Five Serum Proteins Identified Using SELDI-TOF-MS as Potential Biomarkers of Gastric Cancer

Hai-bo Lu1,*, Jian-hua Zhou1, Yu-yan Ma2, Hai-ling Lu1, Ya-li Tang2, Qing-Yuan Zhang1 and Chang-hong Zhao1,*

1Department of Oncology, The Affiliated Tumor Hospital of Harbin Medical University and 2Tumor Institute of Heilongjiang Province, Harbin, China

*For reprints and all correspondence: Hai-bo Lu, Department of Oncology, The Affiliated Tumor Hospital of Harbin Medical University, Harbin 150040, China. E-mail: lu_haibo200@126.com. Chang-hong Zhao, Department of Oncology, The Affiliated Tumor Hospital of Harbin Medical University, Harbin 150040, China. E-mail: changhong110@yahoo.com.cn.

Received July 16, 2009; accepted November 22, 2009

Objective: The aims of this study were to detect serum proteomic patterns in gastric cancer serum samples using Surface-enhanced Laser Desorption/ionization-Time-of-flight-Mass Spectrometry ProteinChip array technology, to screen biomarker candidates, to build diagnostic models and to evaluate their clinical significance.

Methods: Serum samples from patients with gastric cancer and normal healthy control subjects (n = 125) were analysed using surface-enhanced laser desorption/ionization technology. The spectra were generated on weak cation exchange (WCX2) chips, and protein peak clustering and classification analyses were established using Ciphergen Biomarker Wizard and Biomarker Pattern software, respectively. The diagnostic models were developed and validated by discriminant analysis. In addition, the results of the surface-enhanced laser desorption/ionization model were compared with the biomarkers carcinoembryonic antigen and carbohydrate antigen 199 in a subset of samples using a microparticle enzyme immunoassay.

Results: Five protein peaks at 2046, 3179, 1817, 1725 and 1929 m/z were automatically chosen as components of the best biomarker pattern for diagnosis of gastric cancer. In addition, we identified a single protein peak at 4665 m/z, which could distinguish between stage I/II and stage III/IV gastric cancer with a specificity and sensitivity of 91.6% (11/12) and 95.4% (21/22), respectively. When this biomarker was validated in the second set of samples, the specificity and sensitivity were 91.7% (11/12) and 86.3% (19/22), respectively.

Conclusions: The present results suggest that serum surface-enhanced laser desorption/ionization protein profiling can distinguish patients with gastric cancer, and in particular stage I/II patients, from normal subjects with a relatively high sensitivity and specificity. Surface-enhanced Laser Desorption/ionization-Time-of-flight-Mass Spectrometry is a potential new diagnostic tool for the screening of gastric cancer.

Key words: biomarker – CA19-9 – CEA – gastric cancer – SELDI-TOF-MS

INTRODUCTION

Gastric cancer is currently the second most common cancer worldwide (1). Patients with gastric cancer generally have a poor prognosis, due largely to the lack of sufficient screening and early diagnostic tools for this disease. Currently in clinic the screening and early diagnosis of gastric cancer relies mainly on ultrasonography, endoscopic sonography, double-contrast radiologic method, computed tomography (CT), or magnetic resonance imaging and tumour markers including carcinoembryonic antigen (CEA), carbohydrate antigen 199 (CA19-9), microsatellite instability (2–5). All these methods, however, lack adequate sensitivity and/or specificity. Thus, it is urgent to search for better methods which provide more valuable information for screening and early diagnosis of gastric cancer.

Recent advances in protein expression analyses or proteomics have shown promise for the establishment of better
biomarkers for disease diagnosis, staging and therapeutic monitoring (6). Because of the marked heterogeneity of gastric cancer, a panel of biomarkers for screening and diagnosis would be most appropriate. Surface-enhanced Laser Desorption/ionization Time-of-flight Mass Spectrometry (SELDI-TOF-MS), an innovative proteomic technology (6,7), has overcome many of the limitations of two-dimensional electrophoresis and Matrix-assisted Laser Desorption/ionization Time-of-flight Mass Spectrometry (MALDI-TOF-MS) (8,9). SELDI-TOF-MS is a high-throughput technique for the analysis of complex biological specimens such as serum. It can detect multiple protein changes simultaneously with both high sensitivity and high specificity (10,11). Recently, SELDI was successfully used to distinguish samples between patients with pancreatic, ovarian and prostate cancer and healthy control subjects (9,12,13) and to detect biomarkers of bladder cancer in urine (14). SELDI-TOF-MS has also been used to successfully detect biomarkers of colorectal cancer in human serum and urine samples (15,16). The aim of the current study was to investigate the application of serum SELDI protein profiling to distinguish gastric cancer patients from a healthy population.

PATIENTS AND METHODS

SUBJECTS

A total of 125 serum samples including 65 pathologically confirmed gastric cancer patients and 60 healthy subjects were collected from the Affiliated Tumor Hospital of Harbin Medical University. The first phase of the study, the comparison phase, compared serum analyses between 34 patients with pathologically confirmed gastric cancer and 30 healthy subjects. The second phase, the validation phase, evaluated the ability of the algorithm established in the first phase to differentiate between serum samples from 31 patients with different stages of confirmed gastric cancer and 30 healthy subjects, and also compared the algorithm results with the results of analyses for two biomarkers in current use, CEA and CA19-9 (5). Informed consent was obtained from every subject prior to the study.

Patient demographics and characteristics of gastric cancer status were also recorded. All patients with gastric cancer were found to have no evidence of other disease. The distribution of clinical stages (17) was as follows: 5 patients had stage I disease, 7 had stage II, 10 had stage III and 12 had stage IV. Among these patients, 30 patients suffered from adenocarcinomas. The average age of the patients (14 males, 20 females) was 55.5 years (range 28–79 years). The healthy control subjects (18 males, 12 females) came from general physical examinations including negative gastroduodenoscopy and had an average age of 54.5 years (range 30–72 years). The two groups were matched for age and sex. Two millilitres of whole blood were collected in the early morning following overnight fasting and stored within 1 h at 4°C. The blood was later centrifuged for 20 min at 4000 rpm, distributed into 100 µl aliquots, and stored at −80°C for later analyses.

MATERIALS AND EQUIPMENT

PBS-II SELDI-TOF-MS, energy-absorbing elements sinapinic acid (SPA), CM-10 protein chips (weak cation + hydrophobic membrane) and the corresponding ProteinChip Software 3.2.1 were from Ciphergen (Fremont, CA, USA). HEPES buffer salt and CHAPS buffer salt were from Sigma (St Louis, MO, USA). All other reagents were of analytical grade and were obtained commercially.

CM-10 CHIP PREPARATION

Protein chips were washed three times in high-performance liquid chromatography (HPLC) water with shaking for 5 min at 400 rpm. Five microlitres of 10 mM HCl was added to samples in covered test tubes and the samples were shaken for 5 min. The chips were rinsed three times with deionized (DI) water, and then the chips were loaded into the test tubes containing HPLC water and were shaken for 5 min. The chip shelf and pad were cleaned using ultrasound, rinsed three times with DI water, and then dried. The chips were loaded into the processor, 200 µl 0.1 M NaAC was added into each orifice to bind the buffer, and the samples were oscillated at 250 rpm at room temperature for 5 min. This procedure was repeated once and the samples were added after the chips were dried.

SURFACE-ENHANCED LASER DESORPTION/IONIZATION PROTEIN PROFILING

The ProteinChip Array Cassette was placed in the Bioprocessor, 150–250 µl of binding solution was added to each well, and the cassette was incubated at room temperature for 5 min with vigorous shaking (250 rpm). After incubation, this procedure was repeated once. After the second incubation, the buffer was removed and 50–150 µl of each sample was added to individual wells. The final protein concentration was 50–2000 µg/ml total protein, diluted in binding buffer. Samples were incubated at room temperature with vigorous shaking for 30 min.

Following incubation, sample solutions were removed and the wells were washed two times with 150–250 µl binding buffer for 5 min, with agitation. After washing, the binding buffer was removed and wells were washed twice with 150–250 µl DI water. The arrays were air-dried for 15–20 min, and then 1 µl of EAM solution (consisting of 75 µl acetonitrile in SPA added at 10 ml/l to trifluoroacetic acid (18) was added to each well, holding down the frame provided with the Cassette Compatible Bioprocessor in order to keep the cassette flat during EAM addition. The cassette was air-dried for 5 min, and then an additional 1 µl of EAM
was added to each well. After the wells were dry, the array was analysed using a ProteinChip Reader.

DATA ANALYSIS

The chips were placed in the Protein Biological system II-C mass spectrometer reader (Ciphergen Biosystems, Inc.) and TOF spectra were generated by 190 laser shots intensity and a detector sensitivity of 9. The optimization range was from 1500 to 20 000 m/z, and a maximum of 200 000 Da. External calibration of the instrument was performed using the all-in-one peptide molecular mass standard (Ciphergen Biosystems, Inc.). We achieved a mass accuracy of 0.1% with this system. Peak detection used Ciphergen Proteinchip Biomarker Wizard Software 3.2.1.

DECISION TREE CLASSIFICATION

Construction of the decision tree classification algorithm was performed by Ciphergen Biomarker Pattern software version 5.0. Classification tree, selected Gini, split the data into two nodes using one rule at a time in the form of peak intensity. The splitting decisions in this case were based on the normalized intensity levels of peaks from SELDI protein expression profile. The process of splitting was continued until terminal nodes were created. After V-fold cross validation 50, the accuracy of each classification tree was then challenged with the samples.

DETECTION OF SERUM CARCINOEMBRYONIC ANTIGEN AND CARBOHYDRATE ANTIGEN 199

The two markers, CEA and CA19-9, were measured in the 64 sera included in this study using a microparticle enzyme immunoassay (MEIA; Abbott, USA). The cut-off values for CEA and CA19-9, recommended by the manufacturers, were 5.0 ng/ml and 37 µg/ml, respectively.

STATISTICAL ANALYSIS

Serum samples from a total of 125 patients, 64 in the comparison phase and 61 in the validation phase, were examined through SELDI-TOF-MS applied decision tree algorithm. In the comparison phase, the algorithm was used to detect the gastric cancer diagnostic accuracy in patients with gastric cancer and healthy volunteers. In the validation phase, it was used to determine the stage of cancer among patients with gastric cancer. In addition, the diagnostic rule from the SELDI-TOF-MS applied decision tree algorithm in the comparison group was applied to the validation patient sample for comparison with the currently used biomarkers CEA and CA19-9.

For other analyses, continuous variables are presented as the mean ± standard deviation and categorical variables are presented as the number of individual and percentage of the population. Two-sample t-tests and Pearson’s χ² tests were performed for comparison of differences between two groups. Comparison of relative peak intensity levels between groups was made using the Student’s t-test. A McNemar’s test was also performed for comparison of the diagnostics of dependent subjects through SELDI-TOF-MS, CEA and CA19-9 diagnostic methods. All statistical analyses were performed with a predetermined significance level of α = 0.05. All statistical analyses were conducted using the statistical software SPSS 15.0 (SPSS Inc, Chicago, IL, USA).

RESULTS

A total of 125 patients, 64 for the comparison phase and 61 for the validation phase, were enrolled in the study. The demographics for all enrolled patients, including normal and gastric cancer patients are summarized in Table 1. The two study phases were conducted in order to first establish the accuracy of gastric cancer diagnosis, and second to validate the model that was created. None of the patient demographics was significantly different between the comparison and the validation groups (Table 1).

SELDI-TOF-MS applied tree algorithm classification (Figure 1) was used to evaluate gastric cancer diagnostic accuracy by classification of serum samples from patients with gastric cancer versus normal controls. The left branch node after the first layer depicts the cases of peak intensity under 1.475, while the right branch node depicts cases equal to or greater than 1.475. The cut-off points for peaks 3179, 1817, 1725 and 1929 m/z were 2.239, −0.856, 1.321 and 1.798, respectively.

Figure 2 presents classification of gastric cancer stage using a SELDI-TOF-MS applied decision tree algorithm. The left branch node after the first level depicts cases of peak intensity under the cut-off value (0.022), while the right branch node depicts cut-off values equal to or greater than the cut-off value. The other cut-off points were 0.348 and 1.509. Nodes 1 and 3 were classified as stage B (III/IV), while nodes 2 and 4 were classified as stage A (I/II). Thus,

<table>
<thead>
<tr>
<th>Table 1. Patient demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Age±</td>
</tr>
<tr>
<td>Gender b</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Gastric cancer b</td>
</tr>
</tbody>
</table>

Data are expressed as mean ± standard deviation for continuous variables and as n (%) for categorical variables.

bP values were established for continuous variables using a two-sample t-test.

bP values were established for categorical variables using Pearson’s χ² test.
values between 0.022 and 0.348 were classified as node 2, stage A; values between 0.348 and 1.509 were classified as node 3, stage B; and values ≥ 1.509 were grouped as node 4, stage A. Representative spectra for the 2046 and 4665 m/z SELDI peaks are presented in Figures 3 and 4, respectively.

The sensitivity and specificity for diagnostic accuracy were defined as the patients with clinically diagnosed gastric cancer which was detected using the SELDI-TOF-MS algorithm, and the normal control patients who were categorized as normal by the same methodology. In the comparison phase of the study, the sensitivity and specificity for the algorithm were 94.1% and 93.3%, respectively (Table 2). The decision tree algorithm applied in SELDI-TOF-MS was also tested for accurate classification of gastric cancer stage in the validation stage (Figure 2). The results showed that the sensitivity and specificity in the validation group were 90.3% and 80.0%, respectively.

Independent of whether patients were part of the comparison or validation group, the results from SELDI-TOF-MS were not significantly different from the individual clinical diagnosis as determined using McNemar’s tests (P = 1.000 for the comparison group and P = 0.508 for the validation group). Based on these results, the SELDI-TOF-MS applied tree algorithm might provide confident diagnostic accuracy for the determination of the presence of gastric cancer. However, differential sensitivity for detecting gastric cancer stage using SELD-TOF-MS was observed in the two study phases, with a higher sensitivity performance in the comparison group (91.7% and 95.4% for early stage (I/II) and late stage (III/IV), respectively) than in the validation group (92.9% (13/14) and 88.2% (15/17) for the early stage and late stage, respectively). Although the sensitivity results in the validation stage were not as high as that in the comparison stage, the results of the SELDI-TOF-MS decision rule were still not significantly different patient clinical diagnoses for either stage.

Figure 1. Classification of serum samples from patients with gastric cancer versus normal controls using a decision tree algorithm. The left branch node after the first layer depicts cases of peak intensity under 1.475, while the right branch node depicts cases equal to or greater than 1.475. The cut-off points for 3179, 1817, 1725 and 1929 m/z were 2.239, −0.856, 1.321 and 1.798, respectively.

Figure 2. Classification of gastric cancer stage using a decision tree algorithm. The left branch node after the first level depicts cases of peak intensity under the cut-off value (0.022), while the right branch node depicts cut-off values equal to or greater than the cut-off value. The other cut-off points were 0.348, 1.509. Nodes 1 and 3 were classified as stage B (III/IV), and nodes 2 and 4 were classified as stage A (I/II). Values ≤ 0.022 were classified as node 1, stage B; values >0.022 and ≤ 0.348 were classified as node 2, stage A; values >0.348 and ≤ 1.509 were classified as node 3, stage B; and values >1.509 were grouped as node 4, stage A.
group ($P = 1.000$ for the comparison group and $P = 0.804$ for the validation group).

We also wished to compare the accuracy of this new diagnostic methodology with that of currently used biomarkers. Table 3 summarizes the sensitivity and specificity for the SELDI-TOF-MS applied tree algorithm in the validation group with those for CEA and CA19-9 in the same population. Results indicate that the highest sensitivity and specificity among these three diagnostic methodologies were observed for the SELDI-TOF-MS applied in tree algorithm. Thus this new method might improve patient gastric cancer diagnostic accuracy. However, despite large numerical differences in sensitivity and specificity between the methods, there were no significant differences as determined using McNemar’s tests ($P = 0.557$ as compared with CEA and $P = 0.473$ as compared with CA19-9).

### DISCUSSION

There are currently no satisfactory screening and early diagnostic strategies available for gastric cancer. SELDI is a high-throughput technique used to generate protein expression profiles which, in combination with bioinformatics tools to extract information for biomarker discovery, has been essential for the identification of novel protein biomarkers of disease (6). Indeed, application of this technology...

---

**Figure 3.** Differential expression of 2046 $m/z$ SELDI peaks in serum samples from patients with gastric cancer versus healthy controls. C70: gastric cancer. N29: healthy control. Top: representative spectrometric image; bottom: representative electrophoresis image.

**Figure 4.** Differential expression of 4665 $m/z$ SELDI peaks in serum samples from patients with stage I/II versus stage III/IV gastric cancer. C73a: I stage gastric cancer. C7-a: II stage gastric cancer. C70-b: III stage gastric cancer. C5: IV stage gastric cancer. Top: representative spectrometric image; bottom: representative electrophoresis image.

**Table 2.** The sensitivity of SELDI-TOF-MS applied tree algorithm classification for accurate diagnosis and staging of gastric cancer in the comparison and validation patient populations

<table>
<thead>
<tr>
<th></th>
<th>Comparison group ($n = 64$)</th>
<th>Validation group ($n = 61$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct classification</td>
<td>Incorrect classification</td>
</tr>
<tr>
<td>Normal</td>
<td>93.3% (28/30)</td>
<td>6.7% (2/30)</td>
</tr>
<tr>
<td>Gastric cancer</td>
<td>94.1% (32/34)</td>
<td>5.9% (2/34)</td>
</tr>
<tr>
<td>Stage I/II</td>
<td>91.7% (11/12)</td>
<td>8.3% (1/12)</td>
</tr>
<tr>
<td>Stage III/IV</td>
<td>95.4% (21/22)</td>
<td>4.6% (1/22)</td>
</tr>
</tbody>
</table>

Classification was established using the SELDI-TOF-MS applied tree algorithm. Data are expressed as the percentage ($n$) for each group.

$^a$For gastric cancer diagnostic accuracy, sensitivity was not significantly different between SELDI-TOF-MS results and clinical diagnostics ($P = 1.000$ and $P = 0.508$ for comparison and validation groups, respectively).

$^b$For diagnostic accuracy in determining gastric cancer stage, sensitivity was not significantly different between SELDI-TOF-MS results and clinical diagnostics ($P = 1.000$ and $P = 0.804$ for comparison and validation groups, respectively).
has shown great potential for the early detection of ovarian, prostate and colorectal cancers (10,12,15,16).

There are an increasing number of reports on the application of proteomic techniques to the study of gastric cancer (19–24), few investigators have applied SELDI methodologies to the development of novel biomarkers for gastric cancer diagnosis and classification (22,25–28). In the present study, we examined 34 serum samples from gastric cancer patients and 30 from healthy individuals using the SELDI technique together with the WCX2 protein chip. A classification tree was constructed for use in distinguishing cases of gastric cancer cases from healthy individuals which incorporated five protein peaks at 2046, 3179, 1817, 1725 and 1929 m/z as a marker pattern. When the model was validated in a blinded set of samples, it yielded a sensitivity of 80.0%, specificity of 73.5%. These results are comparable with the sensitivities and specificities reported by other investigators using similar SELDI methodology, but different protein peaks in their models (25,27,28).

Our results using SELDI biomarker analysis were also compared with the current standard gastric cancer diagnostic biomarkers CEA and CA19-9 as measured using MEIA. Although there are no statistical differences between the specificities of CEA, CA19-9 and the SELDI marker pattern, the sensitivity achieved by CEA, CA19-9 individually or in combination was significantly lower than that of the SELDI pattern. These findings are in agreement with two previous reports also using SELDI methodologies (25,27). These results indicate that the SELDI pattern is distinctly superior to CEA and CA19-9, either individually or in combination, in distinguishing patients with gastric cancer from healthy individuals. The application of the SELDI technique in the detection of the protein markers of gastric cancer may overcome the shortcomings of CEA and CA19-9, whose sensitivity and specificity are poor.

Using our technique, an additional protein peak, 4665 m/z, was identified which could distinguish between I/II stage and III/IV stage gastric cancers. The sensitivity of the SELDI marker pattern for I/II stage was significantly higher than III/IV stage, indicating that the pattern may be more effective in discriminating I/II stage cancers than III/IV stage cancers. These findings suggest that this pattern might be more effective at early detection of gastric cancer than any other single or panel of biomarkers currently in clinical use.

To develop a broad biomarker panel for screening a diverse, high-risk population, the results of the current investigation require future studies incorporating a larger sample size in order to broaden and improve the diagnostic value of this new diagnostic methodology. Furthermore, the five peaks and the single 4665 m/z included in the SELDI marker pattern should be investigated using MALDI-MS-MS.

In conclusion, we have found that serum SELDI protein profiling can distinguish gastric cancer patients, especially I/II stage patients, from healthy controls with relatively high sensitivity and specificity. SELDI-TOF-MS is a potential new tool for the screening and identification of patients with gastric cancer.

Funding

This work was supported by grants from the HeiLongJiang Province Nature Science (D0349) and Harbin Science and Technology Plan Program (2005AA9cs116-14).

Conflict of interest statement

None declared.

References


Table 3. The sensitivity and specificity of the classification results by SELDI-TOF-MS applied tree algorithm, CEA and CA19-9 in the validation patient population ($n = 61$)

<table>
<thead>
<tr>
<th>SELDI-TOF-MS algorithm</th>
<th>CEA</th>
<th>CA19-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>90.3% (28/31)</td>
<td>48.4% (15/31)</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>80.0% (24/30)</td>
<td>50.0% (15/30)</td>
</tr>
<tr>
<td>$P$ valuea</td>
<td>–</td>
<td>0.557</td>
</tr>
</tbody>
</table>

*aThe sensitivity and specificity of SELDI-TOF-MS was not significantly different from that of CEA or CA19-9 ($P = 0.557$ and $P = 0.473$, respectively). Data are expressed as percentage ($n$).


