RESEARCH ARTICLE

Comparative Serum Proteomic Analysis of Serum Diagnosis Proteins of Colorectal Cancer Based on Magnetic Bead Separation and MALDI-TOF Mass Spectrometry

Bao-Guo Deng[1]*, Jin-Hua Yao[2], Qing-Yin Liu[3], Xian-Jun Feng[4], Dong Liu[5], Li Zhao[2], Bin Tu[6], Fan Yang[1]

Abstract

Background: At present, the diagnosis of colorectal cancer (CRC) requires a colorectal biopsy which is an invasive procedure. We undertook this pilot study to develop an alternative method and potential new biomarkers for diagnosis, and validated a set of well-integrated tools called ClinProt to investigate the serum peptidome in CRC patients. Methods: Fasting blood samples from 67 patients diagnosed with CRC by histological diagnosis, 55 patients diagnosed with colorectal adenoma by biopsy, and 65 healthy volunteers were collected. Division was into a model construction group and an external validation group randomly. The present work focused on serum proteomic analysis of model construction group by ClinProt Kit combined with mass spectrometry. This approach allowed construction of a peptide pattern able to differentiate the studied populations. An external validation group was used to verify the diagnostic capability of the peptidome pattern blindly. An immunoassay method was used to determine serum CEA of CRC and controls. Results: The results showed 59 differential peptide peaks in CRC, colorectal adenoma and health volunteers. A genetic algorithm was used to set up the classification models. Four of the identified peaks at m/z 797, 810, 4078 and 5343 were used to construct peptidome patterns, achieving an accuracy of 100% (> CEA, $P < 0.05$). Furthermore, the peptidome patterns could differentiate the validation group with high accuracy close to 100%. Conclusions: Our results showed that proteomic analysis of serum with MALDI-TOF MS is a fast and reproducible approach, which may provide a novel approach to screening for CRC.

Keywords: peptidome patterns - colorectal cancer - diagnosis - magnetic beads - MALDI-TOF mass spectrometry

Asian Pac J Cancer Prev, **14 (10)**, 6069-6075

Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in males and the second in females, with over 1.2 million new cancer cases and 608,700 deaths estimated to have occurred in 2008 (Jemal et al., 2011). A majority of CRC are either locally or distantly invasive at diagnosis, restricting treatment options and reducing survival rates, whereas the 5-year survival rate is extremely favorable if detected at an early stage and successfully resected (Ciccolallo et al., 2005; Jemal et al., 2010). Therefore, early diagnosis is of importance for CRC patient prognosis (Edwards et al., 2010). Although several screening techniques, such as colonoscopy, fecal occult blood testing (FOBT), and analysis of various serial markers are recommended, the early diagnosis rate of CRC is still comparatively low (Smith et al., 2011). So it remains to be an urgently necessity to explore effective biomarker for diagnosis of CRC.

Proteomics is a new, exciting and largely unexplored area in CRC. Preliminary studies have shown that this technique may provide a novel noninvasive means of diagnosing CRC, and it may have additional value as a prognostic tool (Boja et al., 2012; Cadeco et al., 2012; Fan et al., 2012; Fan et al., 2012; Garay et al., 2012; Heckman-Stoddard et al., 2012). Serum proteome analysis has the potential to facilitate disease diagnosis and therapeutic monitoring, because serum is more easily accessible and widely collected sample, which contains >10,000 different proteins and peptides (Fan et al., 2011; Pan et al., 2011; Fan et al., 2012; Fan et al., 2012). These proteins and peptides are from almost every tissue and cell, and the changes in their quantity and quality are specific not only to the tissues affected but also to the disease process itself. Advances in mass spectrometry (MS) now permit the display of hundreds of small- to medium sized peptides

*[1]Department of Microbiology, Xinxiang Medical University, [2]Paediatrics Intensive Care Unit, [4]Department of Respiratory Medicine, [5]Department of Dermatology, the First Affiliated Hospital of Xinxiang Medical University, Xinxiang, [3]The Clinical Laboratory, No.150 Central Hospital of PLA, [6]Import-Export Inspection and Quarantine Bureau of Luoyang, Luoyang, China *For correspondence: dbgyjh@163.com*

using only microliters of serum (Huang et al., 2009; Fan et al., 2012; Fan et al., 2012) Magnetic bead (MB), based on nanomaterials, has been developed and was considered as a promising material for convenient and efficient enrichment of peptides and proteins in biological samples (Whiteaker et al., 2007; Yao et al., 2008). Combination of MALDI-TOF MS and MB enables high throughput and sensitive investigation of peptides and proteins.

In the current study, we used a novel technology platform, called ClinProt (Bruker Daltonics, Ettlingen, Germany), comprising a weak cationic-exchanger magnetic beads (WCX-MB) based sample separation, MALDI-TOF MS for peptide profiling acquisition, and a bioinformatics package for inspection and comparison of data sets to create "disease-specific" peptidome patterns models, which could serve as a powerful tool for the diagnosis of CRC (Baumann et al., 2005; Shin et al., 2007; West-Norager et al., 2007). The resulting spectra between groups were analysed using post-processing software ClinProt 2.2 and patterns recognition Genetic Algorithm (GA). Diagnostic model, comprised by 4 differentially expressed peptides, were established and validated by the GA, by which different groups were discriminated effectively. The diagnostic model obtained in this manner was further verified in blinded CRC, colorectal adenoma disease and health volunteer samples. Thus, the preliminary work was completed for an early diagnosis and differential diagnosis of CRC from an integrated perspective of peptide mass patterns.

Materials and Methods

Reagents and instruments

The AutoFlex III MALDI-TOF MS, MTP 384 target plate polished steel, α-cyano-hydroxycinnamic (CHCA), MB-WCX kit and peptide calibration standard were purchased from Bruker Daltonics (Leipzig, Germany). Trifluoroacetic acid (TFA) was purchased from Alfa Aesar (Ward Hill, MA, USA). Acetonitrile (ACN) was acquired from Sigma (St. Louis, MO, USA). Diagnostic Kit of Carcinoembryonic antigen (CEA) (ELISA) was acquired from Roche Diagnostics GmbH (Sandhofer Strasse, Germany).

Table 1. Clinical Characteristics of Colorectal Cancer Patients Recruited in Model Construction Group and External Validation Group

Clinical characteristics	Model construction group(n=33)	External valuation group(n=34)	*P**
Gender: female/male	20/13	20/14	0.882
Age (years, $\overline{\chi}$±SD)	63±9.85	61.18±10.54	0.854
TNM stage			0.874
I	3	3	
II	18	18	
III	9	11	
IV	3	2	
Tumor localization			0.727§
Colon	14	13	
Rectum	19	21	

*Determined using Pearson Chi-Square;§Given the limited sample size, samples were combined as early stage (TNM stageIand II) and advanced stage (TNM stage III and IV)

Patients and sample collection

Valid records of the First Attached Hospital of Xin Xiang Medical College Cancer Center were searched for patients with a histological diagnosis of CRC and colorectal adenoma whose serum samples had been obtained at least in recent 3 years. Sixty-five health volunteers (blood donor volunteers), 67 CRC patients, 65 colorectal adenoma patients were enrolled with the permission of the Local Ethical Commission, and blood was collected after informed consent from the patients. Enteroscopy were performed in all health subjects to exclude the presence of incidental colon and rectum mass. Colorectal adenoma patients were diagnosed by enteroscopy and biopsy.

CRC patients underwent clinical staging, surgical excision of the lesion, and were followed up. Pathologic samples were classified according to the 2004 tumor-node-metastasis stage classification (Compton et al., 2004). Serum samples were prepared by collecting blood in a vacuum tube and allowing it to clot for 30 minutes at room temperature. About 1 mL of serum was obtained after centrifugation at 2000 rpm for 10 minutes, and it was stored in small aliquots at -80°C until analysis.

Study design

The data set including 65 healthy volunteers, 55 colorectal adenoma and 67 CRC patients was randomly split into two groups, the clinical characteristics of CRC patients were shown in Table 1. The first group, called model construction group (including 32 health volunteers, 28 colorectal adenoma and 33 CRC patients) was used for the identification of signals related to peptides expressed differentially in CRC patients compared with controls and diagnosis patterns recognition. The second group, called external validation group (including 33 health volunteers, 27 colorectal adenoma and 34 CRC patients) was used for the independent patterns validation of the diagnosis patterns blindly. The accuracy of the peptide model was compared with that of CEA.

The gender ratio (male/female) of health volunteers, colorectal adenoma patients and CRC patients was 1.21, 1.32 and 1.11, respectively. The mean age (years) of health volunteers, colorectal adenoma and CRC patients was 55.13±8.32, 58.33±16.43, and 60.25±12.92. The difference of age and gender of health volunteers in model construction group and external validation group were not significantly. No significant differences were observed for age and gender between CRC and health volunteers, and no significant differences were observed for TNM stage of CRC between model construction group and external evaluation group (Table 1).

Sample purification

We used MB-WCX for peptidome separation of samples following the standard protocol by the manufacturer (Ketterlinus et al., 2005). Step 1, 10μL of WCX-MB binding solution and 10μL of WCX-beads were combined in a 0.5-mL microfuge tube after thoroughly

vortexing both reagents. Step 2, 5μL of serum sample was added to the microfuge tube containing 10μL of WCX-MB binding solution and 10μL of WCX-beads, and mixed by pipetting up and down. Step 3, microfuge tubes were then placed in a magnetic bead separator (MBS) and agitated back and forth 10 times. The beads were collected on the wall of the tubes in the MBS 1 minute later. Step 4, the supernatant was removed carefully by using a pipette. Step 5: 100μL of WCX-MB wash buffer was added to tubes, which were agitated back and forth in the MBS 10 times. The beads were collected on the wall of the tubes, and supernatant was removed carefully by using a pipette. After three washes, 5μL of WCX-MB elution buffer was added to disperse beads in tubes by pipetting up and down. The beads were collected on wall of tubes for 2 minutes and the clear supernatant was transferred into fresh tubes, then 5μL WCX-MB stabilization solutions were added to the collected supernatant, mixing intensively by pipetting up and down, the mixture was then ready for spotting onto MALDI-TOF MS targets and measurement. Finally, prior to the MALDI-TOF MS analysis, we prepared targets by spotting 1μL of the proteome fraction on the polished steel target (Bruker Daltonics). After air drying, 1μL of 3 mg/mL CHCA in 50% ACN and 50% Milli-Q with 2% TFA was applied onto each spot, and the target was air dried again (co-crystallization). The peptide calibration standard (1 pmol/μL peptide mixture) was applied for calibrating the machine.

Mass spectrometry analysis

For proteome analysis, we used a linear Autoflex III MALDI-TOF-MS with the following setting: ion source 1, 20.00kV; ion source 2, 18.60kV; lens, 6.60kV; pulsed ion extraction, 120ns; Ionization was achieved by irradiation with a crystal laser operating at 200.0Hz. For matrix suppression, we used a high gating factor with signal suppression up to 600 Da. Mass spectra were detected using linear positive mode. Mass calibration was performed with the calibration mixture of peptides and proteins in the mass range of 1000-12000Da. We measured three MALDI preparations (MALDI spots) for each MB fraction. For each MALDI spot, 1600 spectra were acquired (200 laser shots at 8 different spot positions). Spectra were collected automatically using the Autoflex Analysis software (Bruker Daltonik) for fuzzy controlled adjustment of critical instrument settings to generate raw data of optimized quality.

The criteria for protein mass peak detection (m/z) were as follows: signal-to-noise ratio (S/N) >5, a 2-Da peak width filter, and a maximum peak number of 200. The intensities of the peaks of interest were normalized with the peak intensity of an ACTH internal standard. More than 10% of the molecular weight was sieved in simultaneous samples, with the discrepancy of identical spinnacle in different samples <0.3% after removal of the initial data noise.

Bioinformatics and statistical analysis

The ClinProt Tools software 2.2 (Bruker Daltonik) was used for analysis of all serum sample data derived from either patients or normal controls. Data analysis began with raw data pretreatment, including baseline subtraction of spectra, normalization of a set of spectra, internal peak alignment using prominent peaks, and a peak picking procedure. The pretreated data were then used for visualization and statistical analysis in ClinProt Tools. Statistically significant different quantity of peptides was determined by means of Welch's t-tests. The significance was set at $p < 0.05$. Class prediction model was set up by GA. A classify peptidome patterns was constructed. To determine the accuracy of the class prediction, firstly, a cross-validation was implemented. Twenty percent of model construction group were randomly selected sample as a test set, and the rest samples were taken as a training set in the class predictor algorithm. Secondly, designed as double blind test, the samples of external valadiation group were classified by the classify peptidome patterns constructed by GA.

Detection of CEA

The serum CEA of 34 CRC, 27 colorectal adenoma and 33 health volunteers included in model valadiation group was detected using an electrochemilum inescent immunoassay method following the standard protocol by the manufacturer (The methods were omitted). The sample was diagnosed as CRC (CEA≥5 ng/ml), otherwise diagnosed as health volunteers (CEA< 5 ng/ml).

Statistical methods, evaluation of assay precision

We analyzed each spectrum obtained from MALDI-TOF MS with AutoflexAnalysis and ClinProt TM software (Bruker Daltonics), the former to detect the peak intensities of interest and the latter to compile the peaks across the spectra obtained from all samples. This allowed differentiation between the cancer and control samples. To evaluate the precision of the assay, we determined within- and between-run variations by use of multiple analyses of bead fractionation and MS for 2 plasma samples. For within- and between-run variation, we examined 3 peaks with various intensities. We determined within-run imprecision by evaluating the CVs for each sample, using 8 assays within a run, then determined between-run imprecision by performing 8 different assays over a period of 7 days. SPSS16.0 was used for analysis of the clinical characteristics of volunteers using χ^2 test or t test. The significance was set at $P < 0.05$. Also, SPSS 16.0 was used to compare the accuracy of the peptidome models and CEA.

Results

For the reproducibility of the protein profiling, Within- and between-run reproducibility of 2 samples were determined with the WCX-MB fractionation and MALDI-TOF MS analysis. In each profile, 3 peaks with different molecular masses were selected to evaluate the precision of the assay. Despite varying peptide masses and spectrum intensities, the peak CVs were all <5% in the within-run and <10% in the between-run assays. These values were consistent with the reproducibility data for the Protein Biology System reported by the manufacturer (Bruker Daltonik). Firstly, we evaluated the differences

Table 2. Statistic of the 59 Differential Peptide Peaks in Colorectal Cancer, Colorectal Adenoma and Health Volunteer Group

Mass	Colorectal cancer (Average±SD)	Colorectal adenoma (Average±SD)	Health volunteer (Average±SD)	*P*
4210.4	89.03±26.49	59.36±17.69	59.4±26.89	0.043
759.87	87.33±23.37	54.31±16.79	61.67±16.75	0.023
760.4	87.12±23.36	48.7±14.61	52.56±13.3	0.005
4215.59	80.03±24.34	22.36±7.67	28.82±15.41	<0.001
5917.99	62.37±28.94	7.83±1.16	14.32±13.53	<0.001
786.12	56.42±17.57	32.95±12.92	35.92±14.59	0.028
5906.51	56.3±26.94	21.98±7.11	34.49±28.35	0.002
786.85	55.91±17.81	32.68±12.8	36.21±14.85	0.031
2662.1	52.46±18.77	25.54±12.82	15.04±12.13	0.003
1618.25	42.37±32	4.7±1.3	23.1±15.97	<0.001
810.84*	30.63±10.98	17.06±5.82	14.82±7.87	0.008
4094.16	30.24±15.91	8.34±4.50	12.3±7.8	0.002
4091.23	30.24±15.91	13.04±6.78	15.92±10.62	0.019
811.68	30.08±11.21	18.84±6.35	14.11±7.31	0.009
1467.15	29.92±25.48	4.52±2.15	11.57±6.41	<0.001
1546.94	28.66±14.01	9.60±3.04	16.78±14.83	<0.001
797.76*	22.81±20.75	119.31±31.26	11.56±3.92	0.007
9326.19	22.07±9.52	2.55±2.65	14.26±13.98	<0.001
3242.81	19.41±9.39	7.03±2.84	6.77±6.6	0.002
4649.53	17.5±6.8	2.97±2.65	10.67±4.88	0.001
7788.6	17.31±7.76	0.08±0.10	0.39±0.32	<0.001
2953.91	16.44±7.36	6.83±2.61	7.99±4.07	0.004
1866.77	14.18±7.21	4.16±1.79	2.83±1.17	<0.001
4645.79	14.17±5.48	3.70±2.82	11.89±5.29	0.007
5343.62*	13.87±6.96	1.33±0.86	2.19±2.52	<0.001
5338.47	12.65±6.23	4.92±2.00	6.51±6.28	0.005
3263.33	11.95±5.45	4.00±0.70	5.53±2.98	<0.001
3265.04	11.86±5.45	3.07±0.49	5.06±2.83	<0.001
2106.31	10.62±3.97	7.78±2.42	6.76±2.78	0.045
2864.56	10.45±11.83	3.07±0.84	3.79±6.45	0.009
4972.38	9.74±7.65	6.02±9.76	0.93±1.03	0.006
1450.79	9.49±5.47	4.16±0.80	3.87±1.86	<0.001
3952.11	9.14±5.13	3.83±2.86	3.03±1.77	0.007
3955.46	9.12±5.13	2.12±1.44	2.05±0.92	0.00178
4057.53	8.96±5.12	6.30±4.64	2.22±0.76	0.004
4055.44	8.65±4.85	8.09±7.29	2.37±0.74	0.004
3885.75	8.18±3.03	1.41±0.88	4.39±1.42	<0.001
1520.78	6.35±3.42	2.03±0.48	2.71±1.22	<0.001
7769.65	6.11±3.62	0.72±0.66	7.55±2.63	<0.001
3193.79	6.00±2.52	2.17±0.71	2.25±1.42	<0.001
4267.77	5.52±3.5	2.82±1.18	2.59±1.61	0.02
4126.53	5.41±3.32	2.47±1.39	2.03±1.34	0.011
1779.76	5.33±2.72	2.15±0.77	1.51±0.56	0.002
2884.72	5.06±6.35	2.03±1.04	1.35±0.95	0.038
4272.96	4.95±3.29	1.07±0.12	1.14±0.59	<0.001
2933.65	4.77±2.15	1.94±0.77	2.62±1.14	0.005
2935.22	4.73±2.13	1.92±0.81	2.47±1.05	0.005
4078.95	4.48±1.61	1.75±1.07	1.65±1.08	0.005
4075.52	4.47±1.57	1.87±1.28	1.81±1.04	0.006
2901.59	3.51±4.87	1.1±0.26	0.93±0.92	0.029
2646.38	2.98±1.17	1.66±0.89	0.85±0.48	0.002
4175.08	2.86±1.28	2.04±1.02	1.14±0.56	0.008
3938.59	2.75±1.40	1.68±1.70	0.73±0.31	0.00351
2740.65	2.69±1.11	2.13±1.36	1.01±0.34	0.004
2992.21	2.31±2.14	1.87±1.53	0.65±0.41	0.024
4160	1.80±0.80	1.60±0.73	0.84±0.42	0.016
5073.04	1.50±1.03	1.19±0.73	0.41±0.27	0.009
4251.24	1.42±0.86	4.35±2.41	0.71±0.41	0.032
4253.52	1.41±0.87	2.25±0.64	0.54±0.34	0.008

*Differential peptide peaks used for diagnosis model construction

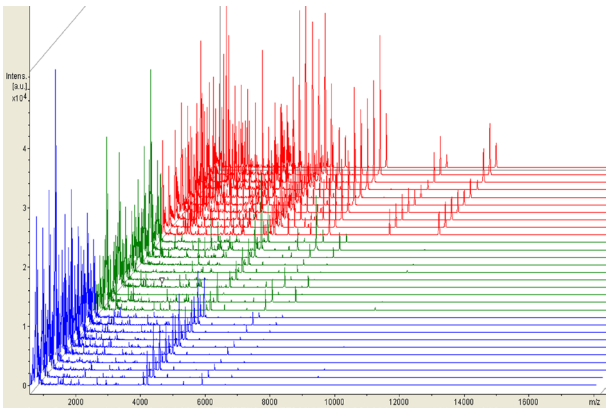

Figure 1. View of the Aligned Mass Spectra of the Serum Protein Profile of Model Construction Group (Red represent 10 colorectal cancer patients, green represent 10 colorectal adenoma patients, and blue represent 10 healthy volunteers) **Obtained by MALDI-TOF after Purification with WCX Magnetic Beads**

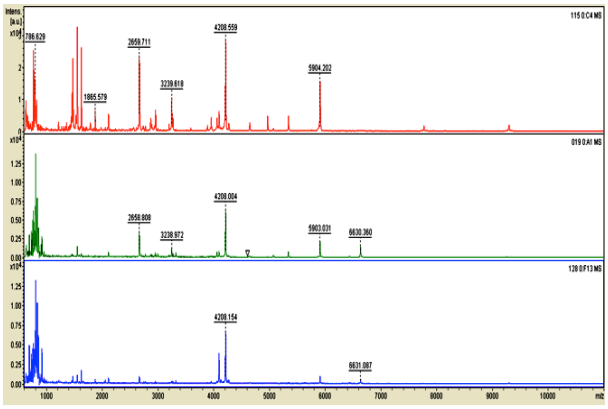

Figure 2. Representative Mass Spectrum (Red represent colorectal cancer patients, green represent colorectal adenoma patients, and blue represent healthy volunteers)

of the serum proteome profiles of CRC in comparison to colorectal adenoma and health volunteers. The mass spectra from 1 to 18 kDa were obtained by MALDI-TOF MS in linear mode. The representative mass spectra of pre-fractionated serum of model construction group are reported in Figure 1. On average about 86 signals common to the three groups have been detected in this mass range and about 59 were identified by the ClinProt software with a statistically different area (P<0.05 by Wilcoxon analysis) in model construction population, including 57 up-regulated and 2 down-regulated peptides. The differential peptide peaks were shown in Table 2.

Classification models were developed to classify samples between CRC and controls (colorectal adenoma and health volunteers). The use of individual peaks as diagnostic biomarker for CRC was addressed using GA analysis. First, we conducted comparison between CRC and controls (colorectal adenoma and health volunteers). Second, all detected peaks were analysed by ClinProt 2.2 to generate cross-validated classification models. The optimized model resulted in the following correct classification of samples. Four peptide ion signatures (m/z 797, 810, 4078 and 5343) was provided as a class prediction for a cross-validation set to discriminate CRC from controls (colorectal adenoma and health volunteers),

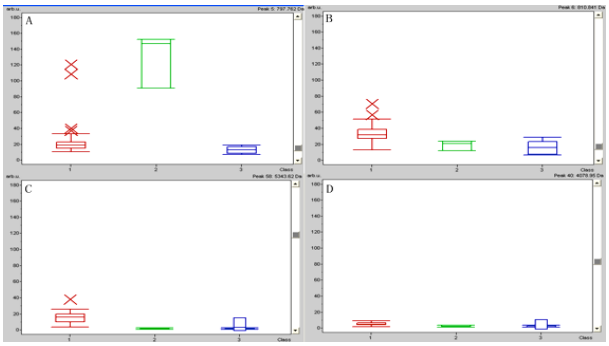

Figure 3. Box-and-whiskers Plot Calculated from the Areas of the 4 Signals Used in the Cluster for the 3 Studied Populations. (Red represent colorectal cancer patients, green represent colorectal adenoma patients, and blue represent healthy volunteers)

which achieved a recognition capacity of 100% and a cross-validation of 100%. Regions of the mass spectra obtained at 800 resolving are reported in Figure 2.

Preliminary statistical analysis was carried out for each single marker. Areas of these four the different peptide ion signatures (peak A at m/z 797 ($P = 0.0069$), peak B at m/z 810 ($P < 0.00812$), peak C at m/z 4078 ($P < 0.00456$) and peak D at m/z 5343 ($P < 0.00001$)) in the spectra of CRC were statistically different from those of controls (colorectal adenoma and health volunteers) (Figure 3). Combination of the two peaks allowed to yielding a specificity of 100% (60/60), and a sensitivity of 100% for CRC (33/33).

To verify the accuracy of the established GA classification model with the adopted peptides, we introduced another group of samples (not used in model construction), which consisted of 34 CRC, 27 colorectal adenoma and 33 health volunteers. As a result, the model correctly classified 97.06% (33/34) of CRC (sensitivity), 100% (27/27) of colorectal adenoma (specificity) and 93.94% (31/33) of health volunteers (specificity), which surpassed that of CEA (a specificity of 52.94% (18/34), and a sensitivity of 60% (36/60)).

Discussion

The usefulness of multiple markers for diagnosis, prognosis and for predicting the risk of developing diseases or their complications is now widely recognized (Issaq et al., 2007; Issaq et al., 2008). Various proteomic approaches have been applied to biomarker discovery using biological fluids. It is being interestingly recognized that low mass weight peptides, such as S100A8 and fibrinogen, play an important role in physiological and pathological process and could be used as relevant biomarker candidates (Cheng et al., 2005; Tolson et al., 2006). Recently mass spectrum that directly detects and differentiates short peptides has offered a promising approach for peptidomic biomarker discovery (Alagaratnam et al., 2008; Chinello et al., 2010; Dai et al., 2010; Du et al., 2010; Liu et al., 2010).

Advancements in genomics, proteomics and bioinformatics have improved our understanding of the cause, carcinogenesis and progression of the disease (Cho et al., 2007). Proteomic profiling is based upon the fact that proteins represent the dynamic state of the cells, reflecting earlier pathophysiological changes in the disease more accurately than genomic sequencing (Hudler et al., 2010). Proteomic patterns should assist in the detection of tumor biomarkers, as well as in evaluating the efficacy of anticancer drugs. Unfortunately, the proteome associated with CRC early diagnosis is currently poorly understood. In this study, an extensive proteomic analysis in the serum of patients with CRC was performed. A standardized serum preparation method for MALDI-TOF-MS was utilized based on weak cation magnetic beads and was able to identify many valuable, low-abundance protein masses of interest. MALDI-TOF-MS is capable of detecting proteins that can aid in the diagnosis of many common types of cancer. Serum proteomics profiling may also help predict the response to treatment, in addition to improving our understanding of diagnosis (Gemoll et al., 2010). Such analyses have identified several clues concerning the markers of diagnosis. Wang et al identified two serum protein biomarkers (3961 and 5200 m/z) useful for monitoring colorectal cancer (Wang et al., 2009). Another recent study confirmed the decreased serum levels of apolipoprotein C-1 in CRC (Engwegen et al., 2008). Additionally, protein markers associated with lymph node metastases in colorectal and prostate cancer have been profiled (Pang et al., 2010; Fan et al., 2011).

In present study, by integrating the purification of short peptides with WCX-MB, detection of peak intensity with MALDI-TOF MS, and profile analysis with ClinProt Tool software 2.2, we have successfully detected a series of short peptides that differentially expressed in the serum of patients with CRC. A case control comparative analysis between CRC and controls (colorectal adenoma and health volunteer) was performed. Peptidomic maps associated with the disease were drawn. The results show that, compared to controls, CRC sharing 59 significantly differentiated peptides, including 57 up-regulated and 2 down-regulated peptides. Current knowledge of cellular regulation indicates that many networks operate at the epigenetic, transcriptional and translational levels. Genomic and proteomic technologies will help further understand the intracellular signaling and gene transcription systems as well as the protein pathways that connect extracellular microenvironment to the serum or plasma macro-environment of cancer (Cho et al., 2007). These 59 interesting significantly differentiated peptides may provide further evidence for understanding the occurrence and progress of colorectal cancer. In particular, the prominent peptides that have a greater than twofold change in intensity, such as m/z 4215, 5917, 2662 and 1618, may be defined as the leading differential peptides associated with colorectal cancer, worthy of further sequence determination and function analysis.

By using the GA analysis, classification model were developed to classify samples between CRC from controls (colorectal adenoma and health volunteer). A cluster of 4 peptides at m/z 797, 810, 4078 and 5343 achieved a recognition capacity and a cross-validation of close to 100% to discriminate CRC from controls (colorectal adenoma and health volunteer). Blinded verification of the GA classification model proved to correctly classify

97.06% (33/34) of CRC (sensitivity), 100% (27/27) of colorectal adenoma (specificity) and 93.94% (31/33) of health volunteers (specificity), which surpassed that of CEA (a specificity of 52.94% (18/34), and a sensitivity of 60% (36/60)). This demonstrated that the GA would be effective in facilitating the construction of a sensitive and specific diagnostic model.

According to our knowledge, this study is one of the few study to screen CRC related short peptides in sera by combining WCX-MB and MALDI-TOF-MS. The classification model we have set up have application in providing alternatives for CRC diagnosis or differential diagnosis, and may provide a better understanding of the pathogenesis in CRC or help in tailoring the use of chemotherapy to each patient, finally resulting in an improvement in patient outcome. Despite of the high sensitivity and specificity, the number of specimens analyzed in this study was relatively small, which may limit the validity of the results.

In conclusion, we directly profiled peptidome patterns from WCX-MB–purified serum samples with MALDI-TOF MS, and constructed a peptidome model that differentiated CRC from control (colorectal adenoma and health volunteer) with high sensitivity and specificity. The next step of our study will be to analyse larger patient cohorts and to run blinded samples to confirm the usefulness of our currently identified peptides for CRC diagnosis. After this confirmation, we will then isolate and identify the biomarkers of the interest and study their biological role in CRC pathogenesis.

Acknowledgements

The author (s) declare that they have no competing interests.

References

Alagaratnam S, Mertens BJ, Dalebout JC, et al (2008). Serum protein profiling in mice: identification of Factor XIIIa as a potential biomarker for muscular dystrophy. *Proteomics*, **8**, 1552-63.

Baumann S, Ceglarek U, Fiedler GM, et al (2005). Standardized approach to proteome profiling of human serum based on magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Clin Chem*, **51**, 973-80.

Boja ES, Rodriguez H (2012). Mass spectrometry-based targeted quantitative proteomics: achieving sensitive and reproducible detection of proteins. *Proteomics*, **12**, 1093-110.

Cadeco S, Williamson AJ, Whetton AD (2012). The use of proteomics for systematic analysis of normal and transformed hematopoietic stem cells. *Curr Pharm Des*, **18**, 1730-50.

Cheng AJ, Chen LC, Chien KY, et al (2005). Oral cancer plasma tumor marker identified with bead-based affinity-fractionated proteomic technology. *Clin Chem*, **51**, 2236-44.

Chinello C, Gianazza E, Zoppis I, et al (2010). Serum biomarkers of renal cell carcinoma assessed using a protein profiling approach based on ClinProt technique. *Urology*, **75**, 842-47.

Cho WC (2007). Contribution of oncoproteomics to cancer biomarker discovery. *Mol Cancer*, **6**, 25.

Cho WC (2007). Nasopharyngeal carcinoma: molecular biomarker discovery and progress. *Mol Cancer*, **6**, 1.

Ciccolallo L, Capocaccia R, Coleman MP, et al (2005). Survival differences between European and US patients with colorectal cancer: role of stage at diagnosis and surgery. *Gut*, **54**, 268-73.

Compton CC, Greene FL (2004). The staging of colorectal cancer: 2004 and beyond. *CA Cancer J Clin*, **54**, 295-308.

Dai Y, Hu C, Wang L, et al (2010). Serum peptidome patterns of human systemic lupus erythematosus based on magnetic bead separation and MALDI-TOF mass spectrometry analysis. *Scand J Rheumatol*, **39**, 240-6.

Du J, Yang S, Lin X, et al (2010). Use of anchorchip-time-of-flight spectrometry technology to screen tumor biomarker proteins in serum for small cell lung cancer. *Diagn Pathol*, **5**, 60.

Edwards BK, Ward E, Kohler BA, et al (2010). Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. *Cancer*, **116**, 544-73.

Engwegen JY, Depla AC, Smits ME, et al (2008). Detection of colorectal cancer by serum and tissue protein profiling: a prospective study in a population at risk. *Biomark Insights*, **3**, 375-85.

Fan NJ, Gao CF, Wang CS, et al (2012). Discovery and verification of gelsolin as a potential biomarker of colorectal adenocarcinoma in the Chinese population: Examining differential protein expression using an iTRAQ labelling-based proteomics approach. *Can J Gastroentero*, **26**, 41-7.

Fan NJ, Gao CF, Wang XL (2011). Identification of regional lymph node involvement of colorectal cancer by serum SELDI proteomic patterns. *Gastroenterol Res Pract*, 2011, 784967.

Fan NJ, Gao CF, Wang CS, et al (2012).Identification of the up-regulation of TP-alpha, collagen alpha-1 (VI) chain, and S100A9 in esophageal squamous cell carcinoma by a proteomic method. *J Proteomics*, **75**, 3977-86.

Fan NJ, Gao CF, Zhao G, Wang XL, Qiao L (2012). Serum peptidome patterns for early screening of esophageal squamous cell carcinoma. *Biotechnol Appl Biochem*, **59**, 276-82.

Fan NJ, Gao CF, Zhao G, Wang XL, Liu QY (2012).Serum peptidome patterns of breast cancer based on magnetic bead separation and mass spectrometry analysis. *Diagn Pathol*, **7**, 45.

Garay JP, Gray JW (2012). Omics and therapy - a basis for precision medicine. *Mol Oncol*, **6**, 128-39.

Gemoll T, Roblick UJ, Auer G, Jornvall H, Habermann JK (2010).SELDI-TOF serum proteomics and colorectal cancer: a current overview. *Arch Physiol Biochem*, **116**, 188-96.

Heckman-Stoddard BM (2012). Oncology biomarkers: discovery, validation, and clinical use. *Semin Oncol Nurs*, **28**, 93-8.

Huang YJ, Xuan C, Zhang BB, et al (2009). SELDI-TOF MS profiling of serum for detection of nasopharyngeal carcinoma. *J Exp Clin Cancer Res*, **28**, 85.

Hudler P, Gorsic M, Komel R (2010). Proteomic strategies and challenges in tumor metastasis research. *Clin Exp Metastasis*, **27**, 441-51.

Issaq HJ, Xiao Z, Veenstra TD (2007): Serum and plasma proteomics. *Chem Rev*, **107**, 3601-20.

Jemal A, Bray F, Center MM, et al (2011). Global cancer statistics. *CA Cancer J Clin*, **61**, 69-90.

Jemal A, Siegel R, Xu J, Ward E (2010).Cancer statistics2010. *CA Cancer J Clin*, **60**, 277-300.

Ketterlinus R, Hsieh SY, Teng SH, Lee H, Pusch W (2005).

Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software. *Biotechniques*, **Suppl**, 37-40.

Liu LH, Shan BE, Tian ZQ, et al (2010). Potential biomarkers for esophageal carcinoma detected by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Clin Chem Lab Med*, **48**, 855-61.

Pang J, Liu WP, Liu XP, et al (2010).Profiling protein markers associated with lymph node metastasis in prostate cancer by DIGE-based proteomics analysis. *J Proteome Res*, **9**, 216-26.

Pan C, He N, Zhao M, et al (2011). Subdividing the M1 stage of liver metastasis for nasopharyngeal carcinoma to better predict metastatic survival. *Med Oncol*, **28**, 1349-55.

Shin S, Cazares L, Schneider H, et al (2007). Serum biomarkers to differentiate benign and malignant mammographic lesions. *J Am Coll Surg*, **204**, 1065-71; discussion 1071-3.

Smith RA, Cokkinides V, Brooks D, et al (2011). Cancer screening in the United States, 2011: A review of current American Cancer Society guidelines and issues in cancer screening. *CA Cancer J Clin* , **61**, 8-30.

Tolson JP, Flad T, Gnau V, et al (2006). Differential detection of S100A8 in transitional cell carcinoma of the bladder by pair wise tissue proteomic and immunohistochemical analysis. *Proteomics*, **6**, 697-708.

Wang Q, Shen J, Li ZF, et al (2009).Limitations in SELDI-TOF MS whole serum proteomic profiling with IMAC surface to specifically detect colorectal cancer. *BMC Cancer*, **9**, 287.

West-Norager M, Kelstrup CD, Schou C, et al (2007). Unravelling in vitro variables of major importance for the outcome of mass spectrometry-based serum proteomics. *J Chromatogr B Analyt Technol Biomed Life Sci*, **847**, 30-7.

Whiteaker JR, Zhao L, Zhang HY, et al (2007). Antibody-based enrichment of peptides on magnetic beads for mass-spectrometry-based quantification of serum biomarkers. *Anal Biochem*, **362**, 44-54.

Yao N, Chen H, Lin H, Deng C, Zhang X (2008). Enrichment of peptides in serum by C (8)-functionalized magnetic nanoparticles for direct matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis. *J Chromatogr A*, **1185**, 93-101.

Zethelius B, Berglund L, Sundstrom J, et al (2008).Use of multiple biomarkers to improve the prediction of death from cardiovascular causes. *N Engl J Med*, **358**, 2107-16.