	9.078	0.011	0.047	3.336	1.113	↓	0.642 (0.751, 0.534)
3	08431	PC (20:4/18:0)	832.583	11.051	< 0.001	< 0.001	3.019	1.238	↑	0.729 (0.632, 0.827)
4	08169	PC (18:3/18:1)	782.570	10.301	0.005	0.025	2.582	1.084	↑	0.691 (0.586, 0.796)
5	08226	PC (18:3/P-18:0)	768.590	10.842	< 0.001	< 0.001	2.571	1.151	↑	0.731 (0.63, 0.831)
6	00734	Indoleacrylic acid	188.071	3.883	0.002	0.012	2.536	1.163	↓	0.671 (0.777, 0.566)
7	13420	PC (o-18:0/20:4)	796.621	11.516	< 0.001	< 0.001	2.271	1.189	↑	0.728 (0.628, 0.828)
8	10382	LysoPC (16:0)	496.341	8.730	0.007	0.034	2.222	1.117	↓	0.655 (0.763, 0.548)
9	09225	PE (20:0/18:2)	772.584	10.799	< 0.001	< 0.001	1.916	1.231	↑	0.764 (0.669, 0.858)
10	00159	L-Phenylalanine	166.087	3.402	< 0.001	< 0.001	1.901	1.210	↓	0.752 (0.848, 0.657)
11	13302	Phenylalanylphenylalanine	313.155	5.048	0.003	0.016	1.681	1.270	↓	0.679 (0.785, 0.574)
12	08659	PC (22:5/16:0)	830.567	10.581	0.002	0.012	1.618	1.190	↑	0.687 (0.582, 0.792)
13	13122	LysoPC (P-18:0)	508.377	9.050	0.008	0.037	1.522	1.135	↓	0.667 (0.777, 0.558)
14	00687	L-Leucine	132.102	2.642	< 0.001	< 0.001	1.320	1.275	↓	0.764 (0.86, 0.669)
15	05065	Oleoylcarnitine	426.358	8.081	0.001	0.007	1.310	1.245	↓	0.717 (0.82, 0.614)
16	10395	LysoPC (20:4)	544.340	8.684	0.010	0.044	1.286	1.118	↓	0.632 (0.743, 0.522)
17	07871	PC (14:0/18:0)	734.570	10.788	0.003	0.016	1.245	1.146	↑	0.693 (0.59, 0.796)
18	12107	SM (d18:1/24:1)	835.666	12.581	0.004	0.021	1.093	1.180	↑	0.694 (0.59, 0.799)
19	12102	SM (d18:1/20:0)	781.619	11.409	< 0.001	< 0.001	1.000	1.265	↑	0.735 (0.636, 0.834)

[a] VIP, variable importance in the project; FDR, the false discovery rate; HMDB, the Human Metabolome Database; AUC, areas under the curve; CI, confidence interval.

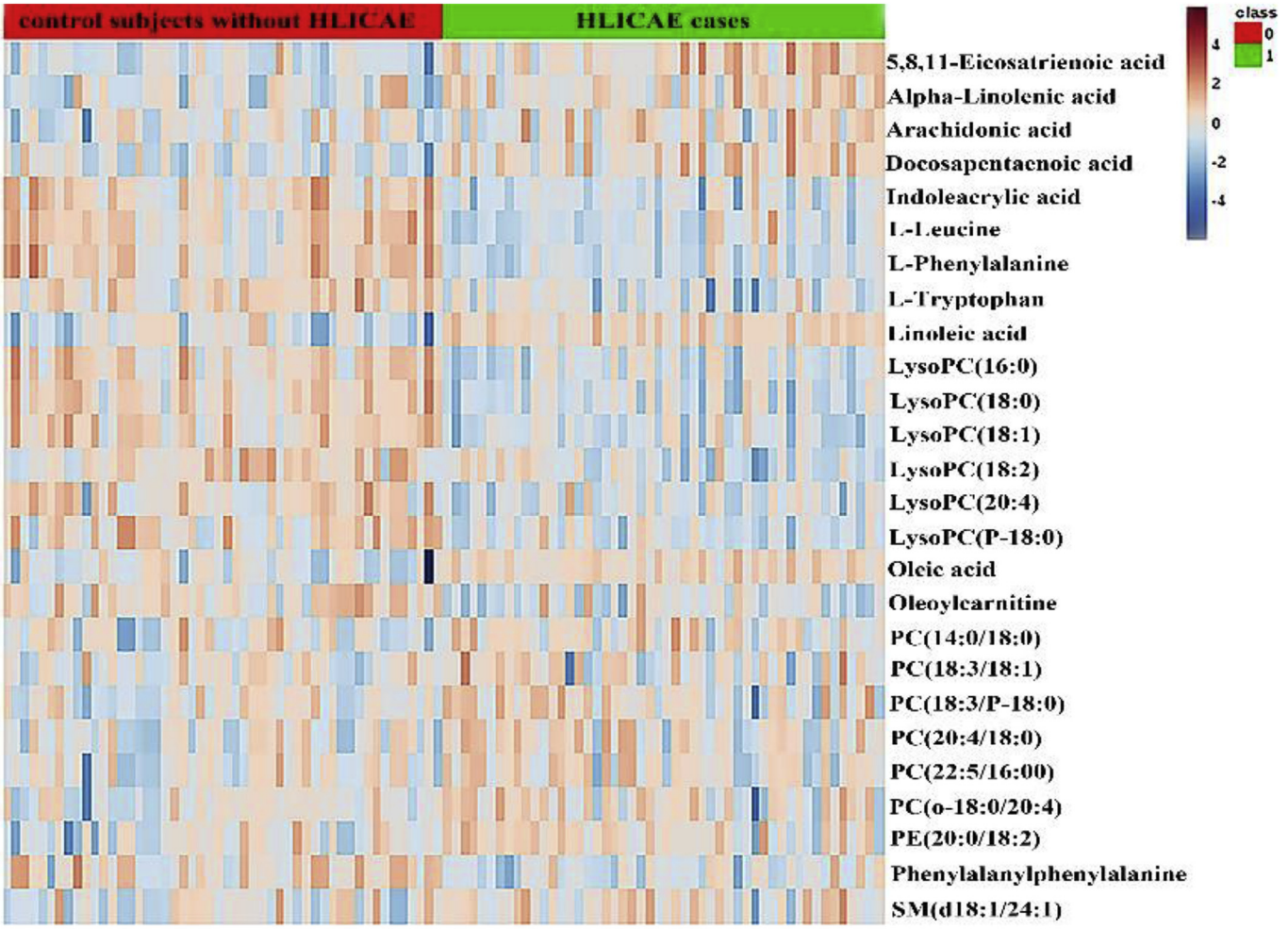

Fig. 4. Heatmap visualization analysis of differential metabolites (Discovery phase). This figure presents relative peak areas of the final differential metabolites between the cases of health lesions induced by chronic arsenic exposure (HLICAE) and subjects without HLICAE in the discovery phase. With colors changing from orange to blue, orange means up-regulated metabolite in the serum sample, and blue means down-regulated metabolite in the serum sample. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

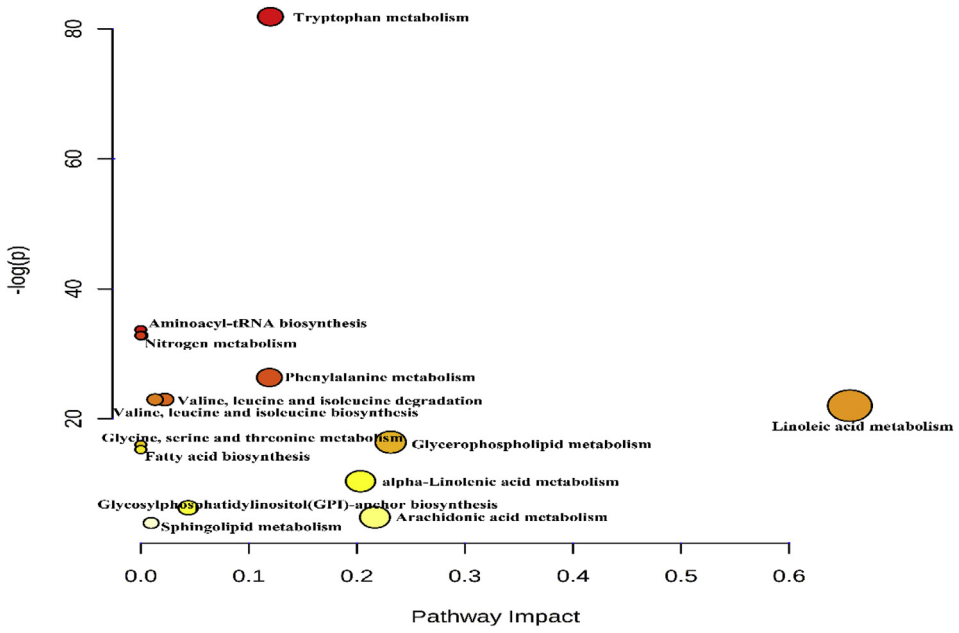

Fig. 5. Disturbed Metabolic Pathways in Chronic Arsenic Exposure Population (Discovery Phase). The disturbed metabolic pathways show various changed metabolism between cases and controls. Circles represent major metabolisms related with health lesions induced by chronic arsenic exposure (HLCAE), according to *p-value* and pathway impact score from pathway analysis. Color gradient signifies the significance of the pathway ranked by *p-value* (yellow means higher *p-value* and red means lower *p-value*). Circle size signifies the pathway impact score (larger means higher impact score). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

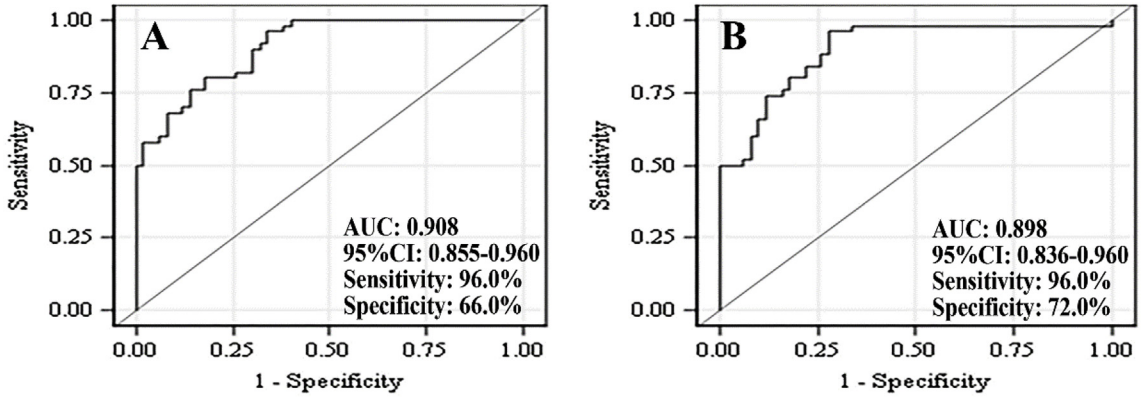

Fig. 6. The early identification value of the metabolomic-based biomarkers via the receiver-operating characteristic (ROC) curves for distinguishing the health lesions induced by chronic arsenic exposure (HLICAE) from normal participants in the discovery phase in negative (A) and positive (B) ion modes.

treated unsatisfactorily, and the associated health damages can't be fully recovered in their late stages but are preventable. Therefore, early detection has been accepted to be crucial to prevent the permanent damages due to NCDs as early as possible. The results of ROC analysis are believed to be a sound evidence in the early detection of disease of interest though the cases are diagnosed patients.

Recent rapid advance of high-throughput technologies including metabolomics provides unprecedented rich information to the early detection of diseases. Nowadays, metabolomics is accepted as an attractive biomarker discovery tool, and several related studies were performed in the past several decades. Nagato and co-authors reported that 48 h exposure to sub-lethal concentration of arsenic (49 μg/L) resulted in a metabolic shift in comparison to the controls (Nagato et al., 2013). Some metabolomics studies suggested that arsenic exposure could affect the seleno-proteins level (Garcia-Sevillano et al., 2014c), altered the gut microbiome community and substantially disturbed its metabolic profiles (Lu et al., 2014). Zhang et al. observed obvious dose-dependent urinary metabolome changes with arsenic exposure in a Chinese adult male cohort (Zhang et al., 2014). A recent study reported that 18 differential serum metabolites were identified between male Sprague Dawley rats exposed to arsenic and the controls (Wang et al., 2015). Furthermore, the altered metabolites might induce the changed metabolism in rats exposed to arsenic, such as cyclic phosphatidic acid (CPA), lysophosphatidylcholine (LysoPC), L-palmitoylcarnitine and deoxygcholylglycine (Wang et al., 2014). Consistent with our findings, all of these studies revealed that the altered metabolites might serve as available biomarkers on HLICAE early recognition in high-risk population. As most of the current evidence was obtained from animal experiments, this report might be the first serum metabolomics research on the early identification of HLICAE in a population-based study.

4.1. Lipids metabolism

Lysophosphatidylcholine (LPC, LysoPC) is a class of phospholipids, existing in the cell membrane (≤3%) and blood plasma (8–12%) (Munder et al., 1979). LPC is mainly formed through hydrolyzing phospholipids by secretory phospholipase A2-II (sPLA2-II) and hydrolyzing phosphatidylcholine (PC) by phospholipases A1 and A2 (PLA1 and PLA2) (Klingler et al., 2016; Kougias et al., 2006). Metabolomics studies of human plasma revealed the association between plasma LPC and type 2 diabetes, obesity and insulin resistance (Barber et al., 2012; Drogan et al., 2015; Wallace et al., 2014; Wang-Sattler et al., 2012). Moreover, the level of LPC was negatively correlated with the inflammatory markers, such as C-reactive protein (CRP), tumor necrosis factor-α (TNF-α), resistin and monocyte chemotactic protein 1 (MCP-1), suggesting a potential anti-inflammatory effect of LPC (Wallace et al., 2014).

In the present study, we found that six kinds of LPC, including LysoPC (P-18:0), LysoPC (16:0), LysoPC (18:0), LysoPC (18:1) LysoPC (18:2) and LysoPC (20:4), decreased in the HLICAE population compared to their counterparts. The decreased levels of the above long chain LysoPCs indicated that the transformation process of LPC might be disrupted by arsenic exposure. Due to the potential anti-inflammatory effect of LPC, the decreased level of LPC might trigger or amplify arsenic-induced inflammatory responses, like conjunctivitis, gingivitis, stomatitis and colitis (Wallace et al., 2014). However, other reports found that LysoPC (m/z = 450–600) increased in the plasma of mice (*Mus musculus*) exposed to iAs (Garcia-Sevillano et al., 2014a; Garcia-Sevillano et al., 2014b).

It is well believed that phosphatidylcholine (PC) is the major phospholipid in cellular membrane of mammals and play important roles in crucial cellular functions in the body, such as signal transduction, apoptosis, necrosis, and protein sorting (Kawai et al., 1974). In this study, as compared to the controls, the levels of PCs, including PC (22:5/16:0), PC (o-18:0/20:4), PC (20:4/18:0), PC (18:3/18:1), PC (18:3/P-18:0) and PC (14:0/18:0), largely increased in the HLICAE population. The ascending levels of PCs might affect cellular functions by arsenic exposure. Unlike our findings, other research revealed that PC (m/z = 700–850) came down in the plasma of arsenic-exposed mice (*Mus musculus*) (Garcia-Sevillano et al., 2014a; Garcia-Sevillano et al., 2014b).

You et al. reported that the endogenous PC could generate arachidonic acid (AA) and LPC, deacylated by cytosolic phospholipase A2 (cPLA2) (You et al., 2015). AA is a kind of ω-6 polyunsaturated fatty acid (20:4), participating in the construction of cell membrane and maintaining the stability and function of cell membrane (Fukaya et al., 2007). Oleic acid, a kind of monounsaturated ω-9 fatty acid (18:1), is responsible for the regulation of cholesterol level and interferes directly with the inflammatory response that characterizes early atherogenesis (Bowen et al., 2017; Massaro et al., 1999). In contrast with the controls, the elevated levels of oleic acid and AA in the case group were consistent with another study (Garcia-Sevillano et al., 2014a). Since AA could inhibit platelet aggregation (Wong et al., 2016), the increased level of AA might raise the risk of bleeding upon arsenic exposure. And the ascendant level of oleic acid might affect cholesterol level and atherogenesis-induced inflammatory response in the body owing to chronic arsenic exposure. Unlike our findings, a rodent model and human serum research showed AA and oleic acid were negatively correlated in total lipids of rat/human serum (Hostmark and Haug, 2013a; b). This inconsistency between our study and others might be in virtue of the responses to arsenic of different species between humans and rodents. Compared with rodents, humans might have better abilities to respond effectively to confront the arsenic attack. However, this hypothesis still needs to be further validated.

5,8,11-Eicosatrienoic acid (20:3n-9), also named Mead acid, was synthesized from oleic acid in essential fatty acid deficiency. It can be

converted into 3-series leukotrienes, which was a group of important inflammatory mediators and has biphasic effect on platelet aggregation (Hamazaki et al., 2009; Rachelefsky, 1997; Sakuradani et al., 2002). In this study, Mead acid increased in the cases in comparison with the controls. The increased level of 20:3n-9 might cause or increase arsenic induced inflammatory response and bleeding. Perhaps this is one of the novel potential biomarkers, which can be used in the early recognition of the health lesions induced by chronic arsenic exposure.

4.2. Amino acid metabolism

Alteration of amino acid metabolism, a common response to many toxins in animal research, indicated a general response to toxicant exposure rather than a specific biological response to a particular toxicant (Connor et al., 2004). In our study, we observed that L-Tryptophan, L-Leucine, L-Phenylalanine and their derivatives, such as indoleacrylic acid and phenylalanylphenylalanine, decreased in the cases in comparison with the controls. The altered level of L-Tryptophan was consistent with another study (Garcia-Sevillano et al., 2014a). Mellor et al. reported a close association between tryptophan and immune system activation and inflammation (Mellor et al., 2001), which suggested that the levels of inflammation and immune activity might be down-regulated after long-term arsenic exposure. Furthermore, phenylalanine, the precursor of tyrosine, was mainly observed in the human brain and plasma and was believed to be associated with cerebral serotonin and catecholamine metabolism (Mckean, 1972). In addition, phenylalanylphenylalanine (Phe-Phe), a peptide composed of two phenylalanine molecules, was found in multiple myeloma (MM) and chronic lymphocytic leukaemia (CLL) (Piszcz et al., 2013). Owning to the antioxidant effects of these amino acids, the decreased levels of L-Tryptophan and L-Phenylalanine indicated that arsenic exposure population might have lower antioxidant capacity and further trigger or amplify oxidative stress. Meanwhile, L-Leucine, a kind of branched chain amino acids (BCAA), is thought to be critical to human life, especially in stress, energy and muscle metabolism. Zanchi et al. suggested that L-Leucine could modify plasma insulin concentration and directly affect the brain's resting state functional networks such as the food-reward system and appetite regulation (Zanchi et al., 2016).

4.3. Strengths and limitations

Compared to previous studies, the cases enrolled in the present study were screened from a community-based chronic arsenic exposure cohort, which would largely reduce the impact induced by selection bias. The controls were strictly matched with gender and very similar ages (± 1 year) from the same arsenic exposure cohort and would avoid, to some extent, the influences of arsenic exposure, gender, age and other potential confounding factors in distinguishing HLICAE cases from their counterparts. In general, the relationships between the outcomes and indicators are usually estimated with traditional linear regression model, which usually assumes that the outcome linearly related to those indicators. Unfortunately, it is not always the truth. It will greatly, at least partly, affect the “real” relationship and decrease the credible of the estimation. To overcome this problem, spline regression model is developed to improve the robustness of inferences and believed to be much better than traditional linear regression model. (Harring, 2014) In the present study, cubic smoothing spline regression models were applied to examine the real association of 28 detected metabolites with the odds of HLICAE. It is suggested that traditional single biomarker can't identify early disease samples due to its static nature. It may also suffer from low coverage and high false-positive rate or high false-negative rate, which seriously limit its further applications. To overcome these difficulties, network biomarkers from different pathways or phases of the disease attract much attention and achieve better performance because they are considered to be able to distinguish pre-disease state from normal and disease states by even a small number of samples, and therefore has great potential to achieve “real” early diagnosis of complex diseases(Chen et al., 2012b; Liu et al., 2014). In the present study, the 28 detected metabolites were combined as a network biomarker with multiple regression model and could be used in the early detection or screening of HLICAE in population at high risk. Additionally, the metabolites associated with HLICAE development were comprehensively analyzed based on the discovery phase and verified in the validation phase, which would largely increase the credibility of our findings. Finally, to estimate the impacts due to potential insufficient power on our conclusion, the power calculation for paired design study, with type I error as 0.05 and number of pairs as 50, was conducted with 28 detected metabolites separately. The result showed that the minimum sample size was 36 pairs, which met the statistical needs for all 28 differential metabolites in the discovery phase. Based on 50 pairs of participants in the discovery phase, the lowest power among all 28 detected metabolites is 0.921 (Supplementary material: Fig. S5). The sample size of this study, 50 pairs for discovery phase and 44 pairs for validation phase, would well balance the power of tests about the two-stage study. These perhaps are the major strengths of this study.

However, this study may also have some limitations. First, to the best of knowledge, no golden standard for HLICAE diagnosis can be obtained. Some studies revealed that chronic arsenic exposure was significantly related with total cholesterol and could lead to obvious dyslipidemia (Mendez et al., 2016; Waghe et al., 2017). So, we defined HLICAE mainly by “the standard of diagnosis for endemic arsenism of China (WS/T211-2001, 2001, combined with hypercholesterolemia. Second, the controls were selected from the same arsenic-exposed population as the cases, rather than randomly enrolled from a non-arsenic exposure population. This might lead to misclassification bias, to some extent, as the controls might include a few HLICAE in their very early stage. This might not be considered as the best option. However, this research was a case-control study and the outcome were HLICAE. If the controls were non-exposed to arsenic, the probability to develop HLICAE would be zero and the odds ratio (95% CI) of HLICAE occurrence could not be assessed. So, non-arsenic exposure population couldn't be the ideal controls in this study. Furthermore, the present study mainly explored the relationships between the small molecular metabolites and the risk to develop HLICAE. From the perspective of scientific research, the exposure or main research factors should be small molecule metabolites. Arsenic exposure is an important confounding factor rather than a research factor. And the cases and controls were enrolled from the same cohort, having the similar history of arsenic exposure, same eating habits, lifestyle and others. This could largely balance the influences of these important confounding factors on results. Fortunately, we found no significant differences between cases and controls in the above-mentioned potential confounding factors (Table 1). Third, the cases of this study only covered diagnosed HLICAE patients, which might be another limitation. As the majority of HLICAE were chronic diseases, the duration between the onsets and diagnosis were quite long. No typical symptoms and signs in their early stages could be observed. So, it was very difficult to find several appropriate time-points in the process of HLICAE development. This is also an important reason why some metabolomics studies including this research only involve one time-point (Aggio et al., 2016; Drogan et al., 2015; Zhang et al., 2016). Finally, all participants in the present study were chronically exposed to arsenic via drinking water. This might affect, to some extent, our conclusion being extrapolated to other ways of arsenic exposure, such as coal burning, etc.

5. Conclusion

In the present study, based on the non-targeted metabolomics approach with UPLC-QTOF-MS technology platform, 28 serum small molecular metabolites related to arsenic-induced health lesions were picked out in a chronic arsenic-exposed population. Metabolic changes

mainly manifested the disruption of lipid and amino acid metabolisms. Our findings provide new insights on mechanism of health effects probably induced by chronic arsenic exposure, and these potential biomarkers, which mainly involved in lipids and amino acids, could be applied in the early detection for HLICAE. Additional studies are still indispensable for further validation and assessment of the biomarkers identified in our study.

Declarations of interest

None.

Funding

This study was supported in part by the Science and Technology Department of Zhejiang Province grant (2013C33169), the initial Scientific Research Fund (KYQD170301) and the major project of the Eye Hospital of Wenzhou Medical University (YNZD201602).

Acknowledgments

In the present study, we sincerely thank the assistance and cooperation of the Center for Disease Control and Prevention in Wuyuan, Inner Mongolia and thank again all the study participants. At point, we also express our highest respect and sincere appreciation to the field team at the School of Public Health & Management, Wenzhou Medical University, due to their hard work and dedication. Additionally, our heartfelt thanks must also go to Beijing Masspeaks Technology Co., Ltd, for their careful assay of all the serum samples and their guidance about the metabolomics data analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2018.12.010.

References

Aggio, R.B.M., Costello, B.D., White, P., Khalid, T., Ratcliffe, N.M., Persad, R., Probert, C.S.J., 2016. The use of a gas chromatography-sensor system combined with advanced statistical methods, towards the diagnosis of urological malignancies. J. Breath Res. 10 (1), 017106.

Barber, M.N., Risis, S., Yang, C., Meikle, P.J., Staples, M., Febbraio, M.A., Bruce, C.R., 2012. Plasma lysophosphatidylcholine levels are reduced in obesity and type 2 diabetes. PLoS One 7 (7), e41456.

Bowen, K.J., Kris-Etherton, P.M., Shearer, G.C., West, S.G., Reddivari, L., Jones, P.J.H., 2017. Oleic acid-derived oleoylethanolamide: a nutritional science perspective. Prog. Lipid Res. 67, 1–15.

Chen, J., Zhao, X., Fritsche, J., Yin, P., Schmitt-Kopplin, P., Wang, W., Lu, X., Haring, H.U., Schleicher, E.D., Lehmann, R., Xu, G., 2008. Practical approach for the identification and isomer elucidation of biomarkers detected in a metabonomic study for the discovery of individuals at risk for diabetes by integrating the chromatographic and mass spectrometric information. Anal. Chem. 80 (4), 1280–1289.

Chen, J.W., Wang, S.L., Wang, Y.H., Sun, C.W., Huang, Y.L., Chen, C.J., Li, W.F., 2012a. Arsenic methylation, GSTO1 polymorphisms, and metabolic syndrome in an arseniasis endemic area of southwestern Taiwan. Chemosphere 88 (4), 432–438.

Chen, L., Liu, R., Liu, Z.P., Li, M., Aihara, K., 2012b. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. Sci. Rep. 2, 342.

Chen, X.S., Guo, X.J., He, P., Nie, J., Yan, X.Y., Zhu, J.Q., Zhang, L.P., Mao, G.Y., Wu, H.M., Liu, Z.Y., Aga, D.N., Xu, P.L., Smith, M., Ren, X.F., 2017. Interactive influence of N6AMT1 and As3MT genetic variations on arsenic metabolism in the population of Inner Mongolia, China. Toxicol. Sci. 155 (1), 124–134.

Connor, S.C., Wu, W., Sweatman, B.C., Manini, J., Haselden, J.N., Crowther, D.J., Waterfield, C.J., 2004. Effects of feeding and body weight loss on the 1H-NMR-based urine metabolic profiles of male Wistar Han rats: implications for biomarker discovery. Biomarkers 9 (2), 156–179.

Dodson, M., Zhang, D., 2016. Chronic arsenic-induced metabolic syndrome: a role for prolonged Nrf2 activation and mitochondrial metabolism. Free Radical Biol. Med. 100, S183.

Drenkard, C., Dunlop-Thomas, C.M., Bao, G.B., Lim, S.S., 2013. Depression screening in patients with systemic lupus erythematosus from the southeastern United States: missing opportunities for early diagnosis and treatment. Arthritis Rheum. 65, S455–S456.

Drogan, D., Dunn, W.B., Lin, W., Buijsse, B., Schulze, M.B., Langenberg, C., Brown, M., Floegel, A., Dietrich, S., Rolandsson, O., Wedge, D.C., Goodacre, R., Forouhi, N.G., Sharp, S.J., Spranger, J., Wareham, N.J., Boeing, H., 2015. Untargeted metabolic profiling identifies altered serum metabolites of type 2 diabetes mellitus in a prospective, nested case control study. Clin. Chem. 61 (3), 487–497.

Fukaya, T., Gondaira, T., Kashiyae, Y., Kotani, S., Ishikura, Y., Fujikawa, S., Kiso, Y., Sakakibara, M., 2007. Arachidonic acid preserves hippocampal neuron membrane fluidity in senescent rats. Neurobiol. Aging 28 (8), 1179–1186.

Garcia-Sevillano, M.A., Contreras-Acuna, M., Garcia-Barrera, T., Navarro, F., Gomez-Ariza, J.L., 2014a. Metabolomic study in plasma, liver and kidney of mice exposed to inorganic arsenic based on mass spectrometry. Anal. Bioanal. Chem. 406 (5), 1455–1469.

Garcia-Sevillano, M.A., Garcia-Barrera, T., Gomez-Ariza, J.L., 2014b. Application of metallomic and metabolomic approaches in exposure experiments on laboratory mice for environmental metal toxicity assessment. Metallomics 6 (2), 237–248.

Garcia-Sevillano, M.A., Garcia-Barrera, T., Navarro-Roldan, F., Montero-Lobato, Z., Gomez-Ariza, J.L., 2014c. A combination of metallomics and metabolomics studies to evaluate the effects of metal interactions in mammals. Application to Mus musculus mice under arsenic/cadmium exposure. J. Proteomics 104, 66–79.

Guo, X., Cui, H., Zhang, H., Guan, X., Zhang, Z., Jia, C., Wu, J., Yang, H., Qiu, W., Zhang, C., Yang, Z., Chen, Z., Mao, G., 2015. Protective effect of folic acid on oxidative DNA damage: a randomized, double-blind, and placebo controlled clinical trial. Medicine (Baltim.) 94 e1872.

Guo, Z., Hu, Q., Tian, J., Yan, L., Jing, C., Xie, H.Q., Bao, W., Rice, R.H., Zhao, B., Jiang, G., 2016. Proteomic profiling reveals candidate markers for arsenic-induced skin keratosis. Environ. Pollut. 218, 34–38.

Hamazaki, T., Suzuki, N., Widyowati, R., Miyahara, T., Kadota, S., Ochiai, H., Hamazaki, K., 2009. The depressive effects of 5,8,11-eicosatrienoic acid (20:3n-9) on osteoblasts. Lipids 44 (2), 97–102.

Harring, J.R., 2014. A spline regression model for latent variables. Educ. Psychol. Meas. 74 (2), 197–213.

Hostmark, A.T., Haug, A., 2013a. Percentage oleic acid is inversely related to percentage arachidonic acid in total lipids of rat serum. Lipids Health Dis. 12, 40.

Hostmark, A.T., Haug, A., 2013b. Percentages of oleic acid and arachidonic acid are inversely related in phospholipids of human sera. Lipids Health Dis. 12, 106.

Hughes, M.F., Beck, B.D., Chen, Y., Lewis, A.S., Thomas, D.J., 2011. Arsenic exposure and toxicology: a historical perspective. Toxicol. Sci. 123 (2), 305–332.

Kawai, K., Fujita, M., Nakao, M., 1974. Lipid components of two different regions of an intestinal epithelial cell membrane of mouse. Biochim. Biophys. Acta 369 (2), 222–233.

Klingler, C., Zhao, X., Adhikary, T., Li, J., Xu, G., Haring, H.U., Schleicher, E., Lehmann, R., Weigert, C., 2016. Lysophosphatidylcholines activate PPARdelta and protect human skeletal muscle cells from lipotoxicity. Biochim. Biophys. Acta 1861 (12), 1980–1992.

Kougias, P., Chai, H., Lin, P.H., Lumsden, A.B., Yao, Q., Chen, C., 2006. Lysophosphatidylcholine and secretory phospholipase A2 in vascular disease mediators of endothelial dysfunction and atherosclerosis. Med. Sci. Monit. 12 (1), RA5–16.

Lin, S., Shi, Q., Nix, F.B., Styblo, M., Beck, M.A., Herbin-Davis, K.M., Hall, L.L., Simeonsson, J.B., Thomas, D.J., 2002. A novel S-adenosyl-L-methionine:arsenic(III) methyltransferase from rat liver cytosol. J. Biol. Chem. 277 (13), 10795–10803.

Liu, R., Wang, X., Aihara, K., Chen, L., 2014. Early diagnosis of complex diseases by molecular biomarkers, network biomarkers, and dynamical network biomarkers. Med. Res. Rev. 34 (3), 455–478.

Lu, K., Abo, R.P., Schlieper, K.A., Graffam, M.E., Levine, S., Wishnok, J.S., Swenberg, J.A., Tannenbaum, S.R., Fox, J.G., 2014. Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis. Environ. Health Perspect. 122 (3), 284–291.

Massaro, M., Carluccio, M.A., De Caterina, R., 1999. Direct vascular antiatherogenic effects of oleic acid: a clue to the cardioprotective effects of the Mediterranean diet. Cardiologia 44 (6), 507–513.

Matschullat, J., 2000. Arsenic in the geosphere–a review. Sci. Total Environ. 249 (1–3), 297–312.

Mckean, C.M., 1972. The effects of high phenylalanine concentrations onserotonin and catecholamine metabolism in the human brain. Brain Res. 47 (42), 469–476.

Mellor, A.L., Sivakumar, J., Chandler, P., Smith, K., Molina, H., Mao, D., Munn, D.H., 2001. Prevention of T cell–driven complement activation and inflammation by tryptophan catabolism during pregnancy. Nature 2, 64–68.

Mendez, M.A., Gonzalez-Horta, C., Sanchez-Ramirez, B., Ballinas-Casarrubias, L., Ceron, R.H., Morales, D.V., Terrazas, F.A.B., Ishida, M.C., Gutierrez-Torres, D.S., Saunders, R.J., Drobna, Z., Fry, R.C., Buse, J.B., Loomis, D., Garcia-Vargas, G.G., Del Razo, L.M., Styblo, M., 2016. Chronic exposure to arsenic and markers of cardiometabolic risk: a cross-sectional study in Chihuahua, Mexico. Environ. Health Perspect. 124 (1), 104–111.

Minatel, B.C., Sage, A.P., Anderson, C., Hubaux, R., Marshall, E.A., Lam, W.L., Martinez, V.D., 2018. Environmental arsenic exposure: from genetic susceptibility to pathogenesis. Environ. Int. 112, 183–197.

Miyata, T., Sonoda, K., Tomikawa, J., Tayama, C., Okamura, K., Maehara, K., Kobayashi, H., Wake, N., Kato, K., Hata, K., Nakabayashi, K., 2015. Genomic, epigenomic, and transcriptomic profiling towards identifying omics features and specific biomarkers that distinguish uterine leiomyosarcoma and leiomyoma at molecular levels. Sarcoma 2015, 412068.

Munder, P.G., Modolell, M., Andreesen, R., Weltzien, H.U., Westphal, O., 1979. Lysophosphatidylcholine (lysolecithin) and its synthetic analogues. Immunemodulating and other biologic effects. Springer Semin. Immunopathol. 2 (2), 187–203.

Nagato, E.G., D'Eon J, C., Lankadurai, B.P., Poirier, D.G., Reiner, E.J., Simpson, A.J., Simpson, M.J., 2013. (1)H NMR-based metabolomics investigation of Daphnia magna responses to sub-lethal exposure to arsenic, copper and lithium. Chemosphere 93 (2), 331–337.

Nicholson, J.K., Wilson, I.D., 2003. Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. Nat. Rev. Drug Discov. 2 (8), 668–676.

Nickel, J.C., Rizzo, M., Marchetti, F., Travaglini, F., Trinchieri, A., 2004. Prevalence, characterization, diagnosis and treatment of the prostatitis patient in Italy: an opportunity to compare the European prostatitis patient to the North American experience. J. Urol. 171 (4), S27.

Piszcz, J., Lemancewicz, D., Dudzik, D., Ciborowski, M., 2013. Differences and similarities between LC-MS derived serum fingerprints of patients with B-cell malignancies. Electrophoresis 34 (19), 2857–2864.

Psychogios, N., Hau, D.D., Peng, J., Guo, A.C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., Young, N., Xia, J., Knox, C., Dong, E., Huang, P., Hollander, Z., Pedersen, T.L., Smith, S.R., Bamforth, F., Greiner, R., McManus, B., Newman, J.W., Goodfriend, T., Wishart, D.S., 2011. The human serum metabolome. PLoS One 6 (2), e16957.

Rachelefsky, G., 1997. Childhood asthma and allergic rhinitis: the role of leukotrienes. J. Pediatr. 131 (3), 348–355.

Sakuradani, E., Kamada, N., Hirano, Y., Nishihara, M., Kawashima, H., Akimoto, K., Higashiyama, K., Ogawa, J., Shimizu, S., 2002. Production of 5,8,11-eicosatrienoic acid by a delta5 and delta6 desaturation activity-enhanced mutant derived from a delta12 desaturation activity-defective mutant of Mortierella alpina 1S-4. Appl. Microbiol. Biotechnol. 60 (3), 281–287.

Sreekumar, A., Poisson, L.M., Rajendiran, T.M., Khan, A.P., Cao, Q., Yu, J., Laxman, B., Mehra, R., Lonigro, R.J., Li, Y., Nyati, M.K., Ahsan, A., Kalyana-Sundaram, S., Han, B., Cao, X., Byun, J., Omenn, G.S., Ghosh, D., Pennathur, S., Alexander, D.C., Berger, A., Shuster, J.R., Wei, J.T., Varambally, S., Beecher, C., Chinnaiyan, A.M., 2009. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature 457 (7231), 910–914.

Vlaanderen, J., Moore, L.E., Smith, M.T., Lan, Q., Zhang, L., Skibola, C.F., Rothman, N., Vermeulen, R., 2010. Application of OMICS technologies in occupational and environmental health research; current status and projections. Occup. Environ. Med. 67 (2), 136–143.

Waghe, P., Sarkar, S.N., Sarath, T.S., Kandasamy, K., Choudhury, S., Gupta, P., Harikumar, S., Mishra, S.K., 2017. Subchronic arsenic exposure through drinking water alters lipid profile and electrolyte status in rats. Biol. Trace Elem. Res. 176 (2), 350–354.

Wallace, M., Morris, C., O'Grada, C.M., Ryan, M., Dillon, E.T., Coleman, E., Gibney, E.R., Gibney, M.J., Roche, H.M., Brennan, L., 2014. Relationship between the lipidome, inflammatory markers and insulin resistance. Mol. Biosyst. 10 (6), 1586–1595.

Wang-Sattler, R., Yu, Z., Herder, C., Messias, A.C., Floegel, A., He, Y., Heim, K., Campillos, M., Holzapfel, C., Thorand, B., Grallert, H., Xu, T., Bader, E., Huth, C., Mittelstrass, K., Doring, 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, H.G., de Angelis, M.H., Rathmann, W., Suhre, K., Prokisch, H., Peters, A., Meitinger, T., Roden, M., Wichmann, H.E., Pischon, T., Adamski, J., Illig, T., 2012. Novel biomarkers for pre-diabetes identified by metabolomics. Mol. Syst. Biol. 8, 615.

Wang, C., Feng, R., Li, Y., Zhang, Y., Kang, Z., Zhang, W., Sun, D.J., 2014. The metabolomic profiling of serum in rats exposed to arsenic using UPLC/Q-TOF MS. Toxicol. Lett. 229 (3), 474–481.

Wang, T.J., Larson, M.G., Vasan, R.S., Cheng, S., Rhee, E.P., McCabe, E., Lewis, G.D., Fox, C.S., Jacques, P.F., Fernandez, C., O'Donnell, C.J., Carr, S.A., Mootha, V.K., Florez, J.C., Souza, A., Melander, O., Clish, C.B., Gerszten, R.E., 2011. Metabolite profiles and the risk of developing diabetes. Nat. Med. 17 (4), 448–453.

Wang, X., Mu, X., Zhang, J., Huang, Q., Alamdar, A., Tian, M., Liu, L., Shen, H., 2015. Serum metabolomics reveals that arsenic exposure disrupted lipid and amino acid metabolism in rats: a step forward in understanding chronic arsenic toxicity. Metallomics 7 (3), 544–552.

Wong, W.T., Ismail, M., Tohit, E.R., Abdullah, R., Zhang, Y.D., 2016. Attenuation of thrombosis by crude rice (Oryza sativa) bran policosanol extract: ex vivo platelet aggregation and serum levels of arachidonic acid metabolites. Evid. Based Complement. Alternat. Med. 2016, 7343942.

WS/T211-2001, 2001. Diagnostic Criteria of Endemic Arsenism in China. Standards press of China, Beijing.

Wu, F., Chi, L., Ru, H., Parvez, F., Slavkovich, V., Eunus, M., Ahmed, A., Islam, T., Rakibuz-Zaman, M., Hasan, R., Sarwar, G., Graziano, J.H., Ahsan, H., Lu, K., Chen, Y., 2018. Arsenic exposure from drinking water and urinary metabolomics: associations and long-term reproducibility in Bangladesh adults. Environ. Health Perspect. 126, 017005.

Wu, L.L., Chiou, C.C., Chang, P.Y., Wu, J.T., 2004. Urinary 8-OHdG: a marker of oxidative stress to DNA and a risk factor for cancer, atherosclerosis and diabetics. Clin. Chim. Acta 339 (1–2), 1–9.

You, J.C., Yang, J., Fang, R.P., Hu, N., Zhang, X.D., Zhang, W.Y., Ye, L.H., 2015. Analysis of phosphatidylcholines (PCs) and lysophosphatidylcholines (LysoPCs) in metastasis of breast cancer cells. Prog. Biochem. Biophys. 42 (6), 563–573.

Zanchi, D., Meyer-Gerspach, A.C., Suenderhauf, C., Janach, K., le Roux, C.W., Haller, S., Drewe, J., Beglinger, C., Wolnerhanssen, B.K., Borgwardt, S., 2016. Differential effects of L-tryptophan and L-leucine administration on brain resting state functional networks and plasma hormone levels. Sci. Rep. 6, 35727.

Zhang, J., Shen, H., Xu, W., Xia, Y., Barr, D.B., Mu, X., Wang, X., Liu, L., Huang, Q., Tian, M., 2014. Urinary metabolomics revealed arsenic internal dose-related metabolic alterations: a proof-of-concept study in a Chinese male cohort. Environ. Sci. Technol. 48 (20), 12265–12274.

Zhang, J., Yang, W., Li, S., Yao, S., Qi, P., Yang, Z., Feng, Z., Hou, J., Cai, L., Yang, M., Wu, W., Guo, D.A., 2016. An intelligentized strategy for endogenous small molecules characterization and quality evaluation of earthworm from two geographic origins by ultra-high performance HILIC/QTOF MS(E) and Progenesis QI. Anal. Bioanal. Chem. 408 (14), 3881–3890.

Zhou, Y., Wang, Y., Su, J., Wu, Z., Wang, C., Zhong, W., Liu, X., Cui, L., Zhou, X., Ma, Y., Xin, Y., Zhang, J., Wu, L., Hu, X., Chen, X., Peng, C., Gao, M., 2017. Integration of microRNAome, proteomics and metabolomics to analyze arsenic-induced malignant cell transformation. Oncotarget 8 (53), 90879–90896.