Serum Metabolomic Profiles for Human Gastric Adenocarcinoma Discrimination

Cailin Zhu
the First Affiliated Hospital of Xi'an Jiaotong University
Hui Liu
Chinese Academy of Medical Sciences, Peking Union Medical College
Hongtao Li
Lanzhou Army General Hospital
Weiyun Bi
The Fourth Military Medical University
Bin Bai
The Fourth Military Medical University
Qingchuan Zhao (zhaoqingchuan@aliyun.com)
The Fourth Military Medical University

Research Article

Keywords: metabolomics, mass spectrometry, ROC curve, stomach neoplasms, tumor markers

Posted Date: October 25th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-961365/v1

Serum Metabolomic Profiles for Human Gastric Adenocarcinoma Discrimination

Cailin Zhu[1*], Hui Liu [2*], Hongtao Li[3], Weiyun Bi[4*], Bin Bai[5], ,Qingchuan Zhao[5#]

1. Department of Thoracic Surgery, the First Affiliated Hospital of Xi'an Jiaotong University, No. 277 West Yanta Road, Xi'an, 710061 China

2.Department of Biophysics and Structural Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences & School of Basic Medicine, Peking Union Medical College, Beijing, P.R., China

3.Department of General Surgery, Lanzhou Army General Hospital, Lanzhou Gansu, P.R., China

4. Department of Clinical Skills Training Center, XiJing Hospital, The Fourth Military Medical University, Xi'an, Shaanxi, P.R. China

5. Department of Gastrointestinal Surgery, Xijing Hospital of Digestive Diseases, The Fourth Military Medical University, Xi'an, Shaanxi, P.R. China

*: Contributed equally. #: Corresponding author

Corresponding author: Qingchuan Zhao, Department of Gastrointestinal Surgery, Xijing Hospital of Digestive Diseases. The Fourth Military Medical University, Xi'an, Shaanxi, P.R. 710032, China, E-mail: zhaoqingchuan@aliyun.com. Tel. and Fax: +86-29-8477-1503

Abstract

As one of the most common cancers，Gastric cancer (GC) exhibits high incidence and mortality. Recent studies have indicated that serum

tumor markers can provide valuable diagnostic information for GC. In this study, we examine the clinicopathological significance of preoperative serum tumor markers and serum metabolites on gastric adenocarcinoma patients. We first perform the serum metabolic profiling of 139 gastric adenocarcinoma patients and 156 healthy controls by the Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS). Based on these chemometric results, we then identify three metabolites, namely phosphatidylcholine (PC) (34:1), palmitoyl-L-carnitine and m/z 361.234, as "potential marker metabolites". Finally, we perform the ROC curve analysis on the three potential marker metabolites to distinguish the 12 early gastric adenocarcinoma patients from healthy controls. Our analysis has yielded an AUC of 0.973, with 91.7% on sensitivity and 93.6% on specificity at the best cut-off point. Our study has empirically shown that serum metabolomics is amenable for the minimally invasive diagnosis of human gastric adenocarcinoma.

Keywords: metabolomics; mass spectrometry; ROC curve; stomach neoplasms；tumor markers;

Introduction

Gastric cancer (GC) is the fourth most common malignancy and the second leading cause of cancer death worldwide [1]. The circumstance is even more severe in most Asian countries, such as China, Japan and Korea [2]. However, due to its relatively asymptomatic property in the early stage

and the lack of adequate screening methods, the majority of GC patients can only been diagnosed at the advanced stage, after the occurrence of distant or local metastases. Although combinatorial chemotherapy has been well developed, surgical resection is still the only curative therapy available [3], and its 5-year survival rate remains disappointingly low [4]. Thus, the most effective approach to improve the patient's prognosis is by GC carcinogenesis at the early stage. Currently, gastroscopy is widely used for early screening. Unfortunately, it involves invasive procedures and requires skilled expertise of endoscopists, thus its cost effectiveness still remains disputable. Therefore, it is important to establish a sensitive, noninvasive, reliable serum screening mechanism by using serum biomarkers as the risk indicator of GC.

As the downstream of the transcriptome and proteome, the metabolome is considered to be an important complement to genomics, transcriptomics, and proteomics [5]. Metabolites are the end products of cellular regulatory processes, and their levels reflect the ultimate response of biological systems to genetic or environmental changes [6]. Metabolomics demonstrates high sensitivity in that it is able to amplify the smallest changes in genes and proteins. This technique has been previously shown to be an effective tool for disease diagnosis [7-10], biomarker screening [11-14] and characterization of biological pathways [15]. The most commonly used techniques for characterization studies include

nuclear magnetic resonance (NMR) and mass spectrometry (MS) [16]. Having the highest resolving power of all mass analyzers, the FTICR-MS analyzer has been proven to be highly effective for the analysis of complex mixtures due to its high sensitivity, wide mass and dynamic ranges [17-21].

In this study, we have coupled the FTICR-MS technology with the multivariate statistical analysis to profile serum samples obtained from gastric adenocarcinoma patients and healthy controls. By pattern recognition, we aim to establish a diagnosis model based on the potential marker metabolites.

Materials and Methods

Clinical samples This study has been approved by the ethics committee of Xijing hospital. Informed consent was obtained from all participants, and the research method complies with Chinese laws and regulations. We collected the Serum samples of 139 gastric adenocarcinoma patients from Xijing hospital of digestive diseases. All the patients had primary gastric adenocarcinoma and were not on any medical treatment prior to sample collection. They were compared to a control group consisting of 156 healthy individuals, whose information was collected from the medical examination center at Xijing hospital. Healthy controls were selected via a health check, including blood tests, endoscopic examinations and diagnostic imaging. We excluded the individuals, whose

diagnosis indicated requiring therapy, detailed examination, and/or observations, from the healthy control group. The specific information about the tumor stage and histological differentiation has been shown in Table 1. We have confirmed the diagnosis for all the patients by pathological examinations, and obtained the tissue samples during surgery. The anatomic stage and histological grade were determined according to the American Joint Committee on Cancer (AJCC) TNM staging classification and histological grading for carcinoma of the stomach (Seventh Edition 2010). All serum samples were collected at 6:00 in the morning before breakfast. Standard blood processing for FTICR-MS analysis was performed in the following manner: we collected whole blood by blood collection tubes and stored them at room temperature for up to 2h; we then centrifuged the samples at 4°C, 2000 rpm for 8 min; the corresponding serum was finally aliquoted into micro-centrifuge tubes and quickly frozen and stored in a -80°C freezer for further analysis. To separate the serum from the blood cells, the blood samples for the tumor marker assays were centrifuged at 1000 × *g* for 10 min. Serum CEA, AFP, CA19-9, and CA12-5 levels were measured using the Roche cobas Elecsys 601 Automatic electrochemiluminescence immunoassay analyzer. The cut-off values for serum CEA, AFP, CA19-9, and CA12-5 were set at 5 ng/ml, 8.1 ng/ml, 20 U/ml and 35 U/ml, respectively.

Chemicals We have purchased Acetonitrile from Thermo Fisher ,

formic acid from TEDIA, Reserpine from Sigma (America), Phosphatidylcholine (PC) (34:1) from Avanti polar lipids (Alabaster AL), and Palmitoyl-L-carnitine chloride from Sigma (America). We filtered the distilled water through a Milli-Q system (Millipore, Billerica, MA).

Serum samples preparation. After thawing the serum on ice, we added acetonitrile (400 μl) to the serum (100 μl) in a 2 ml microfuge tube. The sample was then vortex-mixed for 45 seconds and stored at -20°C overnight. The mixtures were centrifuged at 14,000 rpm for 10 min at 4°C, and then 10 μl of the supernatant from each sample was vortex-mixed with 190 μl of the internal standard solution (54 ng/ml of reserpine in water; 50% acetonitrile (v/v); 0.2% formic acid (v/v)). Each sample was stored in an auto-sampler. Quality control (QC)[22] samples were prepared by mixing equal amounts of the serum samples from 30 healthy controls. All the serum samples were randomized along with 20 QC samples. The QC samples were inserted and analyzed in every 15 samples to ensure the stability and repeatability of the MS system.

FTICR-MS analysis Metabolite analysis was performed using a 9.4 T apex-ultra™ hybrid Qh-FTICR mass spectrometer (Bruker Daltonics, Billerica, MA) equipped with an electrospray ionization source. The sample was introduced into the ion source by a syringe pump at a flow rate of 3 μl/min. The mass spectrometer was operated in the positive ion mode with the capillary voltage of 4000 V, the

spray shield of 3500 V, the drying gas flow of 5.0 L/min, the drying gas temperature of 180 °C, the nebulizing gas flow of 1.0 L/min, the time of flight of 0.0012 s, the source accumulation of 0.05 s and the ion accumulation time of 2.5 s. The mass spectrum was acquired from 4 scans for an averaged spectrum in the m/z range of 100 – 1000 with 2 M acquisition size, and the max resolution was 260,000 at m/z 400. In the tandem MS experiments, argon was used as the collision gas.

Data preprocessing We obtained the Raw MS data by the Apex Control 3.0.0 software (Bruker Daltonics, Billerica, MA). To detect the peak, we executed the FTMS peaks finder algorithm with the signal-to-noise ratio set at >5, the relative intensity at 0.2% and the absolute intensity thresholds at 10,000. For peak detection and alignment, we removed the variables having too many missing values according to the 80% principle[23], and only kept the variables with more than 80% non-zero measurement values in the control and gastric adenocarcinoma group on the peak list. The missing values were filled with the mean of the variable, and the peak intensities of all the metabolites were normalized to the total intensity of the sample [24]. After normalization, the data were then exported to the SIMCA-P v 11.5 software (Umetrics AB, Sweden) for multivariate data analysis. Unsupervised principal component analysis (PCA) was initially used to visualize the general separation. At the next step, supervised orthogonal partial least squares-discriminant analysis

(OPLS-DA) was performed to model the difference between the GC patients and healthy controls. We selected Potential biomarkers based on the Variable Importance in the Project (VIP) and Z-score [Z-score $= \frac{X_i - \overline{X}_{control}}{\mathrm{SD}_{\mathrm{control}}}$] [25]. Furthermore, we performed an independent t-test using SPSS 16.0 for windows to confirm the significant differences of the marker metabolites between the gastric adenocarcinoma and healthy control group, with the level of statistical significance set as $p<0.05$. The receiver operating characteristic (ROC) curve was employed to demonstrate the distinguishing ability of the potential marker metabolites. We calculated the area under the curve (AUC) for the ROC curve for the combination of potential marker metabolites. To identify the potential marker metabolites, we used the METLIN database.

Results

Validation of analytical performance Our PCA analysis performed on all of the samples have revealed that the QC samples were tightly clustered in PCA scores plots. This observation confirmed the reproducibility of our method. We have also cross-validated the variations in the levels of metabolites in the QC samples. It can be observed that the relative standard deviation (RSD) of all the metabolite peaks in the QC samples ranged from 3.17 % to 17.57 %. These results clearly demonstrated the stability and reproducibility of FTICR MS measurement.

Serum Metabolomics of Human Gastric Adenocarcinoma A total

of 218 aligned individual peaks were obtained from samples after the removal of internal standards due to lack of too many value variables. It can be observed that the gastric adenocarcinoma group and healthy controls were not well distinguished in the PCA score plots. The likely reason is that PCA was performed using the unsupervised data analysis method, which separates the samples based on their essential attributes; however, the obtained human serum samples are extremely complex. To address the limitation of PCA, we used the more advanced method of OPLS-DA for further analysis, which makes use of class information to maximize the separation among the different classes. It can be observed that the OPLS-DA model exhibits satisfactory modeling and predictive ability to separate the GC and healthy control samples ($R^2X_{(cum)}$ =0.494, $R^2Y_{(cum)}$=0.889, $Q^2_{(cum)}$=0.813). All the OPLS-DA models have been validated with a test of 200 times permutation, in which all the $R^2_{(cum)}$ and $Q^2_{(cum)}$ values calculated from the permuted data were lower than the original ones in the validated plot, and all the $Q^2_{(cum)}$ intercepted the y-axis below zero (as shown in Figure 2). To test the prediction ability of this OPLS-DA model, the gastric adenocarcinoma and healthy control samples were randomly and equally divided into two sets. In total, 80% of the samples were used as the training set, and the remaining 20% were used as the test set. All the samples in the training set were correctly categorized to the control group or gastric adenocarcinoma group by the OPLS-DA model

(as shown in Figure 1). On the test set, only one sample in the test set (from the healthy control group) was categorized to the wrong group. Accordingly, the achieved sensitivity and specificity are 100% and 96.7% respectively. However, the test failed to distinguish between different pathological stages (I-IV) , T stages (1-4) and histological grades (well, moderately, and poorly, respectively) of gastric adenocarcinoma serum.

Selection of potential marker metabolites According to the Variable Importance in the Project (VIP), we selected a total of 72 variables with the VIP value greater than 1 (VIP>1). The independent t-test results indicated that all the 72 variables had significantly different values ($p<0.05$) between the gastric adenocarcinoma and healthy control samples. Furthermore, by the rule of Z-score, which specifies that the variable whose absolute values of the Z-score are greater than 2 in more than 1/3 of the samples in the experimental group can be defined as a potential marker metabolite[26], we identified only three variables as potential marker metabolites. We annotated these serum metabolites using the METLIN database and further analyzed them using MS/MS. The m/z 760.585 and m/z 400.341 were phosphatidylcholine (PC) (34:1) and palmitoyl-L-carnitine. However, the m/z 361.234 was not identified. We thus performed Binary logistic regression [27] and ROC analysis to identify the potential marker metabolites and examine their ability to predict gastric adenocarcinoma [28]. As shown in Table 2, the results of the ROC analysis

revealed that the variable was accurately ranked according to their AUC. At the best cut-off point, the achieved sensitivity and specificity were 91.4% and 91.7% respectively (as shown in Figure 3). The ROC curve analyses of the three potential marker metabolites resulted in an AUC of 0.971. To distinguish the 12 early gastric adenocarcinoma patients from healthy controls, the ROC curve analyses of the three potential marker metabolites generated an AUC of 0.973 (as shown in Table 3). In addition, it achieved the sensitivity of 91.4% and the specificity of 91.7% at the best cut-off point (as shown in Figure 4).

Positive rates of tumor markers It can be observed that the preoperative serum positive rates of CEA, AFP, CA19-9, and CA12-5 were 15.1%, 5.8%, 17.3%, and 3.6%, respectively. The serum value of CEA was ranged from 0.2 to 117.3 ng/ml (with the mean of 5.22 ng/ml, and the median of 1.15 ng/ml), AFP from 0.24 to 248.48 ng/ml (with the mean of 4.43 ng/ml, and the median of 1.37 ng/ml), CA19-9 from 0.6 to 651.1 U/ml (with the mean of 26.95 U/ml, and the median of 5.85 U/ml), and CA12-5 from 1 to 61.7 U/ml (with the mean of 8.84 U/ml, and the median of 5.92 U/ml).

Discussion

Since Gold and Freedman first reported the discovery of carcinoembryonic antigen (CEA) in 1965 [29], various serum biomarkers, such as CEA [30], CA19-9[31], CA72-4[31-33], CA50[34], CA125[35] and

CA242[36], have been investigated for their applicability and feasibility in gastric cancer screening. Because none of the individual biomarkers could show sufficient diagnostic value, gastric cancer screening usually requires a combination of multiple biomarkers, and may also involve the exploration of new biomarkers[37]. Using the highly sensitive characterization techniques, serum characterization profiles provide the good opportunity to explore new marker metabolites for the early diagnosis of human gastric cancer.

Previous work on human GC metabolomics [38- 42] have proposed models to separate cancer from healthy controls and cancer tissues from normal mucosae. However, these studies were based on small-sized samples, and none of them attempted the early diagnosis of GC. In contrast, our study was based on the much larger-sized samples. It used FTICR-MS to compare the serum metabolic fingerprints of 139 gastric adenocarcinoma patients to 156 healthy controls. In the FTICR MS analysis, we used the biological QC and PCA to ensure the validity of the data. The observed close clustering of QC samples have validated the robustness of our metabolic profiling platform with regard to various parameters, including the stability of the detector response, accuracy of the sample injection and reproducibility. In the case of large and complex data, chemometric analysis was usually necessary for obtaining an interpretable model for the complex intercorrelation between data [43]. The

results obtained from the permutation tests have shown that our proposed supervised chemometric model was valid and the separation of gastric adenocarcinoma patients and healthy controls was not due to chance correlation. The lower variations in the level of QC samples further confirmed that the perturbations of the metabolites were due to biological differences, but not chance.

In the exploration of disease mechanism, it is usually necessary to identify as many differential metabolites as possible from a variety of pathways. In contrast, for effective diagnosis of diseases, it is usually required to curtail the number of biomarkers and maintain the constance of diagnosis ability. In this study, three identified variables were selected based on the VIP values (VIP>1) estimated by the OPLS-DA model, with the *p*-values ($p<0.05$) obtained from Student's t test. We found that 72 ions are of great value for the detection of gastric adenocarcinoma. Using the Z-score analysis, we successfully narrowed down these 72 metabolites to only three. All the three serum metabolites were then annotated using the METLIN database, and two of them were confirmed using MS/MS and standards. Finally, we used the binary logistic regression method to evaluate the capacity of three potential marker metabolites as the indicators of gastric adenocarcinoma. Our results have demonstrated that the model can achieve high sensitivity and specificity. They provided the preliminary evidence for these metabolites to be used as effective markers

for the diagnosis of gastric adenocarcinoma.

Wu et al. reported that a diagnostic model of 18 metabolites might be used to distinguish gastric cancer from normal mucosae [44] with the AUC value of the ROC at 0.963. More recently, Song et al. proposed a diagnostic model based on 44 metabolites that achieved the sensitivity and specificity of 93.3% (28/30) and 96.7% (29/30) respectively[45]. Our current study achieved the similar distinguishing power based on only three metabolites. Its AUC of the ROC for a discriminant score was 0.971, and its sensitivity and specificity were 91.4% and 91.7% respectively. Furthermore, it can distinguish 12 early gastric adenocarcinoma from healthy controls with the AUC of 0.973, the sensitivity of 91.4% and the specificity of 91.7%. It can be reasonably inferred that either metabolic perturbations are not directly associated with the different pathological stages of gastric adenocarcinoma, or when tumor occurs, it demonstrates a different metabolic profile from normal tissue.

In our study, significantly increased levels of phosphatidylcholine (PC) (34:1), palmitoyl-L-carnitine and m/z 361.234 were observed in the serum samples obtained from gastric adenocarcinoma patients compared to those obtained from the healthy controls. Phosphatidylcholine (PC) is the major phospholipid of eukaryotic cell membranes. Its increased concentration in the tumor indicates an intensified cell membrane synthesis and accelerated cell replication [46, 47]. Carnitine could transport

activated fatty acids from the cytosol to the mitochondria for β-oxidation and reserve energy in the form of acylcarnitine. Tumorigenesis is a process of high-energy expenditure, which may elevate the original concentration of carnitine to a new level [48].

The results of the present characterization study have demonstrated that the proposed diagnosis model based on three potential marker metabolites exhibits high sensitivity and specificity. Thus, the validation of these results using large-scaled serum samples obtained from early gastric cancer patients is very important. In addition, these metabolic features indicated the perturbation of lipid metabolism in gastric adenocarcinoma, which can provide useful clues on how to further explore the mechanism and identify the therapeutic targets of gastric adenocarcinoma.

Acknowledgments

This study was funded by grant 81170432, which was obtained from the National Natural Science Foundation of China.

Data availability statement

The raw data used to support the findings of this study are restricted by the Xijing Hospital Ethics Committee in order to protect patient privacy. Data are available from me for researchers who meet the criteria for access to confidential data.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68(6):394-424.

2. Irino T, Takeuchi H, Terashima M, et al. Gastric Cancer in Asia: Unique Features and Management. Am Soc Clin Oncol Educ Book 2017; 37:279-291.

3. Cai Z, Yin Y, Shen C, et al. Comparative effectiveness of preoperative, postoperative and perioperative treatments for resectable gastric cancer: A network meta-analysis of the literature from the past 20 years. Surg Oncol 2018; 27(3):563-574.

4. Aoyama T, Yoshikawa T. Adjuvant therapy for locally advanced gastric cancer. Surg Today 2017; 47(11):1295-1302.

5.Nicholson, J. K.; Lindon, J. C. Systems biology: Metabonomics. Nature 2008, 455 , 1054-6.

6. Holmes, E.; Wilson, I. D.; Nicholson, J. K., Metabolic phenotyping in health and disease. *Cell* **2008,** 134, (5), 714-7.

7.Mlynarz, P.; Barg, W.; Deja, S.; Jankowska, R., Application of characterizi in COPD diagnosing. *Pol Merkur Lekarski* **2012,** 33, (196), 207-12.

8. Pan X, Fei G, Lu J, et al. Measurement of Blood Thiamine Metabolites for Alzheimer's Disease Diagnosis. EBioMedicine 2016; 3:155-162.

9. Dona AC, Coffey S, Figtree G. Translational and emerging clinical applications of metabolomics in cardiovascular disease diagnosis and treatment. Eur J Prev Cardiol 2016; 23(15):1578-89

10 Li, F.; Qin, X.; Chen, H.; Qiu, L.; Guo, Y.; Liu, H.; Chen, G.; Song, G.; Wang, X.; Li, F.; Guo, S.; Wang, B.; Li, Z., Lipid profiling for early diagnosis and progression of colorectal cancer using direct-infusion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom 2013, 27, (1), 24-34.

11 Emwas AH, Roy R, McKay RT, et al. Recommendations and Standardization of Biomarker Quantification Using NMR-Based Metabolomics with Particular Focus on Urinary Analysis. J Proteome Res 2016; 15(2):360-73.

12.Suzuki, N., Mass spectrometry-based quantitative analysis and biomarker discovery. *Yakugaku Zasshi* **2011,** 131, (9), 1305-9.

13.Laborde, C. M.; Mourino-Alvarez, L.; Akerstrom, F.; et al., Potential blood biomarkers for stroke. Expert Rev Proteomics 2012, 9, (4), 437-49.

14. Mishur, R. J.; Rea, S. L., Applications of mass spectrometry to metabolomics

and metabonomics: detection of biomarkers of aging and of age-related diseases. *Mass Spectrom Rev* **2012,** 31, (1), 70-95.

15.Nicholson, J. K.; Connelly, J.; Lindon, J. C.; Holmes, E., Metabonomics: a platform for studying drug toxicity and gene function. Nat Rev Drug Discov 2002, 1, (2), 153-61.

16. Madsen, R.; Lundstedt, T.; Trygg, J., Chemometrics in metabolomics—a review in human disease diagnosis. Anal Chim Acta 2010, 659, (1-2), 23-33.

17.Allwood, J. W.; Parker, D.; Beckmann, M.; et al., Fourier Transform Ion Cyclotron Resonance mass spectrometry for plant metabolite profiling and metabolite identification. *Methods Mol Biol* **2012,** 860, 157-76.

18. Payne, T. G.; Southam, A. D.; Arvanitis, T. N.; Viant, M. R., A signal filtering method for improved quantification and noise discrimination in fourier transform ion cyclotron resonance mass spectrometry-based metabolomics data. *J Am Soc Mass Spectrom* **2009,** 20, (6), 1087-95.

19. Miura, D.; Tsuji, Y.; Takahashi, K.; et al., A strategy for the determination of the elemental composition by fourier transform ion cyclotron resonance mass spectrometry based on isotopic peak ratios. *Anal Chem* **2010,** 82, (13), 5887-91.

20. Hao-Yang Wang, Xu Chu, Zhi-Xiong Zhao; et al., Analysis of low molecular weight compounds by MALDI-FTICR-MS. J Chromatogr B 2011, 879, 1166-79.

21. Lorkiewicz, P.; Higashi, R. M.; Lane, A. N.; Fan, T. W., High information throughput analysis of nucleotides and their isotopically enriched isotopologues by direct-infusion FTICR-MS. Metabolomics 2012, 8, (5), 930-939.

22.Pasikanti, K. K.; Esuvaranathan, K.; Ho, P. C.; Mahendran, R.; et al., Noninvasive urinary metabonomic diagnosis of human bladder cancer. *J Proteome Res* **2010,** 9, (6), 2988-95

23. Bijlsma, S.; Bobeldijk, I.; Verheij, E. R.; et al., Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. *Anal Chem* **2006,** 78, (2), 567-74.

24. Tan, Y.; Yin, P.; Tang, L.; et al., Metabolomics study of stepwise hepatocarcinogenesis from the model rats to patients: potential biomarkers effective for

small hepatocellular carcinoma diagnosis. *Mol Cell Proteomics* **2012,** 11, (2), M111.010694.

25. Sreekumar, A.; Poisson, L. M.; Rajendiran, T. M.; et al., Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* **2009,** 457, (7231), 910-4.

26 Du Z; Zhang, L.; Liu, S., Application of liquid chromatography-mass spectrometry in the study of metabolic profiling of cirrhosis in different grades. *Se Pu* **2011,** 29, (4), 314-9.

27. Xue, A.; Scarlett, C. J.; Chung, L.; et al., Discovery of serum biomarkers for pancreatic adenocarcinoma using proteomic analysis. *Br J Cancer* **2010,** 103, (3), 391-400.

28 Huang, Z.; Lin, L.; Gao, Y.; Chen, Y.; Yan, X.; Xing, J.; Hang, W., Bladder cancer determination via two urinary metabolites: a biomarker pattern approach. Mol Cell Proteomics 2011, 10, (10), M111.007922.

29.Gold, P.; Freedman, S. O., Specific carcinoembryonic antigens of the human digestive system. *J Exp Med* **1965,** 122, (3), 467-81.

30 Chung, H. W.; Kim, J. W.; Lee, J. H.; et al., Comparison of the validity of three biomarkers for gastric cancer screening: carcinoembryonic antigen, pepsinogens, and high sensitive C-reactive protein. *J Clin Gastroenterol* **2009,** 43, (1), 19-26.

31. Kodama, I.; Koufuji, K.; Kawabata, S.; et al., The clinical efficacy of CA 72-4 as serum marker for gastric cancer in comparison with CA19-9 and CEA. *Int Surg* **1995,** 80, (1), 45-8.

32. Ucar, E.; Semerci, E.; Ustun, H.; et al., Prognostic value of preoperative CEA, CA 19-9, CA 72-4, and AFP levels in gastric cancer. *Adv Ther* **2008,** 25, (10), 1075-84.

33.Mattar, R.; Alves, D. A. C.; DiFavero, G. M.; et al., Preoperative serum levels of CA 72-4, CEA, CA 19-9, and alpha-fetoprotein in patients with gastric cancer. *Rev Hosp Clin Fac Med Sao Paulo* **2002,** 57, (3), 89-92.

34. Pectasides, D.; Mylonakis, A.; Kostopoulou, M.; Papadopoulou, M.; Triantafillis, D.; Varthalitis, J.; Dimitriades, M.; Athanassiou, A., CEA, CA 19-9, and CA-50 in monitoring gastric carcinoma. *Am J Clin Oncol* **1997,** 20, (4), 348-53.

35.Bukowski, J.; Gozdz, S.; Sluszniak, J.; Sluszniak, A.; Korejba, W.; Zielinski, A.; Banasinska, E., [Carcinoembryonic antigen, Ca 19-9 and Ca 125 antigen levels in the sera of patients with cancer of the stomach]. *Pol Tyg Lek* **1988,** 43, (5), 136-9.

36. Carpelan-Holmstrom, M.; Louhimo, J.; Stenman, U. H.; Alfthan, H.; Haglund, C., CEA, CA 19-9 and CA 72-4 improve the diagnostic accuracy in gastrointestinal cancers. *Anticancer Res* **2002,** 22, (4), 2311-6.

37.Corona G, Cannizzaro R, Miolo G, Caggiari L, De Zorzi M, Repetto O, Steffan A, De Re V. Use of Metabolomics as a Complementary Omic Approach to Implement Risk Criteria for First-Degree Relatives of Gastric Cancer Patients. INT J MOL SCI 2018, 19(3).

38.Chan AW, Mercier P, Schiller D, et al. (1)H-NMR urinary metabolomic profiling for diagnosis of gastric cancer. Br J Cancer 2016; 114(1):59-62.

39.Kuligowski J, Sanjuan-Herraez D, Vazquez-Sanchez MA, et al. Metabolomic Analysis of Gastric Cancer Progression within the Correa's Cascade Using Ultraperformance Liquid Chromatography-Mass Spectrometry. J Proteome Res 2016; 15(8):2729-38.

40.Wang H, Zhang H, Deng P, et al. Tissue metabolic profiling of human gastric cancer assessed by (1)H NMR. BMC Cancer 2016; 16:371.

41. Liang Q, Wang C, Li B. Metabolomic Analysis Using Liquid Chromatography/Mass Spectrometry for Gastric Cancer. Appl Biochem Biotechnol 2015; 176(8):2170-84.

42.Chan AW, Gill RS, Schiller D, Sawyer MB. Potential role of metabolomics in diagnosis and surveillance of gastric cancer. World J Gastroenterol 2014; 20(36):12874-82.

43. Ikeda, A., et al., Serum metabolomics as a novel diagnostic approach for gastrointestinal cancer. Biomed Chromatogr, 2012. 26,(5), 548-58.

44. Wu, H.; Xue, R.; Tang, Z.; Deng, C.; Liu, T.; Zeng, H.; Sun, Y.; Shen, X., Metabolomic investigation of gastric cancer tissue using gas chromatography/mass spectrometry. *Anal Bioanal Chem* **2010,** 396, (4), 1385-95.

45. Song, H.; Peng, J. S.; Dong-Sheng, Y.; Yang, Z. L.; Liu, H. L.; Zeng, Y. K.; Shi, X. P.; Lu, B. Y., Serum metabolic profiling of human gastric cancer based on gas chromatography/mass spectrometry. Braz J Med Biol Res 2012, 45, (1), 78-85.

46. Ruiz-Cabello, J.; Cohen, J. S., Phospholipid metabolites as indicators of cancer cell function. *NMR Biomed* **1992,** 5, (5), 226-33.

47. Kuang, Y.; Salem, N.; Tian, H.; Kolthammer, J. A.; Corn, D. J.; Wu, C.; Wang, F.; Wang, Y.; Lee, Z., Imaging lipid synthesis in hepatocellular carcinoma with [methyl-11c]choline: correlation with in vivo metabolic studies. *J Nucl Med* **2011,** 52,

(1), 98-106

48. Vaz, F. M.; Wanders, R. J., Carnitine biosynthesis in mammals. Biochem J 2002, 361, (Pt 3), 417-29.

Table 1

Summary of clinical and demographic characteristics of the gastric adenocarcinoma patients and healthy controls used in this study

characteristics	Gastric adenocarcinoma patients	Healthy controls
Gender		
male	100	113
female	39	43
Age(years)		
mean±SD	57.4±12.2	55.2±9.7
Rang	24-79	31-79
AJCC Anatomic stage(2010)		
Ⅰ	23	
Ⅱ	49	
Ⅲ	64	
Ⅳ	3	
AJCC Histologic grade(2010)		
well	23	
moderately	47	
poor	69	

Table 2 Results of ROC analysis of metabolites

Metabolite	AUC	Sensitivity	Specificity
1 (m/z 760.58)	0.870	71.9%	88.5%
2 (m/z 400.34)	0.826	79.9%	68.6%
3 (m/z 361.23)	0.792	49.6%	98.7%
1+2	0.922	87.1%	85.3%

1+3	0.958	89.9%	88.5%
2+3	0.893	77.0%	87.8%
1+2+3	0.971	91.4%	91.7%

Table 3 serum tumor markers and clinicopathologic factors of the patients

Factors(n)	**CEA(+) n%**	***p***	**AFP (+) n%**	***p***	**CA19-9(+)n%**	***p***	**CA12-5 (+) n%**	***p***
Sex		0.319		0.546		0.527		0.922
Male(100)	17(17)		7(7)		16(16)		3(3)	
Female(39)	4(10.3)		1(2.6)		8(20.5)		2(5.1)	
Age(years)		0.063		1.000		0.031		0.408
≤50(36)	2(5.6)		2(5.6)		2(5.6)		0(0)	
>50(103)	19(18.4)		6(5.8)		22(21.4)		5(4.9)	
Histologic grade		0.063		0.134		0.249		1.000
Well(23)	7(30.4)		3(13.0)		5(21.7)		1(4.3)	
Moderately(47)	7(14.9)		3(6.4)		11(23.4)		2(4.3)	
Poor(69)	7(10.1)		2(2.9)		8(11.6)		2(2.9)	
Anatomic stage		0.373		0.639		0.799		0.057
I+II (23+49)	9(12.5)		3(4.2)		13(18.1)		0(0)	
III+IV (64+3)	12(17.9)		5(7.5)		11(16.4)		5(7.5)	
T stage		0.024		0.295		0.006		0.584
T1+T2(12+17)	0(0)		0(0)		0(0)		0(0)	
T3+T4(106+4)	21(19.1)		8(7.3)		24(21.8)		5(4.5)	
N stage		0.336		0.609		0.284		0.052
N0+ N1 (38+35)	9(12.3)		3(4.1)		15(20.5)		0(0)	
N2+ N3 (28+38)	12(18.2)		5(7.6)		9(13.6)		5(7.6)	

Figure 1. The OPLS-DA prediction model of gastric adenocarcinoma. An OPLS-DA model was constructed using data from 111 gastric adenocarcinoma patients and 125healthy controls (training set), this model was then used to predict gastric

adenocarcinoma of a further 59 samples including 28 gastric adenocarcinoma and 31 healthy controls(test set) that were not used in the construction of the model.

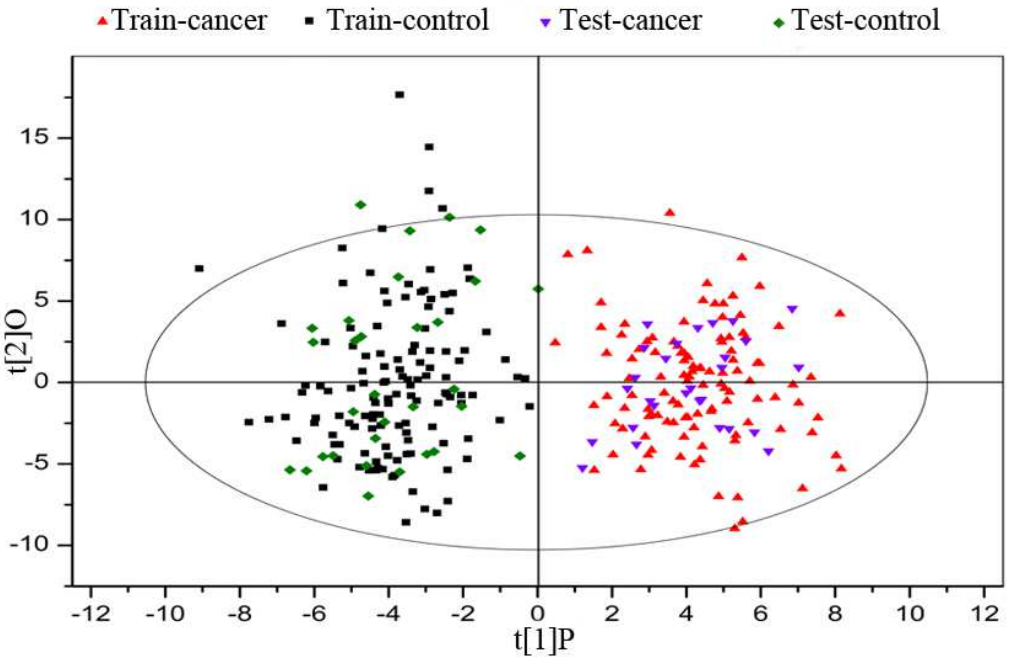

Figure 2. Validation plot obtained from 200 permutation tests

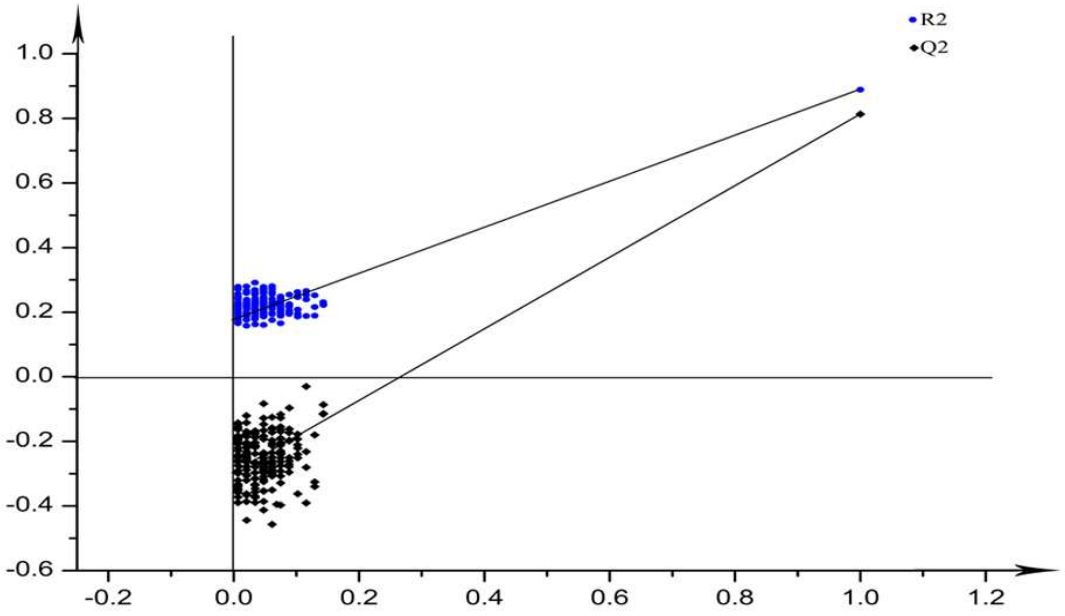

Figure 3. Diagnostic efficacy evaluation using ROC curves for potential biomarkers. 1, 2, 3, and 1+2+3 correspond to metabolite 1, 2, 3, and metabolite 1+2+3 in Table 2.

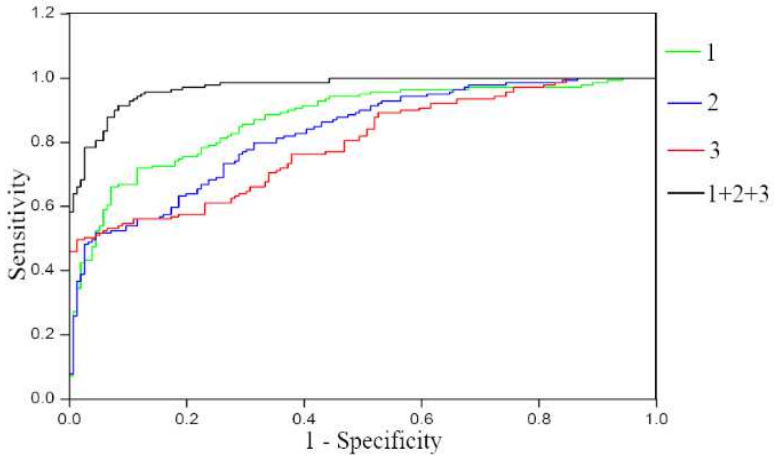

Figure 4. Diagnostic efficacy evaluation for early gastric adenocarcinoma using ROC curves of potential biomarkers. 1, 2, 3, and 1+2+3 correspond to metabolite 1, 2, 3, and metabolite 1+2+3 in Table 3.