RESEARCH COMMUNICATION

SELDI-TOF MS Combined with Magnetic Beads for Detecting Serum Protein Biomarkers and Establishment of a Boosting Decision Tree Model for Diagnosis of Pancreatic Cancer

Jing-Yi Qian[1&], Si-Hua Mou[2&], Chi-Bo Liu[2*]

Abstract

Aim: New technologies for the early detection of pancreatic cancer (PC) are urgently needed. The aim of the present study was to screen for the potential protein biomarkers in serum using proteomic fingerprint technology. Methods: Magnetic beads combined with surface-enhanced laser desorption/ionization (SELDI) TOF MS were used to profile and compare the protein spectra of serum samples from 85 patients with pancreatic cancer, 50 patients with acute-on-chronic pancreatitis and 98 healthy blood donors. Proteomic patterns associated with pancreatic cancer were identified with Biomarker Patterns Software. Results: A total of 37 differential m/z peaks were identified that were related to PC (P < 0.01). A tree model of biomarkers was constructed with the software based on the three biomarkers (7762 Da, 8560 Da, 11654 Da), this showing excellent separation between pancreatic cancer and non-cancer., with a sensitivity of 93.3% and a specificity of 95.6%. Blind test data showed a sensitivity of 88% and a specificity of 91.4%. Conclusions: The results suggested that serum biomarkers for pancreatic cancer can be detected using SELDI-TOF-MS combined with magnetic beads. Application of combined biomarkers may provide a powerful and reliable diagnostic method for pancreatic cancer with a high sensitivity and specificity.

Keywords: Pancreatic cancer - biomarker - protein - SELDI-TOF-MS - diagnosis

Asian Pacific J Cancer Prev, **13**, 1911-1915

Introduction

Pancreatic cancer is the fifth leading cause of cancer death and has the lowest survival rate among solid tumors (DiMagno et al., 1999; Kim and Simeone, 2011). Pancreatic cancer occurs without clear and specific symptoms in early phase, however with high degree of malignancy and is prone to metastasis, leading to the late diagnosis in patients – many patients died within 1 year after the diagnosis. Among the 10-20% patients eligible for surgical resection treatment, the 5 years survival rate is 3-4% (Li et al., 2004). Currently the most efficient diagnosis approach is the radiological detection of small pancreatic cancer with diameter less than 2 cm, which could be surgically removed with reliability. However with ultrasonography the rate of correct diagnosis was only 20-40%; Endoscopic ultrasonography (EUS) could detect small tumors with diameter of 2-3 mm, but would cause trauma through its invasive procedure. Thin slice CT with double-phase scan and enhanced MRI/MRCP could detect small pancreatic cancers; CTA and MRA evaluation on the vasculature could also assess whether the cancer is suitable for surgical resection; while 18F-FDG PET/CT has high sensitivity for diagnosis, but the false positive rate is also high. All these approaches described above significantly improved the techniques for early diagnosis of pancreatic cancer, but the successes rely mainly on the experiences of doctors and these examinations were expensive. The alternative is the use of serum biomarkers, such as the glycoprotein antigen CA19-9 (Singh et al., 2011) which has been extensively used in past studies. However low sensitivity and specificity were reported for this biomarker, with a positive rate of 37.5% (Kim et al., 2004) to pancreatic cancer with diameter less than 2 cm. Therefore the use of new biomarkers with improved diagnosis reliability is of great importance.

Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) has been shown to be effective in identifying new serum biomarker of different cancers such as liver cancer and colorectal Cancer (Liu et al., 2009; Liu et al., 2010; Chibo et al., 2011). The aim of this study is to comparatively analyze pancreatic cancer through serum proteomic profiling using SELDI-TOF MS to seek for new biomarkers and set up a diagnostic model for clinical detection of pancreatic cancer.

Materials and Methods

Patients

The study was performed in Taizhou Municipal Hospital, Zhejiang, China in Aug 2010. Preoperative

[1]*Medical Services Section,* [2]*Department of Clinical Laboratory, Taizhou Municipal Hospital, Taizhou, Zhejiang, China* [&]*Equal contributors* **For correspondence: chibo_liu@163.com*

Table 1. Age and Sexual Distribution of All Study Subjects in the SELDI-TOF Experiments

	age(years) 50-60	61-70	Total	Training set	Test set
Pancreatic cancer	52	33	85	60	25
Healthy controls	79	19	98	60	38
Pancreatitis	41	9	50	30	20
Total	172	61	233	150	83

blood was collected from 85 (ages ranging from 54-70 years old) patients undergoing pancreaticoduodenectomy (Whipple procedure) for resectable infiltrating ductal adenocarcinoma of the pancreas, 50 acute-on-chronic pancreatitis (21 patients with acute pancreatitis and 29 patients with chronic pancreatitis) (ages ranging from 51-70 years old) and 98 controls (healthy volunteers, ages ranging from 52-68 years old) at Taizhou Municipal Hospital and The First Affiliated Hospital of Medical College, Zhejiang University from Jan 2010 to Jul 2010 (Table 1). The diagnoses of pancreatic cancer and acute or chronic pancreatitis were made with endoscopic retrograde cholangiopancreatography (ERCP), CT and ultrasound examination. Among the 85 pancreatic patients, 40 patients were Hermreck stage 1, 32 were stage 2 and 13 were stage 3. The patients were prospectively selected after clinical diagnosis; the gender ratio and age range of the patients recruited in present study were due to the local epidemiology. The studies were approved by the local Ethics Committee of Taizhou Municipal Hospital, and had the informed consent of the patients and volunteers. The patients and serum samples were then divided into two groups: the ''training'' set and the blinded ''test'' set (Table 1). The blood samples were collected in 5 ml BD Vacutainers without anticoagulation and allowed to clot at room temperature for up to 1 hr; the samples were then centrifuged at 4 °C for 5 min at 10000 rpm. The sera were frozen and stored at -80 °C for future analysis.

WCX magnetic beads analysis

Sample pretreatments and proteomic analysis in the proteomic profiling analysis, the serum samples from the diseased and control groups were randomized, and blinded to investigators. Serum samples were pretreated with weak cation exchange (WCX) magnetic beads. 10 μl of each serum sample was mixed with 20 μl of U9 solution (9 mol/L urea, 2% CHAPS, PH 9.0) in a 0.5 ml centrifuge-tube and incubated for 30 min at 4 °C. Denatured serum samples were diluted with 370 μl binding buffer (50 mmol/L sodium acetate, 0.1% Triton X-100, pH 4.0). At the same time, 50 μl of WCX magnetic beads were placed in a PCR-tube and the tube was placed in a magnet separator for 1 min, after which the supernatant was discarded carefully by using a pipette. The magnetic beads were then washed twice with 100 μl binding buffer. Then 100 μl of the diluted serum sample was added to the activated magnetic beads, mixed and incubated for 1 h at 4 °C, after which the beads were washed twice with 100 μl binding buffer.

SELDI-TOF MS

Following binding and washing, the bound proteins were eluted from the magnetic beads using 10 μl of 0.5% trifluoroacetic acid. Then, 5 μl of the eluted sample was diluted in 5 μl of SPA (saturated solution of sinapinic acid in 50% acetonitrile with 0.5% trifluoroacetic acid). Two microliters of the resulting mixture was aspirated and spotted onto the unmodified gold-coated ProteinChip array. After air-drying for 5 minutes at room temperature, protein crystals on the chip were scanned with the ProteinChip (Model PBS IIc) reader (Ciphergen) to determine the masses and intensities of all peaks over the range m/z 1,000 to 50,000. The reader was set up as follows: mass range (1,000 to 50,000 Daltons), optimized mass range (1,000 to 20,000 Daltons), laser intensity (200), and sensitivity (9). Mass calibration was performed using an all-in-one peptide reference standard which contained vasopressin (1084.2Da), somatostatin (1637.9Da), bovine insulin β chain (3495.9 Da), human insulin recombinant (5807.6Da), hirudin (7033.6Da) (Ciphergen Biosystems, Fremont, CA, USA). The default background subtraction was applied, and the peak intensities were normalized using the total ion current from a mass charge of 1000 to 50,000Da. A biomarker detection software package (Ciphergen Biomarker Wizards, Ciphergen Biosystems, Inc) was used to detect protein peaks. Protein peaks were selected based on a first pass of signal-noise ratio of 3 and a minimum peak threshold of 20% of all spectra. This process was completed with a second pass of peak selection at 0.2% of the mass window, and the estimated peaks were added. These selected protein peaks were averaged as clusters and were exported to a commercially available software package (Biomarker Patterns, Ciphergen Biosystems, Fremont, CA, USA) for further classification analysis.

Detection and Statistical Data Analysis

The profiling spectra of serum samples from the training set were normalized using total ion current normalization with Ciphergen's ProteinChip Software (version 3.1). Peak labeling was performed with Biomarker Wizard software 3.1 (Ciphergen Biosystems, Fremont, CA, USA). A two-sample t-test was used to compare mean normalized intensities between the case and control groups. The p value was set at 0.01 to be statistically significant. The intensities of selected peaks were then transferred to Biomarker Pattern Software (BPS) to construct the classification tree of pancreatic cancer. Briefly (Liang et al., 2006; Chibo et al., 2010), the intensities of the selected peaks were submitted to BPS as a 'Root note'. Based on peak intensity, a threshold was determined by BPS to classify the root node into two child nodes. If the peak intensity of a blind sample was lower than or equal to the threshold, this peak would be labeled as "left-side child node." Peak intensities higher than the threshold would be marked as "right-side child node." After rounds of decision making, the training set was found to be discriminatory with the least error.

All of the protein peak intensities of samples in the test set were evaluated by BPS using the classification model. The pancreatic cancer and control samples were then discriminated based on their proteomic profile characteristics. The sensitivity was defined as the

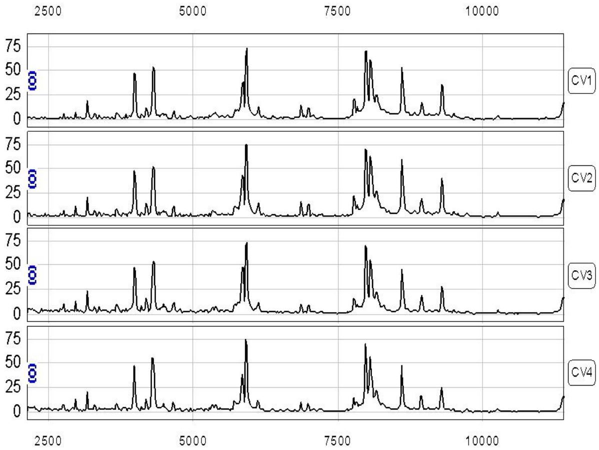

Figure 1. An 4-spot Reproducibility Test Showed Good Reproducibility

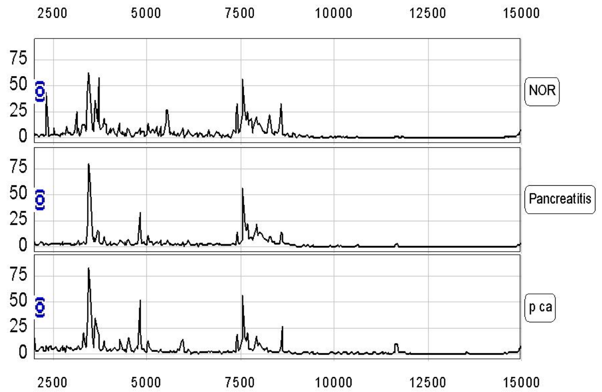

Figure 2. Representative Protein Spectrum of Serum Samples Respectively from the Healthy Controls (NOR), patients with pancreatitis and pancreatic cancer (P ca) detected by SELDI-TOF-MS combined with WCX magnetic beads, showing the protein m/z between 2,000 and 15,000

probability of predicting pancreatic cancer cases, and the specificity was defined as the probability of predicting control samples. A positive predictive value reflected the probability of pancreatic cancer if a test result was positive.

Results

Quality Control and Reproducibility

The quality control (QC) serum sample, 4 mixed serum samples from healthy control subjects with blood type O (2 women and 2 men), was used to determine reproducibility and as a control protein profile for each SELDI Combined With Magnetic Beads experiment. Both the coefficient of variation (CV) for intensity and mass/charge (m/z) were calculated based on duplicate sample testing. The intrachip and interchip CV for intensity were <5%. Both the intrachip and interchip CV for m/z were <0.05%. These values indicated high reproducibility of spectra with SELDI-TOF MS (Figure1).

Detection of the Protein Peaks

Proteomic data from the samples of the training set (consisting of 60 pancreatic cancer, 30 pancreatitis and 60 healthy controls) were analyzed with Biomarker Wizard software 3.1. Up to 239 protein peaks per spot were detected between m/z 2000 and m/z 50000 and it showed the effectiveness of the SELDI technology separation of low molecular weight proteins (<1 5000) (Figure 2).

Table 2. The 37 Discriminating m/z Peaks Among Pancreatic Cancer and Controls

m/z	p	m/z	p	m/z	p
4967.1	5.7×10^{-7}	3378.6	3.7×10^{-4}	4476	8.9×10^{-3}
7762.0*	7.3×10^{-7}	3400.7	4.6×10^{-4}	9713.5	9.1×10^{-3}
3161.6	9.6×10^{-7}	4158.8	7.3×10^{-4}	3288.4	9.5×10^{-3}
3936.2	3.3×10^{-6}	6918.5	8.9×10^{-4}	4794.5	9.9×10^{-3}
6213.1	5.1×10^{-6}	2744.9	9.1×10^{-4}	8087.9	0.001
8560.0*	7.3×10^{-6}	3219.9	9.8×10^{-4}	6361.9	0.001
4137.2	9.5×10^{-6}	3320.6	1.0×10^{-3}	4112.6	0.002
5279.6	1.1×10^{-5}	4093.6	1.0×10^{-4}	4998	0.002
6193.2	1.8×10^{-5}	5807	2.2×10^{-4}	6838.2	0.004
4946.9	4.4×10^{-5}	6305.5	4.0×10^{-4}	15862.1	0.008
11654.0*	8.1×10^{-5}	5056.2	6.4×10^{-4}	5635	0.009
4253.8	1.5×10^{-4}	4299.3	7.6×10^{-4}	N/A	N/A
2942.4	2.3×10^{-4}	4316.2	8.0×10^{-4}	N/A	N/A

m/z means mass-to-charge ratio; P was generated by peak comparison between pancreatic cancer and normal controls; Peaks labeled by *were selected as biomarkers for pancreatic cancer diagnostic model

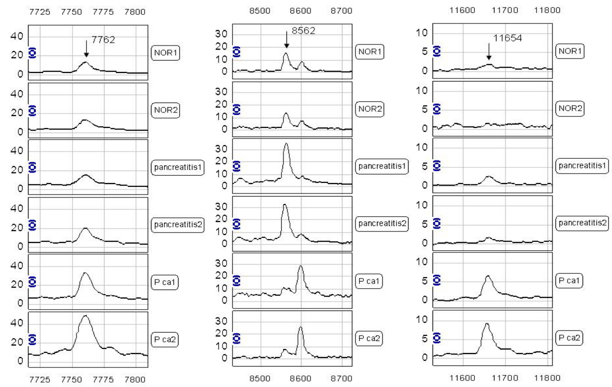

Figure 3. Differential Expression of SELDI Peak m/z 7762, 8560, 11654 in Pancreatic Cancer, pancreatitis and healthy controls. Relative peak intensity is displayed along the y-axis, and mass/charge ratios are shown on the x-axis

Protein Fingerprint Analysis of Serum Samples in Patients with pancreatic cancer, pancreatitis and Healthy Controls

The protein profile of the serum samples from the 60 pancreatic cancer, 30 pancreatitis and 60 healthy controls were extracted by magnetic beads and examined by SELDI-TOF-MS. The data were analyzed by Biomarker Wizard Version 3.1; 37 differential m/z peaks were found from serum samples of the patients with pancreatic cancer, pancreatitis and healthy controls (Table 2).

Protein Fingerprint Analysis of Serum Samples in Patients with pancreatic cancer, pancreatitis and Healthy Controls

The protein profile of the serum samples from 60 pancreatic cancer patients, 30 pancreatitis patients and 60 healthy controls were extracted by magnetic beads and examined by SELDI-TOF-MS. The data were analyzed by Biomarker Wizard Version 3.1; 37 m/z peaks were found to discriminate the patients with pancreatic cancer, pancreatitis and Healthy Controls (Table 2). We identified several biomarkers specific for pancreatic cancer (Figure

Table 3. The Prediction Results of the Diagnostic Model for Pancreatic Cancer

Group	Samples	Accurate %
Training set	pancreatic cancer	93.3
	Control+ pancreatitis	95.6
Blinding set	pancreatic cancer	88
	Control+ pancreatitis	91.4

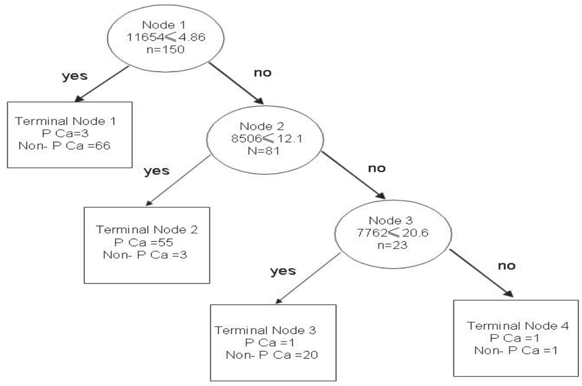

Figure 4. The Decision Trees of Diagnostic Model for Pancreatic Cancer. Each node was represented with different m/z value and the diagnosis result went left or right depending on the detected peaks in test sample. The sensitivity and specificity of diagnosis would significantly increase when several biomarkers were combined in use

3) .Three peaks, m/z 7762Da, 8560Da, 11654Da were then selected to set up the diagnosis tree (Figure 3). At Node l, samples of m/z 11654 with peak intensities lower than or equal to 4.86 went to terminal Node 1, which had 66 controls and 3 PCa samples. Otherwise, samples entered Node 2, which had 24 controls and 57 PCa samples. At Node 2, samples of m/z 8506 with peak intensities lower than or equal to 12.1 went to Node 3, which had 21 controls and 2 PCa samples. The other samples entered terminal Node 2, which had 55 PCa samples and 3 controls. At Node 3, samples of m/z 7762 with peak intensities lower than or equal to 20.6 went to terminal Node 3, which had 20 controls and 1 PCa samples. The other samples went to terminal Node 4, which had 1 PCa samples and 1 controls (Figure 4).

Identification of Biomarker Pattern and Construction of Diagnostic Model

The comparison among different samples showed that the serum profiles from cancer patients and control individuals were very similar in spite of several inter-sample variations. Therefore, the few variations that consistently differentiate these two different groups could be considered as potential disease biomarkers. Here, we used the biomarker wizard function of the ProteinChip software to identify clusters of peaks differentially presented in pancreatic cancer serum samples compared with controls. We obtained 37 different protein peaks in sera (Table 2). To develop biomarker patterns for the diagnosis of pancreatic cancer, the intensities of the protein peaks in the training set were submitted to BPS. A total of three peaks (7762, 8560, 11654) with the highest discriminatory power were automatically selected to construct a classification tree (Figure 3). Figure 4 shows the tree structure and sample distribution. The classification tree discriminated the pancreatic cancer samples from the control samples with a sensitivity of 93.3% and a specificity of 95.6% (Table 3). While in the blind test set, 53 out of 58 true control cases were correctly classified, and 22 out of 25 pancreatic cancer samples were correctly classified as malignant. These results yield a sensitivity of 88% and a specificity of 91.4%.

Discussion

Pancreatic cancer is one of the most common pancreatic tumors and the incidence has increased in past years, accounting for the 6th cause of death by cancer in China. The early diagnosis and treatment are critical to this disease, therefore emphasizing the importance to develop techniques of early detection of pancreatic cancer with high sensitivity and specificity, and with low costs for population examination. Given that fact that during tumor development, multiple genes and proteins altered their expression levels, causing the release of some signaling peptides, metabolites or secreted substances into the circulation, which could be used as the biomarkers for the defined disease. Because proteins with big molecular weight were less likely to enter the circulation, the accurate detection of small molecular weight proteins would provide more evidences in serum-based proteomics for early diagnosis. SELDI or MALDI detection approach when combined with magnetic beads could detect low levels of proteins (up to 1 fmol), and thus offered one feasible approach in early cancer diagnosis (Hortin, 2006; Ward et al., 2006).

There are thousands of different proteins existing in serum, urine and other body fluids to be extracted for the detection of disease biomarkers (Liotta et al., 2003). The dynamic inspection of serum proteomics could be used to screen potential biomarkers and monitor the progress of a given disease, and many new techniques were developed in past decade. For example, SELDI/MALDI combined with magnetic beads has many advantages over traditional proteomic approaches such as 2D-gel electrophoresis and chromatographic separation/purification. The newly developed techniques allowed direct analyses of a small amount of complicated biological samples without prior purification. This technique also allowed examination of multiple samples at the same time, and has very high sensitivity in proteins with low molecular weight or low concentration. Previous studies with this technique achieved many successes in screening of the cancer biomarkers, including liver cancer (Schwegler et al., 2005; Ward et al., 2006), stomach cancer (Ebert et al., 2004), esophagus cancer (Guo et al., 2011), and prostate cancer (Adam et al., 2002; Wagner et al., 2004; Semmes et al., 2005).

In conclusion, the present study showed that the proteomics approaches, such as magnetic beads and SELDI-TOF-MS in combination of bioinformatics tools could facilitate the discovery of new biomarkers, and provide a rapid and accurate mode of analysis for the detection of multiple disease-related proteins simultaneously, reproducibly, as well as in high-throughput manner. With the panel of four selected biomarkers, the diagnostic test achieved high sensitivity and specificity for the detection of pancreatic cancer. It should be noted that in this study each M/Z value may represent many peptides with the same molecular weight. The protein identification studies are yet to be performed. We expect to explore the structure and function of these protein biomarkers for pancreatic cancer in future studies, in a comparative manner with other cancers. Additionally, the study did not investigate the differential biomarkers between healthy subjects and acute-on-chronic pancreatitis patients in our model tree, which could be another interesting topic in our future studies. It is conceivable that some biochemical indices of inflammation such as C-reactive protein, interleukin or indices for pancreatic injury including amylase and lipase could act as such biomarkers.

References

Adam BL, Qu Y, Davis JW, et al (2002). Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res*, **62**, 3609-14.
DiMagno EP, Reber HA, Tempero MA (1999). Epidemiology, diagnosis, and treatment of pancreatic ductal adenocarcinoma. *Gastroenterology*, **117**, 1463–84.
Ebert MP, Meuer J, Wiemer JC, et al (2004). Identification of gastric cancer patients by serum protein profiling. *J Proteome Res*, **3**, 1261-6.
Guo R, Pan C, Shen J, et al (2011). New serum biomarkers for detection of esophageal carcinoma using Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J Cancer Res Clin Oncol*, **137**, 513-9.
Hortin GL (2006). The MALDI-TOF mass spectrometric view of the plasma proteome and peptidome. *Clin Chem*, **52**, 1223-37.
Liu C, Pan C, Shen J, et al (2011). MALDI-TOF MS combined with magnetic beads for detecting serum protein biomarkers and establishment of boosting decision tree model for diagnosis of colorectal cancer. *Int J Med Sci*, **8**, 39-47.
Liu C, Wang H, Pan C, et al (2010). Serum protein fingerprint of patients with gastric cancer by SELDI technology. *African J Biotechnol*, **9**, 2298-304.
Kim JE, Lee KT, Lee JK, et al (2004). Clinical usefulness of Carbohydrate antigen 199 as a screening test for pancreatic cancer in an asymptomatic population. *J Gastroenterol Hepatol*, **19**, 182-6.
Kim EJ, Simeone DM, (2011). Advances in pancreatic cancer. *Curr Opin Gastroenterol*, **27**, 460-6.
Liang Y, Fang M, Li J, et al (2006). Serum proteomic patterns for gastric lesions as revealed by SELDI mass spectrometry. *Exp Mol Pathol*, **81**, 176-80.
Li D, Xie K, Wolff R, et al (2004). Pancreatic cancer. *Lancet*, **363**, 1049-57.
Liotta LA, Ferrari M, Petricoin E (2003). Clinical proteomics: written in blood. *Nature*, **425**, 905.
Liu CB, Liang Y, Pan CQ, et al (2009). Proteome study of differential protein expression in HBV-related primary hepatic carcinoma. *Chem J Chinese Universities*, **30**, 1763-6.
Liu C, Shen J, Pan C, et al (2010). MALDI-TOF MS combined with magnetic beads for detecting serum protein biomarkers and establishment of boosting decision tree model for diagnosis of hepatocellular carcinoma. *Am J Clin Pathol*, **134**, 235-41
Schwegler EE, Cazares L, Steel LF, et al (2005). SELDI-TOF MS profiling of serum for detection of the progression of chronic hepatitis C to hepatocellular carcinoma. *Hepatology*, **41**, 634-42.
Semmes OJ, Feng Z, Adam BL, et al (2005). Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. *Clin Chem*, **51**, 102-12.
Singh S, Tang SJ, Sreenarasimhaiah J, et al (2011). The clinical utility and limitations of serum carbohydrate antigen (CA19-9) as a diagnostic tool for pancreatic cancer and cholangiocarcinoma. *Dig Dis Sci*, **56**, 2491-6.
Wagner M, Naik DN, Pothen A, et al (2004). Computational protein biomarker prediction: a case study for prostate cancer. *BMC Bioinformatics*, **5**, 26.
Ward DG, Cheng Y, N'Kontchou G, et al (2006). Preclinical and post-treatment changes in the HCC-associated serum proteome. *Br J Cancer*, **95**, 1379-83.
Ward DG, Suggett N, Cheng Y, et al (2006). Identification of serum biomarkers for colon cancer by proteomic analysis. *Br J Cancer*, **94**, 1898-905.