Identification of metabolic biomarkers to diagnose epithelial ovarian cancer using a UPLC/QTOF/MS platform

Lijun Fan, Wang Zhang, Mingzhu Yin, Tao Zhang, Xiaoyan Wu, Haiyu Zhang, Meng Sun, Zhenzi Li, Yan Hou, Xiaohua Zhou, Ge Lou & Kang Li


To link to this article: https://doi.org/10.3109/0284186X.2011.648338

Published online: 27 Jan 2012.

Submit your article to this journal

Article views: 2420

View related articles

Citing articles: 8 View citing articles
Identification of metabolic biomarkers to diagnose epithelial ovarian cancer using a UPLC/QTOF/MS platform

LIJUN FAN¹, WANG ZHANG¹, MINGZHU YIN², TAO ZHANG¹, XIAOYAN WU¹, HAIYU ZHANG¹, MENG SUN¹, ZHENZI LI¹, YAN HOU¹, XIAOHUA ZHOU³, GE LOU² & KANG LI¹

¹Department of Epidemiology and Biostatistics, Harbin Medical University, Harbin, China, ²Department of Gynecology Oncology, The Third Affiliated Hospital of Harbin Medical University, Harbin, China and ³Department of Biostatistics, School of Public Health and Community Medicine, University of Washington, Seattle, USA

Abstract

Background. Currently available tests are insufficient to distinguish patients with epithelial ovarian cancer (EOC) from normal individuals. Metabolomics, a study of metabolic processes in biologic systems, has emerged as a key technology in the measurements of small molecular metabolites in tissues or biofluids.

Material and methods. To investigate the application of metabolomics on selecting EOC-associated biomarkers, 173 plasma specimens (80 newly diagnosed EOC patients and 93 normal individuals) were analyzed using ultra-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC/QTOF/MS). A two-step strategy was performed to select EOC-associated biomarkers. The first step was to select potential biomarkers in distinguishing 42 cancer patients from 58 normal controls through partial least-squares discriminant analysis (PLS-DA) and database searching, and the second step was to validate the discrimination performance of these biomarkers in a dataset contained 38 EOCs and 35 controls.

Results. Eight candidate biomarkers were selected. The combination of these biomarkers resulted in the area of receiver operating characteristic curve (AUC) of 0.941, a sensitivity of 0.921, and a specificity of 0.886 at the best cut-off point for detecting EOC.

Discussion. Our findings suggested that sharp differences in metabolic profiles exist between EOC patients and normal controls. The identified eight metabolites associated with EOC may be served as novel biomarkers for diagnosis.

Epithelial ovarian cancer (EOC) is the leading cause of death among gynecologic malignancies due to a lack of highly accurate diagnostic tests [1–3]. In the clinic, a common approach to detecting ovarian cancer is to identify cancer-specific biomarkers in the blood, such as cancer antigen 125 (CA125), carcinoembryonic antigen (CEA), and human epididymis protein 4 (HE4) [4]. However, the identification of ovarian cancer by these biomarkers has been suggested to be high false-positive and false-negative [3]. Thus, it is important to find additional screening and diagnostic biomarkers for detecting EOC accurately.

Metabolomics has offered great promise in many types of cancer as an effective and non-invasive diagnostic method [5,6]. Detecting the metabolites in biofluids is a promising approach for cancer diagnosis [7,8]. For this purpose, a comprehensive analysis by chromatography coupled with mass spectrometry (MS) has been used to study primary metabolic patterns and to discover metabolic biomarkers [9,10]. Some metabolites have been validated as useful biomarkers in many kinds of cancers, and several of them are now being used in cancer diagnosis [9,11,12]. In recent years, metabolomics coupled with multivariate analysis methods have been performed in ovarian cancer diagnosis [13–15]. Odunsi et al. revealed different metabolic patterns between EOC women and healthy women using serum samples analyzed by microflow NMR spectroscopy probe [13]. Garcia et al. showed that NMR spectroscopy based metabolomics of serum analysis could be used to discriminate women with ovarian cancer at early stage I/II from healthy controls [14]. Chen et al. analyzed the serum samples from healthy women, patients with benign
ovarian tumors, and patients with EOC using liquid chromatography-mass spectrometry (LC-MS), and found several diagnostic metabolites [15]. However, all these studies employed serum samples and their results were diverse, and further research revealing metabolic differences in plasma samples between EOC patients and normal women are needed.

To discover potential biomarkers in distinguishing newly diagnosed EOC from normal individuals, we conducted a prospective study with plasma specimens analyzed by ultra-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC/QTOF/MS). And we validated the discrimination performance of the selected biomarkers in an additional independent cohort.

Material and methods

Clinical specimens

The plasma samples for this study were collected from the Third Affiliated Hospital, Harbin Medical University. The institutional review board (IRB) at Harbin Medical University approved this research and all participants signed informed consent forms. According to appropriate protocols, patients of various ages with initially diagnosed pelvic masses were enrolled between August 2009 and November 2010. The patients who were diagnosed as EOC were included in the current study; some subjects were excluded based on the following criteria: 1) those who were unable to undergo an operation or ovarian biopsy; 2) those who also had diabetes mellitus and hepatitis; 3) those who had taken antineoplastic drugs before being hospitalized. The EOC status of a study subject was confirmed by either biopsy or post-operative pathology. Meanwhile, the control samples with no ovarian diseases and no evidence of serious diseases (such as other cancers, diabetes mellitus and hepatitis) were collected from the department of gynecology in the same hospital. This type of controls was defined as normal controls. All participants were not taking any treatment before sample collection. The training set included 42 EOC patients and 38 normal controls, which were enrolled between August 2009 and March 2010. The validation set consisted of 38 EOC patients and 35 normal controls were sequentially enrolled between April 2010 and November 2010.

The age at study enrollment, menopausal status, clinical diagnosis, and pathological report of all participants were obtained when they entered the study. The EOC subjects were categorized based on FIGO (International Federation of Gynecology and Obstetrics) stage classification. Clinical information on the participants is provided in Table I.

Plasma specimens

After fasting and avoiding alcohol and medicine for 12 hours, each participant was collected 5 ml of whole blood before cytoreductive surgery and any radiotherapy or chemotherapy by venipuncture. Then the blood was injected into the evacuated blood collection EDTA-K3 (sodium dihydrogen phosphate) tubes. After centrifuged within one hour, the plasma were separated into five 400-μl portions and stored at -80°C until assayed. After the cytoreductive surgeries and histopathological examinations, all samples were taken into the UPLC/QTOF/MS analysis.

Sample preparation and pretreatment

After being thawed and re-cooling in a 4°C water bath, a 300-μl aliquot of the plasma sample was injected into a vial and vortexed for 30 seconds. The solution was extracted with 1200 μl of methanol and vortexed for two minutes. After incubation for 10 minutes at 20°C, the samples were centrifuged at 14000 g for 10 minutes. An aliquot of the supernatant was transferred to a sampling vial to nitrogen-dry at 37°C. The residue was derivatized using a two-step procedure. Three hundred microliters of a mixture of acetonitrile/water (3:1) were added to the vial, which was then kept at 30°C for 10 minutes, followed by vortexing for 60 seconds. The supernatant was then placed into the sampling vial pending UPLC/QTOF/MS analysis.

Chromatography

A 2 μl aliquot of the pretreated sample was injected into a 100 mm × 2.1 mm, 1.7 μm BEH C18 column (Waters, Milford, USA) held at 35°C using an Acquity ultraperformance liquid chromatography system (Waters). The samples were run alternately between five controls and five EOC samples.

Mass spectrometry

Mass spectrometry was performed on a Waters Micromass Q-TOF (Manchester, UK) equipped with an electrospray ionization source (ESI) operating. The source temperature was set at 100°C with a cone gas flow rate of 25 l/h. Meanwhile, the desolvation gas temperature was 300°C with a desolvation gas flow rate of 650 l/h. The capillary voltage and the cone voltage were set at 2.8 kV and 35 V, respectively. Centroid data were collected from 50 to 1000 M/Z with a scan time of 0.4 seconds and an interscan delay of 0.1 seconds and a lockspray frequency of 10 seconds.
Data process

MassLynx software (Waters) was used for system controlling and data acquisition. The UPLC/QTOF/MS raw data were processed by the MarkerLynx Applications Manager Version 4.1 using the following parameters: initial retention time: 0 minutes; final retention time: 10.5 minutes; mass range: 50–1000; an intensity threshold: 10; noise level: 6.00.

The signals in nine quality control samples (mixtures of the plasma samples) were used to describe the analytical characteristics of metabolic profiling of the UPLC/QTOF/MS system. The relative standard deviation (RSD) value of mean, median and 95% peak areas was 9%, 7% and 1.7–24%, respectively.

Statistical analyses

The statistical analysis was divided into preliminary “training” step and subsequent validation step. In the “training” step, partial least-squares discriminant analysis (PLS-DA) was performed to select the specific metabolites associated with EOC using the variable importance in the projection (VIP) values. The top 50 metabolites among all metabolites ranked in descending order by VIP values were kept for further analysis. Based on the VIP value and the ascertainment of metabolite information in online databases, eight metabolites were selected as potential candidates. Subsequently, a random forest (RF) model which is a powerful classifier composed with trees was constructed from the eight candidates in the training set [16,17]. The logistic regression model adjusted for age and menopause was used to estimate the effects of metabolic biomarkers on classifying EOC patients and normal controls. In the validation step, the RF model was used to evaluate the classification performance of the biomarkers for the external validation set. The receiver operating characteristic (ROC) curve, the area of receiver operating characteristic curve (AUC) values, and the analyses of sensitivity and specificity were performed to describe the differences between EOC patients and normal controls.

PLS-DA procedures were performed using SIMCA-P (version 11.5; Umetrics, Sweden), while the ROC analysis and random forest analysis were conducted using R platform (version 2.12.1). The logistic regression was performed by SAS (version 9.1; SAS Institute, Inc., Cary, NC, USA).

Results

Metabolic profiling and biomarkers

The plasma samples from EOC patients and normal individuals were investigated by UPLC/QTOF/MS
(Figure 1A and B), and 11,473 peaks were obtained. The PLS-DA model score plots showed a clear separation trend between the EOC patients and the normal controls, with satisfactory modeling and predictive abilities ($R^2_{Y\text{cum}} 0.995$, $Q^2_{\text{cum}} 0.872$) (Figure 1C). All EOC patients were distinguished from the controls in the predictive component, including seven patients in their FIGO stage I or II. For the purpose of indentifying potential biomarkers, HMDB (http://www.hmdb.ca/), METLIN (http://metlin.scripps.edu/) and KEGG (http://www.genome.jp/kegg/ligand.html) databases were searched using the exact mass of the 50 kept metabolites. The distinction between the measured $M/Z$ values and the exact mass of the components of interest was within 0.01 Da. According to the databases, eight metabolites with only one matched $M/Z$ individually were selected as candidate biomarkers. These biomarkers were loaded far from the origin in the loading plot (Figure 1D). The eight metabolites (Table II) were significantly associated with EOC. They were up- and down-regulated in the EOC patients compared with the normal controls (Figure 2). The selected biomarkers were statistically significant after adjusted for age and menopause (Table III).

Validation set

The ROC analyses results from RF were a sensitivity of 92.1%, a specificity of 88.6%, and an AUC of 0.941 at the best cut-off points (Figure 3). And the six EOC patients who were in their early stages I or II were correctly identified.

---

**Figure 1.** Plots from UPLC/QTOF/MS and SIMCA-P. A and B showed the typical total ion current (TIC) chromatograms of plasma samples obtained from an EOC patient and a normal control, respectively, using UPLC/QTOF/MS; C shows the PLS-DA score plot for the first two components ($t[1]$, $t[2]$), which reveals that a considerable separation was achieved between EOC plasma samples (the red triangle) and normal controls (the black box). D depicts the PLS-DA scatter-plots results. The candidate biomarkers are marked with red square.
Discussion

In this study, we found that the plasma metabolic profiling by UPLC/QTOF/MS achieved a good diagnostic accuracy in distinguishing EOC patients from normal controls, and the combination of the eight selected biomarkers showed an excellent classification performance for the independent validation set. Our results were relatively reliable since: 1) the homogeneity was achieved because all participants were from the department of gynecology in the same hospital; and 2) the classification performance of the candidate biomarkers was evaluated in an external validation set.

In recent years, much attention has been paid to the application of metabolic profiling in diagnosing different cancers [12]. However, no metabolite individually provided high sensitivity and specificity. As shown in our study, a single biomarker had good diagnostic accuracy in distinguishing EOC from controls, and moreover, a combination of several metabolites was able to achieve a better diagnostic accuracy. The eight selected metabolites in this study were different from those of previous studies on this topic [13–15]. Although some metabolites with same M/Zs reported by them (e.g. M/Z 468.39, 466.33 and 520.43) were found in our top 50 ones, these metabolites were not selected as biomarkers temporarily in our study since we could not identify them due to many isomerides for them. In addition, the metabolite CPG (M/Z 403.35) was suggested as the main biomarker by Chen et al. [15], but it was not statistically significant in our study (Wilcoxon rank sum test: p = 0.44).

Age and menopausal status were imbalance between EOC patients and normal controls in this study, while neither of them influenced on the classification performance of the eight biomarkers analyzed in logistics regression model. The main limitation of our study is that the number of EOC patients with early stage (I/II) was small and only normal women were employed as controls but no women with other benign gynecological disease, such as benign ovarian tumors. Much larger cohorts consisting of relatively large number of EOC women with early stages and other than normal controls are needed in the investigation of metabolomics on this topic.

<table>
<thead>
<tr>
<th>Variable</th>
<th>M/Z</th>
<th>RT</th>
<th>AUC</th>
<th>Chemical formula</th>
<th>Common name</th>
<th>EOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>M 8484</td>
<td>519.3967</td>
<td>9.1751</td>
<td>0.700</td>
<td>C30H44O2</td>
<td>Demethylphylloquinone</td>
<td>↓</td>
</tr>
<tr>
<td>M 8573</td>
<td>668.3502</td>
<td>9.2346</td>
<td>0.837</td>
<td>C92H165N3O39</td>
<td>Ganglioside</td>
<td>↑</td>
</tr>
<tr>
<td>M 8574</td>
<td>682.3666</td>
<td>9.2359</td>
<td>0.777</td>
<td>C32H64NO7P</td>
<td>Lysosphospholipids</td>
<td>↑</td>
</tr>
<tr>
<td>M 9612</td>
<td>568.5645</td>
<td>9.7578</td>
<td>0.750</td>
<td>C36H73NO3</td>
<td>Ceramides</td>
<td>↑</td>
</tr>
<tr>
<td>M 9619</td>
<td>308.297</td>
<td>9.7606</td>
<td>0.817</td>
<td>C18H39NO3</td>
<td>Phytosphingosine</td>
<td>↑</td>
</tr>
<tr>
<td>M 10093</td>
<td>596.5956</td>
<td>9.9619</td>
<td>0.679</td>
<td>C38H77NO3</td>
<td>Ceramides</td>
<td>↑</td>
</tr>
<tr>
<td>M 10839</td>
<td>596.5944</td>
<td>10.2617</td>
<td>0.679</td>
<td>C38H77NO3</td>
<td>Ceramides</td>
<td>↑</td>
</tr>
<tr>
<td>M 11026</td>
<td>536.1654</td>
<td>10.3252</td>
<td>0.804</td>
<td>C11H12N2O4</td>
<td>N’-Formylkynurenine</td>
<td>↑</td>
</tr>
</tbody>
</table>

Figure 2. Box plots of eight typical differential metabolites from newly diagnosed EOC patients (the right box) and normal controls (the left box).
The most commonly used biomarkers, such as CA125, do not have satisfactory diagnostic accuracy in distinguishing EOC patients from normal women. The AUC value of CA125 was about 0.889 in Tcherkassova’s study [18], similar with this result, the AUC value of CA125 in discriminating EOCs from controls was 0.89 for our validation set (data not shown). Both of them were lower than the AUC of 0.941 acquired from our eight metabolic biomarkers. Meanwhile, we tried to identify chemical compounds associated with the eight biomarkers individually, and to explain why these biomarkers were useful in discriminating EOC from normal subjects. Based on MS spectra, retention behavior, database queries, and literatures, three candidate metabolic biomarkers (M/Z 568.5645, 596.5956, and 596.5944) were highly suspected to be ceramides (N-acylsphingosine). Ceramides can be synthesized from serine and palmitate in a de-novo pathway and are regarded as important cellular signals for inducing apoptosis [19]. A high intracellular concentration of ceramides will induce apoptosis in some cancer cells, such as prostate cancer [20], and breast cancer [21] cells. The candidate metabolic biomarker with M/Z 682.3666 was highly suspected to be a lysophospholipid (LyP). LyPs are bioactive lipids that elicit a wide range of cellular responses, such as cell survival, differentiation, and migration [22]. LyP has been shown to react by stimulating cell migration in prostate cancer [23]. Several studies have also shown that LyP is a potential biomarker for many kinds of cancers including EOC [24,25], which were consistent with our findings. The metabolite with M/Z 668.3502 was most likely to be ganglioside, which is composed of a glycosphingolipid (ceramide and oligosaccharide) together with one or more sialic acids. Ganglioside is a type of glycosphingolipid. The roles of ganglioside in neuroectodermal tumor have been reported [26,27], but no information is related to ovarian cancer. The metabolite with M/Z 308.297 was highly suspected as phytosphingosine, and its high regulated in EOC requires further investigation. The metabolites with M/Z 536.1654 and 519.3967 were highly suspected to be N’-Formylkynurenine and demethylphylloquinone, respectively, both of them were not reported connecting with EOC. The roles of these two metabolites in EOC cancer require further investigation.

In conclusion, the metabolic profiling could be used to separate EOCs from normal individuals. The combination of the eight candidate biomarkers provided excellent classification performance for the external cohort, and these biomarkers could be served as novel biomarkers in distinguishing between EOC patients and normal individuals in clinic.

Acknowledgements

The study was supported by a grant from the National Natural Science Foundation of China (NSFC 81172767). The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.
**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

**References**


