PO Box 2345, Beijing 100023, China
www.wjgnet.com
wjg@wjgnet.com

ELSEVIER

*World J Gastroenterol* 2005;11(40):6249-6253
***World Journal of Gastroenterology*** ISSN 1007-9327




# Preliminary study on proteomics of gastric carcinoma and its clinical significance

Hong-Gang Qian, Jing-Shen, Hong Ma, Hua-Chong Ma, Ya-Hui Su, Chun-Yi Hao, Bao-Cai Xing, Xin-Fu Huang, Cheng-Chao Shou

Hong-Gang Qian, Jing-Shen, Hong Ma, Ya-Hui Su, Chun-Yi Hao, Bao-Cai Xing, Xin-Fu Huang, Cheng-Chao Shou, Peking University School of Clinical Oncology and Beijing Cancer Hospital, Beijing 100036, China
Hua-Chong Ma, Beijing Chaoyang Hospital, Beijing 10020, China
Supported by the Major State Basic Research Development Program of China (973 Program), No. 20004CB518702
Correspondence to: Dr. Hong-gang Qian, Department of Surgery, Peking University School of Clinical Oncology and Beijing Cancer Hospital, 52 Fucheng lu, Haidian District, Beijing 100036, China. qianhg@sohu.com
Telephone: +86-10-88121122-2099 Fax: +86-10-88141293
Received: 2005-03-21 Accepted: 2005-04-18

## Abstract

**AIM:** To explore the preliminary identification of serum protein pattern models that may be novel potential biomarkers in the detection of gastric cancer.

**METHODS:** A total of 130 serum samples, including 70 from patients with gastric cancer and 60 from healthy adults, were detected by surface-enhanced laser desorption and ionization time-of-flight mass spectrometry (SELDI-TOF-MS). The data of spectra were analyzed by Biomarker Patterns Software (BPS). Thirty serum samples of gastric cancer patients and 30 serum samples of healthy adults were grouped into the training group to build models, and the other 70 samples were used to test and evaluate the models. The samples of the test group were judged only with their peaks' height and were separated into cancer group or healthy control group by BPS automatically and the judgments were checked with the histopathologic diagnosis of the samples.

**RESULTS:** Sixteen mass peaks were found to be potential biomarkers with a significant level of $P<0.01$. Among them, nine mass peaks showed increased expression in patients with gastric cancer. Analyzed by BPS, two peaks were chosen to build the model for gastric cancer detection. The sensitivity, specificity, and accuracy of the model were 90%, 36/40, 86.7%, 26/30, and 88.6%, 62/70, respectively, which were greatly higher than those of clinically used serum biomarkers CEA (carcinoembryonic antigen), CA19-9 and CA72-4. Stage I/II gastric cancer samples of the test group were all judged correctly.

**CONCLUSION:** The novel biomarkers in serum and the established model could be potentially used in the detection of gastric cancer. However, large-scale studies should be carried on to further explore the clinical impact on the model.



**Key words:** SELDI-TOF-MS; Serum; Gastric cancer

Qian HG, Shen J, Ma H, Ma HC, Su YH, Hao CY, Xing BC, Huang XF, Shou CC. Preliminary study on proteomics of gastric carcinoma and its clinical significance. *World J Gastroenterol* 2005; 11(40): 6249-6253
http://www.wjgnet.com/1007-9327/11/6249.asp

## INTRODUCTION

Gastric cancer is the second leading cause of cancer death in the world and is predicted still to be one of the leading causes of all deaths in the near future[1,2]. Almost 40% of gastric cancer cases occur in China, where it is the most common cancer. According to WHO IARC (International Agency for Research on Cancer) database and Globocan 2002 database, about 393 000 cases are diagnosed with gastric cancer in China, and of these about 308 000 have died. However, few specific tumor markers with high sensitivity could be useful in the diagnosis of gastric cancer. The sensitivity of existing serum biomarkers carcinoembryonic antigen (CEA), CA19-9 or CA72-4 is only about 40%[3-6].

Proteomic technologies, especially the surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) technology, are providing the tools needed to discover and identify the disease-associating biomarkers[7-10]. Ciphergen Biosystems has developed the ProteinChip technology coupled with SELDI-TOF-MS to facilitate protein profiling of complex biological mixtures[11-14]. In SELDI-TOF-MS analysis, a nitrogen laser desorbs the protein/energy-absorbing molecule mixture from the array surface, enabling the detection of the proteins captured by the array[15,16]. It has been reported recently that the promising SELDI-TOF-MS technology could be used for the diagnosis of cancer by analyzing the protein markers in serum[17,18]. Moreover, the SELDI-TOF-MS technology could be used to analyze other kinds of sample, such as nipple aspirate fluid, urine,

serum-free conditioned medium of cell culture, *etc*.[19-21].

In this study, not only 16 peaks were identified as the potential biomarkers with a significant level of $P<0.01$, but also a model of two markers combination was established, which can be used in the detection of gastric cancer. In the double-blind test, this model was evaluated. The sensitivity, specificity, and accuracy are much better than those of any other biomarker used clinically.

## MATERIALS AND METHODS

### Serum samples

A total of 130 serum specimens from 60 healthy adults and 70 patients with gastric cancer were obtained from Tissue Bank of Beijing Cancer Hospital and Beijing Chaoyang Hospital in 2002-2003. All patients of gastric cancer were confirmed histopathologically. Serum samples of gastric cancer had been collected from patients preoperatively and without chemotherapy. The median age of the patients with gastric cancer was 64 years (range, 32-81 years), and the median age of the control group was 52 years (range, 15-84 years). The TNM staging of 70 patients with gastric cancer are shown in Table 1. All samples were obtained with the consent of the patient and Institutional Review Board approval.

**Table 1** Samples of control and gastric cancer staged by TNM classification

| | Training group | Test group | Total |
|---|---|---|---|
| Gastric cancer | 30 | 40 | 70 |
| Stage I | 6 | 6 | |
| Stage II | 4 | 5 | |
| Stage III | 12 | 16 | |
| Stage IV | 8 | 13 | |
| Control | 30 | 30 | 60 |
| Total | 60 | 70 | 130 |

### Preparation of serum samples for SELDI analysis

All samples were stored at -80 °C until use. WCX2 (Weak Cation Exchange) chips and software were provided by Ciphergen Biosystems (Fremont, CA, USA). The array spots of WCX2 chips were preactivated with binding buffer (100 mmol/L sodium acetate, pH 4.5) at room temperature for 5 min in a humidifying chamber with gently shaking, and repeated once. Each serum sample was first diluted 1:2 with 1% DTT U9 (DTT: dithiothreitol; U9: 9 mol/L urea, 2% Chaps, 50 mmol/L Tris-HCl, pH 9.0), and then diluted 1:12 in binding buffer. One hundred microliters of each diluted sample was spotted onto preactivated WCX2 protein array chips and incubated in a humidfiying chamber for 1 h at room temperature. The chips were washed twice with binding buffer and once with HPLC $H_2O$, air-dried, then sequentially treated by 0.5 μL 100% saturated sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid) solution twice. The sinapinic acid solution was 50% acetonitrile and 0.5% trifluoroacetic acid. The chips were analyzed by the Ciphergen ProteinChip Reader (model PBSII). All reagents were provided by Sigma Co.

### Ciphergen ProteinChip SELDI-TOF-MS analysis

The mass spectra of proteins were generated by using 112 laser shots at a laser intensity of 175-180.The optimized detection size range was between 3 000 and 10 000 u, with a high mass to 100 000 u. The laser was focused at by optimization center. The starting detector sensitivity was set at 9, and the mass deflector was set to 800 u.

### Statistical analysis

The data were analyzed with Ciphergen ProteinChip Software version 3.1.1 (Ciphergen Biosystems). For comparison, the mass peaks of test group ($n$=70) were normalized to that of all 60 training samples using one peak as common calibrate before analysis. The Biomarker Wizard application (Ciphergen Biosystems) was used to automatically detect mass peaks to cluster. Peak labeling was completed by using first-pass 7 S/N, minimum peak threshold 30% of all, second-pass 3 S/N, with 0.3% of the cluster mass window, and estimated peaks were added. The cluster data was analyzed by using Biomarker Patterns™ Software (BPS) 4.0.1 (Ciphergen Biosystems).

## RESULTS

By using Biomarker Wizard application, more than 100 mass peaks were identified in the samples of the training group, and 16 qualified mass peaks were identified with a significant level of $P<0.01$ (Table 2). Nine mass peaks, 7 567, 15 117, 15 326, 15 847, 7 934, 4 524, 8 052, 6 982, 4 714 u, showed increased expression in patients with gastric cancer. The other seven mass peaks, 5 252, 2 675, 5 546, 5 340, 4 177, 3 995, 3 245 u, showed decreased expression in patients with gastric cancer. A representative pseudogel view of specific candidate gastric cancer tumor markers and a stacked trace view of candidate markers from diseased *vs* control individuals are shown in Figure 1.

All clusters were exported to BPS 4.0.1 and analyzed.

**Table 2** Mass peaks found in the model group

| M/Z | *P* | Mean-cancer | Mean-control |
|---|---|---|---|
| 7 567 | 0.0000000037 | 9.92 | 2.14 |
| 15 117 | 0.0000001024 | 10.14 | 1.69 |
| 15 326 | 0.0000002675 | 2.57 | 0.60 |
| 15 847 | 0.0000286434 | 7.08 | 2.32 |
| 5 252 | 0.0000742542 | 1.59 | 4.24 |
| 2 675 | 0.0006036111 | 1.16 | 2.29 |
| 5 546 | 0.0008794540 | 2.09 | 3.97 |
| 7 934 | 0.0012685350 | 8.04 | 3.25 |
| 4 524 | 0.0013355919 | 2.63 | 1.66 |
| 5 340 | 0.0013355919 | 14.90 | 27.46 |
| 8 052 | 0.0016378394 | 3.80 | 1.86 |
| 4 177 | 0.0022105930 | 2.80 | 6.73 |
| 3 995 | 0.0052019227 | 1.35 | 1.97 |
| 6 982 | 0.0074510883 | 2.78 | 1.78 |
| 4 714 | 0.0088747087 | 3.10 | 2.08 |
| 3 245 | 0.0096738522 | 1.49 | 2.26 |

The model was generated by using Gini method with favor-even-splits 0. The v-fold cross-validation was set to 11, while the other options remained as defaults. The relative cost of the model tree is 0.200, shown in Figure 2. Two peaks, 7 567 and 5 252 u, were chosen to make the model tree. And the judgment of cancer or healthy control was made according to the rules of the model tree (data not shown). The double-blind test group samples were normalized to the training group using the same procedure and the same parameters. Seventy test samples were judged only with their peaks' height of the two mass ranges and were separated into cancer group or healthy control group by BPS automatically. The judgments were checked with the histopathologic diagnosis of the test samples. The results demonstrated that the sensitivity, specificity, and accuracy were 90% (95% confidence interval 76.9-96.0%), 86.7% (70.3-94.7%), and 88.6% (79.0-94.1%), respectively, shown in Table 3.

**Table 3** Statistical summary of the model (by BPS)

| Group | Sensitivity (%) | Specificity (%) | Accuracy (%) |
|---|---|---|---|
| Training | 96.7 | 96.7 | 96.7 |
| Test | 90.0 | 86.7 | 88.6 |

## DISCUSSION

Recent advances in genomics and proteomics hold great potential for diagnostic, prognostic, and therapeutic applications[22-24]. They help to discover new therapeutic target, design rational individual drug and obtain early-detection biomarkers[25-28]. Proteomic analysis to identify biomarkers has been reported for the detections of several kinds of cancer, such as pancreatic cancer, ovarian cancer, renal cancer, *etc.*[29-32]. However, to our best knowledge gastric cancer has not found potential biomarkers, which can be used for detection as yet. Although the serum tumor-related antigen, such as CEA, CA72-4 and CA19-9 have been examined as routine in some clinics, their sensitivity and specificity for gastric cancer were too low to be used alone for diagnosis, which limited their diagnostic value. SELDI-TOF-MS technology provides a better and easier tool to identify cancers with barely 3 μL serum.

In this study, we found a novel panel of biomarkers and a model built with them. The sensitivity, specificity, and accuracy of the model in test group were 90%, 86.7%, and 88.6%, respectively, which are greatly higher than those of CEA , CA72-4 and CA19-9 whose molecular weight are more than 100 000 u. This made it possible that the model can be used as biomarkers in the detection for gastric cancer.

The four miss-judged (false negative) test samples of the gastric cancer group were obtained from patients with poorly differentiated tumor, and in stage III or IV, are shown in Table 4. Stage I/II gastric cancer samples of the test group were all judged correctly. It suggested that the model can also be used for early detection, not only for advanced gastric cancer. In the test group, 14 gastric cancer samples obtained from patients were with poorly differentiated tumor. The miss-judged samples may reflect the biological heterogeneity of tumor, which leads the change in serum proteins and influences the judgment. In the test group, 3 of 17 female patients of gastric cancer

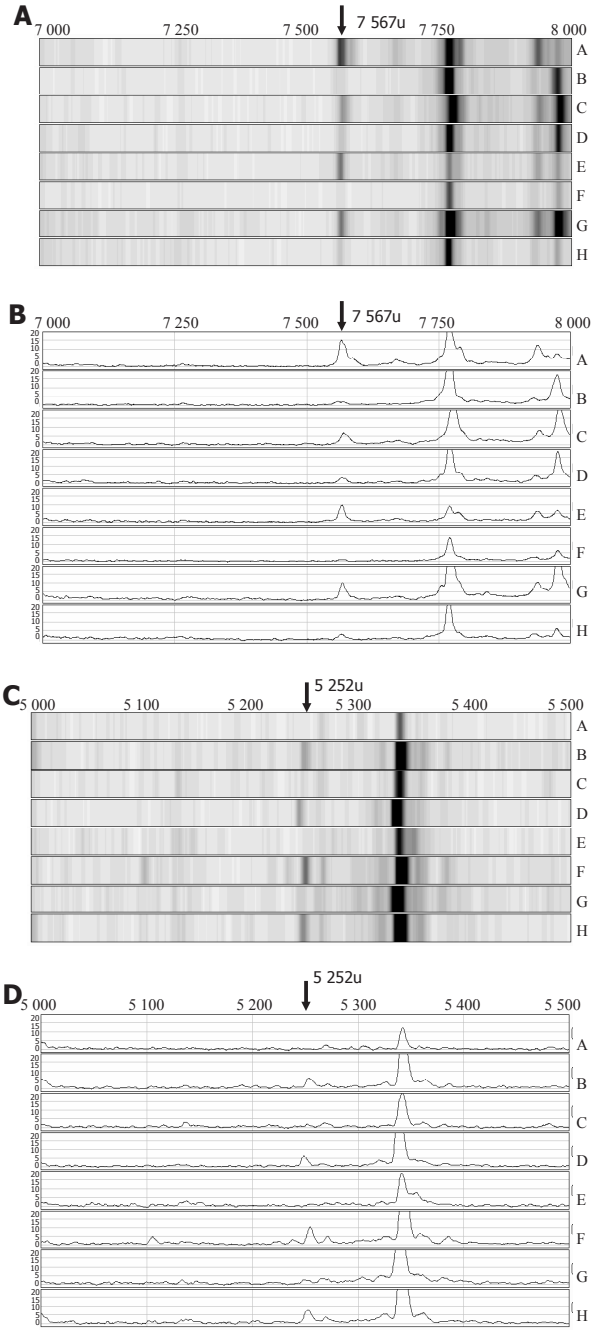
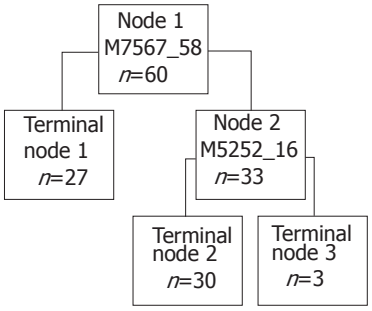



**Figure 1** Detection of differentially expressed gastric cancer-associated serum proteins within a WCX2 chip. The arrows direct to the potential cancer markers detected in the mass spectra, 7 567 u (**A**, **B**) and 5 252 u (**C**, **D**), which were significantly different in gastric cancer samples compared with healthy controls. A, C, E and G were gastric cancer samples and the others were controls. **A** and **C:** A stacked trace view of candidate markers from diseased *vs* control individuals. **B** and **D:** A representative pseudogel view of SELDI-TOF-MS analysis of serum samples.

**A**




**B**

Tree Sequence

| Tree number | Terminal nodes | Resubstitution relative cost | Cross-Validated relative cost | Complexity |
|---|---|---|---|---|
| 1 | 5 | 0.300 ± 0.091 | 0.000 | 0.000 |
| 2 | 3 | 0.200 ± 0.077 | 0.067 | 0.017 |
| 3 | 2 | 0.200 ± 0.077 | 0.167 | 0.050 |
| 4 | 1 | 1.000 ± 0.000 | 1.000 | 0.417 |

**Figure 2** Model tree established by BPS using the training group and relative cost. **A:** Two peaks, 7 567 and 5 252 u, were chosen to make the model tree. n represents the number of the samples that belong to the node; **B:** The relative cost of the model tree is 0.200.

**Table 4** Status of the miss-judged test samples

| Sex | Age (yr) | Differentiation of tumorgy | TNM | Stage |
|---|---|---|---|---|
| 1 Female | 32 | Poor | T4N1M1 | IV |
| 2 Male | 58 | Poor, partially with signet ring cell | T3N2M0 | III |
| 3 Female | 63 | Poor | T4N2M1 | IV |
| 4 Female | 50 | Poor, partially with signet ring cell | T4N1M0 | III |

and 2 of 10 female healthy adults were miss-judged. Therefore, the sensitivities of the female group and the male group are 82.4% and 95.7%, and the specificities are 80% and 90%, respectively. However, the sensitivities had no statistical difference between sexes ($P$=0.394), neither did the specificities ($P$=0.407). Large-scale studies should be carried on to further explore the clinical impact on the model.

Although the model was composed of two peaks, 7 567 u had the highest importance score (data not shown), and can contribute to the model independently with cross-validated relative cost 0.200, as shown in Figure 2B. The peak was significantly higher in the gastric cancer group than in the control group. Therefore, it is more important. We now have focused on the purification and identification of it by using tandem mass spectrometry. If antibodies against the specific proteins are available, we will perform ELISA or immunohistochemistry to further verify the biomarker and make the cancer detection cheaper and easier. Besides, more samples are needed to confirmthat this peak is a specific biomarker of gastric cancer, not of other types of cancer.

## REFERENCES

1 **Pisani P,** Parkin DM, Bray F, Ferlay J. Estimates of the worldwide mortality from 25 cancers in 1990. *Int J Cancer* 1999; **83:** 18-29
2 **Murray CJ,** Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997; **349:** 1498-1504
3 **Gaspar MJ,** Arribas I, Coca MC, Díez-Alonso M. Prognostic value of carcinoembryonic antigen, CA 19-9 and CA 72-4 in gastric carcinoma. *Tumour Biol* 2001; **22:** 318-322
4 **Marrelli D,** Roviello F, De Stefano A, Farnetani M, Garosi L, Messano A, Pinto E. Prognostic significance of CEA, CA 19-9 and CA 72-4 preoperative serum levels in gastric carcinoma. *Oncology* 1999; **57:** 55-62
5 **Carpelan-Holmström M,** Louhimo J, Stenman UH, Alfthan H, Haglund C. CEA, CA 19-9 and CA 72-4 improve the diagnostic accuracy in gastrointestinal cancers. *Anticancer Res* 2002; **22:** 2311-2316
6 **Takahashi Y,** Takeuchi T, Sakamoto J, Touge T, Mai M, Ohkura H, Kodaira S, Okajima K, Nakazato H. The usefulness of CEA and/or CA19-9 in monitoring for recurrence in gastric cancer patients: a prospective clinical study. *Gastric Cancer* 2003; **6:** 142-145
7 **Issaq HJ,** Conrads TP, Prieto DA, Tirumalai R, Veenstra TD. SELDI-TOF MS for diagnostic proteomics. *Anal Chem* 2003; **75:** 148A-155A
8 **Wiesner A.** Detection of tumor markers with ProteinChip technology. *Curr Pharm Biotechnol* 2004; **5:** 45-67
9 **Tang N,** Tornatore P, Weinberger SR. Current developments in SELDI affinity technology. *Mass Spectrom Rev* 2004; **23:** 34-44
10 **Shiwa M,** Nishimura Y, Wakatabe R, Fukawa A, Arikuni H, Ota H, Kato Y, Yamori T. Rapid discovery and identification of a tissue-specific tumor biomarker from 39 human cancer cell lines using the SELDI ProteinChip platform. *Biochem Biophys Res Commun* 2003; **309:** 18-25
11 **Caputo E,** Moharram R, Martin BM. Methods for on-chip protein analysis. *Anal Biochem* 2003; **321:** 116-124
12 **Weinberger SR,** Dalmasso EA, Fung ET. Current achievements using ProteinChip Array technology. *Curr Opin Chem Biol* 2002; **6:** 86-91
13 **Fung ET,** Thulasiraman V, Weinberger SR, Dalmasso EA. Protein biochips for differential profiling. *Curr Opin Biotechnol* 2001; **12:** 65-69
14 **Issaq HJ,** Veenstra TD, Conrads TP, Felschow D. The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochem Biophys Res Commun* 2002; **292:** 587-592
15 **Kozak KR,** Amneus MW, Pusey SM, Su F, Luong MN, Luong SA, Reddy ST, Farias-Eisner R. Identification of biomarkers for ovarian cancer using strong anion-exchange ProteinChips: potential use in diagnosis and prognosis. *Proc Natl Acad Sci USA* 2003; **100:** 12343-12348
16 **Petricoin EF,** Liotta LA. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr Opin Biotechnol* 2004; **15:** 24-30
17 **Cazares LH,** Adam BL, Ward MD, Nasim S, Schellhammer PF, Semmes OJ, Wright GL. Normal, benign, preneoplastic, and malignant prostate cells have distinct protein expression profiles resolved by surface enhanced laser desorption/ionization mass spectrometry. *Clin Cancer Res* 2002; **8:** 2541-2552
18 **Li J,** Zhang Z, Rosenzweig J, Wang YY, Chan DW. Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem* 2002; **48:** 1296-1304
19 **Coombes KR,** Fritsche HA, Clarke C, Chen JN, Baggerly KA, Morris JS, Xiao LC, Hung MC, Kuerer HM. Quality control and peak finding for proteomics data collected from nipple aspirate fluid by surface-enhanced laser desorption and ionization. *Clin Chem* 2003; **49:** 1615-1623
20 **Schaub S,** Wilkins J, Weiler T, Sangster K, Rush D, Nickerson P. Urine protein profiling with surface-enhanced laser-

desorption/ionization time-of-flight mass spectrometry. *Kidney Int* 2004; **65:** 323-332

21 **Sasaki K,** Sato K, Akiyama Y, Yanagihara K, Oka M, Yamaguchi K. Peptidomics-based approach reveals the secretion of the 29-residue COOH-terminal fragment of the putative tumor suppressor protein DMBT1 from pancreatic adenocarcinoma cell lines. *Cancer Res* 2002; **62:** 4894-4898

22 **Cho WC,** Yip TT, Yip C, Yip V, Thulasiraman V, Ngan RK, Yip TT, Lau WH, Au JS, Law SC, Cheng WW, Ma VW, Lim CK. Identification of serum amyloid a protein as a potentially useful biomarker to monitor relapse of nasopharyngeal cancer by serum proteomic profiling. *Clin Cancer Res* 2004; **10:** 43-52

23 **Espina V,** Dettloff KA, Cowherd S, Petricoin EF, Liotta LA. Use of proteomic analysis to monitor responses to biological therapies. *Expert Opin Biol Ther* 2004; **4:** 83-93

24 **Srinivas PR,** Kramer BS, Srivastava S. Trends in biomarker research for cancer detection. *Lancet Oncol* 2001; **2:** 698-704

25 **Chambers G,** Lawrie L, Cash P, Murray GI. Proteomics: a new approach to the study of disease. *J Pathol* 2000; **192:** 280-288

26 **Hobbs SK,** Shi G, Homer R, Harsh G, Atlas SW, Bednarski MD. Magnetic resonance image-guided proteomics of human glioblastoma multiforme. *J Magn Reson Imaging* 2003; **18:** 530-536

27 **Sauter ER,** Zhu W, Fan XJ, Wassell RP, Chervoneva I, Du Bois GC. Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer. *Br J Cancer* 2002; **86:** 1440-1443

28 **Zhukov TA,** Johanson RA, Cantor AB, Clark RA, Tockman MS. Discovery of distinct protein profiles specific for lung tumors and pre-malignant lung lesions by SELDI mass spectrometry. *Lung Cancer* 2003; **40:** 267-279

29 **Koopmann J,** Zhang Z, White N, Rosenzweig J, Fedarko N, Jagannath S, Canto MI, Yeo CJ, Chan DW, Goggins M. Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. *Clin Cancer Res* 2004; **10:** 860-868

30 **Bandera CA,** Ye B, Mok SC. New technologies for the identification of markers for early detection of ovarian cancer. *Curr Opin Obstet Gynecol* 2003; **15:** 51-55

31 **Rogers MA,** Clarke P, Noble J, Munro NP, Paul A, Selby PJ, Banks RE. Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: identification of key issues affecting potential clinical utility. *Cancer Res* 2003; **63:** 6971-6983

32 **Fels LM,** Buschmann T, Meuer J, Reymond MA, Lamer S, Röcken C, Ebert MP. Proteome analysis for the identification of tumor-associated biomarkers in gastrointestinal cancer. *Dig Dis* 2003; **21:** 292-298

**Science Editor** Wang XL and Guo SY **Language Editor** Elsevier HK