Original Article

Serum biomarkers for lymph node metastasis in patients with triple-negative breast cancer by proteomics

Ai-Na Liu, Hua-Jun Qu, Cai-Yan Yu, Jian-Nan Liu, Ai-Hong Jiao, Ping Sun

Department of Oncology, Yantai Yuhuangding Hospital, Yantai, China

Received November 23, 2016; Accepted December 30, 2016; Epub March 1, 2017; Published March 15, 2017

Abstract: The present study was aimed to investigate the differences of serum proteins in triple-negative breast cancer (TNBC) patients with and without lymphatic metastasis for identification of serum biomarkers for TNBC lymphatic metastasis. Thirty pairs of serum samples from TNBC patients with and without lymphatic metastasis were analyzed by two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Differentially expressed proteins were screened using PDQuest™ software and identified by using Mascot software search engine. Two-dimensional gel electrophoresis profiles were obtained. A total of 14 differentially expressed protein spots were found in the sera of two study groups. Among them, 10 were identified by mass spectrometry. By combing database searching and literature review, it is speculated that the upregulated protein transthyretin and clusterin may be the candidate biomarkers for early diagnosis of lymphatic metastasis of TNBC patients. Thus, the 14 differentially expressed proteins, especially transthyretin and clusterin, might provide useful information to early serologic diagnosis and differential diagnosis of lymphatic metastasis of triple-negative breast cancer.

Keywords: Breast cancer, lymphatic metastasis, proteomics, mass spectrometry, prognosis

Introduction

Triple-negative breast cancer (TNBC) refers to any breast cancer that does not express the genes for estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor-2 (Her-2). Triple negative breast cancer (TNBC) is a recently identified group of breast cancers with dismal prognosis, which represent about 15% breast cancer cases [1]. Due to its high invasiveness, poor prognosis, early relapse and metastasis, most patients present with lymphatic or homogenous metastasis at diagnosis [2].

The occurrence of lymphatic metastasis in TNBC is a critical indicator to clinical staging, prognostic prediction and postoperative adjuvant therapy. It has been reported that the survival rate had significant differences between patients with a small tumor (< 2 cm) without lymphatic metastasis and patients with large tumor (> 5 cm) with axillary lymph node metastasis [3]. Thus, detection of metastasis factors at the early stage can improve the local disease control rate and long-term survival of TNBC patients. The clinical prevalent diagnostics utilized to identify lymph node metastasis include palpation diagnoses, lymph node biopsy, imaging examination and sentinel node biopsy. Sentinel lymph node biopsy is a currently available predictor to axillary lymph node metastasis; however, its utility is compromised by resultant complications [4] and additional injury to the patient. The efficiency and accuracy of lymphatic metastasis diagnosis depends on metastasis size and location and the surgical skills of medical staff. Therefore, these diagnostic methods are not suitable for all patients, especially for neoadjuvant chemotherapy patients [5].

Up to now, proteomic strategies have been used to identify candidate serum biomarkers associated with metastatic progression in various of human cancers [6-9]. It is apparent that
2D-DIGE (two-dimensional fluorescence difference gel electrophoresis) in combination with MALDI-TOF-MS (matrix-assisted laser desorption ionization-time-of-flight mass spectrometry) is a basic technology to separate and identify proteins. Some studies have confirmed differential protein expression in breast cancers with and without lymphatic metastasis and proteins that altered between primary carcinomas and matched metastatic lymph nodes using proteomics technology and approaches [10-12]. However, there has been little research or report on serum protein of TNBC with lymphatic metastasis used as biological markers by employing mass spectrometry techniques. Therefore, we resorted to proteomic technology to identify the serum-specific biomarkers closely associated with lymphatic metastasis in TNBC, and provide further molecular insights of its aggressive biology to provide evidence for early diagnosis and treatment of TNBC lymphatic metastasis.

Materials and methods

Specimens

Sixty patients received modified radical mastectomy during January 2012 and January 2013 in Yantai Yuhuangding Hospital, whose integrate data were well kept were enrolled in the study. Among them, 30 were TNBC with lymphatic metastasis (LM+TNBC) and the other 30 were TNBC without lymphatic metastasis (LM-TNBC). Patients were all females with a median age of 51 years (age range: 32~66 years). All of them had received chemotherapy and radiotherapy. The study was approved by the Research Ethics Boards.

A 3 ml of fasting blood was drawn from each subject in the morning, placed at room temperature for 2 hours, and centrifuged at 4000 rpm/min for 10 min when the serum began to separate. Serum was collected and kept at -80°C in a refrigerator.

Extraction of serum proteins and removal of highly abundant proteins

Serum samples of the two groups were mixed in equal amount for processing by Albumin and IgG Removal kit in the following procedures: 15 μl of serum sample was mixed with 750 μl of affinity gel, rotated at 250 r/min for 30 min at room temperature until homogeneity was achieved. The gel/sample mixture was pipetted into a centrifuge tube equipped with a filter, centrifuged at approximately 6500 g for 5 minutes. Then the filtrate, about 500 μl was collected from each sample.

Sample lysis and Cydye labeling

100 μl of serum free from highly-abundant proteins and IgG was pipetted and added pre-cooled acetone at a ratio of 1:4, mixed gently, and precipitated at -20°C for 2 hours. The precipitated proteins were added 300 μl of lysis buffer (30 mM Tris, 7M Urea, 2 M Thiourea and 4% CHAPS) for protein extraction, then the protein solution was adjusted to a pH about 8.5. Cydye (Amersham, GE Healthcare) stock solution and working solutions were prepared as follows: stock solutions were made by 2 nmol of Cydye reagent in 2 μl of DMF; working solutions were made by mixing stock solution and DMF at a ratio of 2:3 before labeling. Imagequant reagent kit was used to quantify protein solution. An internal pool standard of was prepared by mixing same amount of all samples, and 50 μg of the resulting mixture was labeled with 1 μl of Cy2 working solution. 50 μg of serum protein sample from LM+TNBC group was labeled with Cy3 and 50 μg of serum protein sample from LM-TNBC group was labeled with Cy5 at -4°C. After thirty minutes of labelling, 1 μl of 10 nM lysine solution was added to stop the reaction.

Two-dimensional gel electrophoresis

Samples labeled with Cy2, Cy3 and Cy5 were mixed and added rehydration buffer. Strips (24 cm, pH 3-10) rehydrated with such solutions were performed isoelectric focusing (IEF) and SDS- polyacrylamide gel electrophoresis (SDS-PAGE) in traditional 2-D electrophoresis protocol, in the dark with low fluorescence glass plates in the second dimension.

Gel image scanning and analysis

After electrophoresis, samples were performed laser confocal scanning by Typhoon trio scanner at the excitation/emission of 488/520, 532/580 and 633/670 nm respectively. Differentially expressed protein spots were analyzed by Bio-Rad’s PDQuest™ Version 7.0.
Serum biomarkers for lymph node metastasis in TNBC

Differential spots that are more than twice up or down regulated were marked. DIGE gels were post-stained with Coomassie blue R-350. Different spots were sectioned for mass spectrometry.

**Mass spectrometry**

The procedure includes: 1) the gel piece was excised by a flat excision tip, placed in an Eppendorf tube and rinsed 3 times with water. 2) The gel was destained by 50% ACN/25 mM NH4HCO3 at 37°C for 30 min until decoloration is fully achieved. 3) The destined gel piece was immersed in 100% ACN, shrank to small pale piece and dried by vacuum centrifuge concentration. 4) 10 μl of sequencing grade modified trypsin (10 μl/ml dissolved in 25 mM NH4HCO3) was added to saturate the gel. After 30 mins, 50 mM NH4HCO3 was added to submerge the gel and kept at 37°C overnight. 5) The reaction solution was pipetted, added extract solution (50% ACN/5% TFA) and shaken by microwave shaker for 30 min. 6) The reaction plus extract solution was dried by vacuum centrifuge concentration and dissolved in 3 μl of 1% TFA/50% ACN. 7) 10 mg CHCA matrix was dissolved in 500 μl ACN and 500 μl 0.1% TFA. 8) 0.3 μl sample was mixed with 0.3 μl CHCA and spotted in the wells of the plate. 9) ABI 4700 MALDI-TOF mass spectrometer was used to identify differentially expressed protein spots. Peptide mass fingerprinting (PMF) was operated in reflection mode with a laser intensity of 4500 Hz, an M/Z range between 800-3500, and calibrated with the following reference standards: Angiotensin I, Glu1-Fibrinopeptide B, ACTH (clip 1-17) and ACTH (clip 18-39). 10) The peak list of PMF, from which the trypsin peak and keratin peak were screened, was performed protein identification with Mascot software.

**Identification of protein spots and bioinformatics analysis**

The details in methods for identification of protein spots and bioinformatics analysis were described previously [19]. The experimental MS data were matched to a corresponding virtual peptide mass database derived from GPS Explorer™ v3.6, Mascot, the XXX protein database (54,990 sequences; 39,965,472 residues), and NCBI databases (http://www.ncbi.nlm.nih.gov/BLAST; 12,558,787 sequences; 4,287,872,731 residues). Peak list generation was carried out according to default parameters using Flex Analysis Software (Bruker Daltonics, Germany). Protein identification was carried out by MALDI-TOF MS using the Mascot software (http://www.matrixscience.com). Identification required a Mascot confidence interval of 95%. The criteria used to accept protein identifications were based on all MS and MS/MS data, including the extent of sequence coverage, number of peptides matched, and probability score. Protein scores greater than 60 are considered significant (P < 0.05). The function, gene name, and gene ontology (GO) category of each protein were determined using the Mascot v2.1 software protein database search engine. GO analysis for identified proteins based on BLAST results were performed using the UniProt Knowledgebase (Swiss-Prot/TrEMBL) and the GO database. We also analyzed identified proteins within the Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathway maps (www.genome.jp/kegg/).

**Results**

**Clinical characteristics of triple-negative breast cancer patients**

Sixty patients received modified radical mastectomy were enrolled in the study. Among them, 30 were TNBC with lymphatic metastasis (LM+TNBC) and the other 30 were TNBC without lymphatic metastasis (LM-TNBC). Clinical characteristics of triple-negative breast cancer patients were shown in Table 1.
Differential protein spots on 2D-gel electrophoresis and image analysis

Fifty differential spots which were more than 1.5 times up or down regulated were found by the software in the two processed groups, among which 4 were included in Cy3 but disappeared in Cy5. In differential protein expression analysis, Gaussian fit, X streaking, Y streaking, overlap, and linear range of scanner were taken into account by the software. The most significantly 32 differential spots were performed mass spectrometry for identification (Figures 1 and 2). The identified proteins were listed in Table 2. The differential protein results with significant differences from the samples of the two groups identified by MS were listed in Table 3. Serum proteins in LM+TNBC are different from those in LM-TNBC, showing 14 differentially expressed proteins by mass spectrometry analysis. Among them, two proteins were dramatically changed in their expression between the two groups by further analysis and database retrieval, including transthyretin (TTR) and Clusterin (CLU).

Go annotation

GO was performed to understand the biological functions of all the differentially abundant spots. The annotated XXX proteins identified using the NCBI and XXX databases were analyzed using the UniProt Knowledgebase (Swiss-Prot/TrEMBL) and GO database. Each identified protein was classified according to its GO functional annotation. These differentially abundant proteins were mainly involved in biological regulation, metabolic process, response to stress, binding, antioxidant activity, and transporter activity. According to the KEGG metabolic pathway maps of XXX (www.genome.jp/kegg/), the identified proteins including transthyretin (TTR) and Clusterin (CLU) were involved in the glycolytic/gluconeogenesis pathway.

Figure 1. Representative electrophoretogram of serum protein sample from LM+TNBC group.
Discussion

2-D GE and MALDI-TOF-MS have merits of simple manipulation, high resolution, fast speed and good repeatability. This technique has been applied to most tumorous tissue and cell lines, but is rarely reported in serious promotes. Serum samples are ideal for clinical study and application because they are convenient and easy to obtain, large in quantity, and cost-effective. In this study, 14 differential protein spots were found by 2-D GE and MALDI-TOF-MS in the sera of LM+TNBC and LM-TNBC patients. Among them, 13 were identified as differentially expressed proteins; hemoglobin subunit beta manifested itself differentially at two protein spots. Further analysis of the 13 differentially expressed proteins, combined with database search, revealed that TTR and CLU were of significant differential expression between the two groups. They are considered as closely associated with tumorigenesis by literature.

Transthyretin (TTR) is also known as prealbumin which is synthesized primarily in the liver and the retina. As the major serum carrier, TTR binds and transports the thyroid hormones and retinol to play specific biological functions. Differential serum levels of TTR have been linked to several cancers, including ovarian, lung, endometrial and pancreatic cancers [13-17]. TTR in breast cancer is not much reported in domestic or foreign literature. Liang [18] found that TTR was decreased in the breast cancer tissue with 91.7% diagnostic accuracy in the proteomic analysis of 18 cases of breast cancer and normal breast tissue. Previously, our data demonstrated that the expression of TTR in the sera of TNBC patients is obviously decreased compared with that of non-TNBC patients [19]. In contrast with these studies, Chung et al. [20] identified the full-length TTR was also significantly upregulated in the serum of breast cancer patients detected by protein chip. Our present study revealed an up-regulat-
Table 2. Differential protein spots identified by comparing the serum protein expression between LM+TNBC and LM-TNBC samples

<table>
<thead>
<tr>
<th>No.</th>
<th>Spot</th>
<th>Differential protein</th>
<th>Gene</th>
<th>SwissPro Accession No.</th>
<th>Molecular weight (Da)</th>
<th>pI</th>
<th>Up/down regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1101</td>
<td>Clusterin</td>
<td>CLU</td>
<td>P10909</td>
<td>52495</td>
<td>5.88</td>
<td>Up</td>
</tr>
<tr>
<td>2</td>
<td>7603</td>
<td>Serotransferrin</td>
<td>TF</td>
<td>P02787</td>
<td>77064</td>
<td>6.81</td>
<td>Up</td>
</tr>
<tr>
<td>3</td>
<td>1402</td>
<td>Antithrombin-III</td>
<td>SERPINC1</td>
<td>P01008</td>
<td>52602</td>
<td>6.32</td>
<td>Down</td>
</tr>
<tr>
<td>4</td>
<td>5407</td>
<td>Complement C3</td>
<td>C3</td>
<td>P01024</td>
<td>187148</td>
<td>6.02</td>
<td>Down</td>
</tr>
<tr>
<td>5</td>
<td>7303</td>
<td>Complement component C6</td>
<td>C6</td>
<td>P13671</td>
<td>104786</td>
<td>6.39</td>
<td>Up</td>
</tr>
<tr>
<td>6</td>
<td>1402</td>
<td>Vitamin D-binding protein</td>
<td>GC</td>
<td>P02774</td>
<td>52964</td>
<td>5.40</td>
<td>Down</td>
</tr>
<tr>
<td>7</td>
<td>4102</td>
<td>Apolipoprotein A-I</td>
<td>AP01</td>
<td>P02647</td>
<td>30778</td>
<td>5.56</td>
<td>Down</td>
</tr>
<tr>
<td>8</td>
<td>1404</td>
<td>Alpha-1-antitrypsin</td>
<td>SERPINA1</td>
<td>P01009</td>
<td>46737</td>
<td>5.37</td>
<td>Down</td>
</tr>
<tr>
<td>9</td>
<td>3603</td>
<td>Complement factor B</td>
<td>CFB</td>
<td>P00751</td>
<td>85533</td>
<td>6.67</td>
<td>Up</td>
</tr>
<tr>
<td>10</td>
<td>3507</td>
<td>Alpha-2-macroglobulin</td>
<td>A2M</td>
<td>P01023</td>
<td>163291</td>
<td>6.03</td>
<td>Down</td>
</tr>
<tr>
<td>11</td>
<td>3304</td>
<td>Prothrombin</td>
<td>F2</td>
<td>P00734</td>
<td>70036</td>
<td>5.63</td>
<td>Down</td>
</tr>
<tr>
<td>12</td>
<td>4001</td>
<td>Transthyretin</td>
<td>TTR</td>
<td>P02766</td>
<td>15887</td>
<td>5.49</td>
<td>Up</td>
</tr>
<tr>
<td>13</td>
<td>4005</td>
<td>Hemoglobin subunit beta</td>
<td>HBB</td>
<td>P68871</td>
<td>15998</td>
<td>6.74</td>
<td>Down</td>
</tr>
<tr>
<td>14</td>
<td>7002</td>
<td>Hemoglobin subunit beta</td>
<td>HBB</td>
<td>P68871</td>
<td>15998</td>
<td>6.74</td>
<td>Up</td>
</tr>
</tbody>
</table>

Table 3. Local results of differential proteins CLU and TTR from the two groups

<table>
<thead>
<tr>
<th>No.</th>
<th>Spots</th>
<th>Protein name</th>
<th>Gene name</th>
<th>SwissPro Accession Number</th>
<th>MW</th>
<th>PI</th>
<th>Cy3/Cy5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1101</td>
<td>Clusterin</td>
<td>CLU</td>
<td>P10909</td>
<td>52495</td>
<td>5.88</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>4001</td>
<td>Transthyretin</td>
<td>TTR</td>
<td>P02766</td>
<td>15887</td>
<td>5.49</td>
<td></td>
</tr>
</tbody>
</table>

The expression of TTR in the serum of LM+TNBC group compared with in LM-TNBC group, suggesting that TTR might be involved in the lymph node metastasis of TNBC.

Clusterin (CLU) is also known as apolipoprotein J or testosterone repressed prostate message-2 which is a multifunctional glycoprotein and widely distributed in different tissues. The role of clusterin in cancer has been extensively studied and it participates in different pathological stages of cancer progression by complex mechanisms [21-23] Doustjalali et al. [24] reported that CLU was higher in the serum of breast cancer than that in benign breast disease. Redondo et al. [25] found high expression of CLU were associated with large tumor size and negative hormone receptors. Krüger et al. [26] analyzed 141 breast cancer specimens and found that CLU expression was related to prognostic indicators such as high histological tumor grade and high Ki-67 labeling index. Correlation between CLU and TNBC was rarely reported. Such a correlation was proposed by a few retrospective analyses [27, 28]. Zhang [29] reported positive rate of the secretory CLU was 75.4% in 61 TNBC cell specimens, which was much higher than that in non-TNBC or normal breast tissue. Our study revealed that CLU expression was significantly higher in the LM+TNBC group than in LM-TNBC group, indi-
Serum biomarkers for lymph node metastasis in TNBC

In conclusion, the present study screened some proteins associated with lymphatic metastasis in TNBC by proteomics technology. The functions and mechanism of these proteins merits further investigations. We hope that a database of proteins associated with LM+TNBC will be established to facilitate early diagnosis and treatment of breast cancer in the near future.

Acknowledgements

This study was supported by Scientific and Technological Development Program of Yantai Municipality (No. 2012086).

Disclosure of conflict of interest

None.

Address correspondence to: Ping Sun, Department of Oncology, Yantai Yuhuangding Hospital, Yantai 264000, China. Tel: (+86) 13906388003. +86-535-6691999-81903; E-mail: spzhpc@sina.com

References


