SPANDIDOS PUBLICATIONS

Identification of serum proteins as prognostic and predictive markers of colorectal cancer using surface enhanced laser desorption ionization-time of flight mass spectrometry

HELGI H. HELGASON[1], JUDITH Y.M.N. ENGWEGEN[1,4], MARC ZAPATKA[6], ANDREW VINCENT[3], ANNEMIEKE CATS[2], HENK BOOT[2], JOS H. BEIJNEN[4,5] and JAN H.M. SCHELLENS[1,5]

Divisions of [1]Clinical Pharmacology and [2]Gastroenterology and Hepatology, Department of Medical Oncology and [3]Biostatistics, The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam; [4]Department of Pharmacy and Pharmacology, The Netherlands Cancer Institute, Slotervaart Hospital, Amsterdam; [5]Division of Drug Toxicology, Section of Biomedical analysis, Department of Pharmaceutical Sciences, Faculty of Science, Utrecht University, Utrecht, The Netherlands; [6]Department of Theoretical Bioinformatics, German Cancer Research Center, Heidelberg, Germany

Received December 11, 2009; Accepted February 1, 2010

DOI: 10.3892/or_00000828

Abstract. Colorectal cancer (CRC) is the second most common cause of cancer related death. Prognosis is highly dependent on stage at diagnosis making early detection mandatory. This study aimed to identify novel disease specific biomarkers of CRC, validate our previously identified biomarkers of CRC and identify serum biomarkers predicting treatment response and for monitoring. Serum of patients with metastatic CRC was collected, according to a predefined schedule, prior to start of standard first-line chemotherapy with oxaliplatin and capecitabine and serially before each 3 weekly treatment cycle and analyzed for proteomic profile by standardized SELDI-TOF MS. Serum proteomic mass spectrometry data of all subjects were processed using the tbimass R-package and proteomic profiles of CRC patients were compared with those of matched normal control subjects. Furthermore, changes in proteomic profiles during the course of chemotherapy were recorded according to treatment response. In total, 42 patients with advanced CRC were treated and mean follow-up was 13.5 months. The response rate was 50% and the median overall survival 19.5 months (95% CI: 16-23). By comparing CRC patients and healthy controls we identified 13 potential biomarkers of CRC (*m/z* 2.0-31.9 kDa) whereas two proteins, *m/z* 14060 and 28100 Da (apolipoprotein A-I), were highly significant ($p<0.0001$). Comparison of responding and non-responding patients identified 6 proteins potentially predicting response, where of *m/z* 3330 Da was significant ($p=0.007$). Serial analysis identified 2 proteins, *m/z* 2022 and 28100 Da, that changed during chemotherapy in accordance with response. We identified 13 *m/z* values discriminating between CRC patients and healthy controls, including the previously identified apolipoprotein A-I as a candidate biomarker for CRC and treatment monitoring.

Correspondence to: Dr Helgi H. Helgason, Division of Clinical Pharmacology, Department of Medical Oncology, The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands
E-mail: h.helgason@nki.nl

Key words: proteomic, profile, colorectal, cancer, biomarker, prognostic

Introduction

Excluding skin cancer, colorectal cancer (CRC) is the third most commonly diagnosed cancer and the second most common cause of cancer-related deaths in Europe and the United States accounting for about 10% of all cancer-related deaths (1). Epidemiologic and genetic studies suggest that CRC results from a complex interaction between inherited susceptibility and multiple environmental and lifestyle factors. The natural development of CRC, through series of specific mutations, follows gradual progression from benign adenomatous polyps with increasing dysplasia to infiltrating adenocarcinoma and advanced disease (2,3). When CRC is diagnosed and treated at an early stage the 5-year survival rate reaches 90%, but declines to 60% in case of loco-regional disease and 10% in case of metastatic disease at the time of diagnosis (1). Unfortunately, only 40% of all CRC are diagnosed at an early stage making CRC a major cause of morbidity and mortality worldwide. Early detection of CRC, by screening with different techniques, has not reached broad acceptance because of limited sensitivity, high false negative and/or positive results, their invasive character and low patient compliance (4-9). The high curability of early stage disease and the declining survival with higher disease stage

urges the development of sensitive non-invasive screening methods with widespread applicability. Serum biomarkers are among the most promising future screening tools but also of great interest in the search of better predictive markers of chemotherapy response and treatment tailoring.

With surface-enhanced laser desorption/ionization-time of flight mass spectrometry (SELDI-TOF MS) technology we are able to analyze the relative expression levels of proteins over a wide range of molecular weights in biological samples. Differences in protein expression level could then be used to identify for example disease state by its proteomic profile or 'fingerprint' from biological samples such as serum. Proteomic analysis also avoids overlooking post-translational modifications and is ideal to analyze changes occurring during chemotherapy. In a retrospective analysis in patients with advanced colorectal cancer we previously identified serum peptides that differentiated between normal controls (NC) and patients with CRC (10).

We hypothesized that we would be able to: i) identify novel peptides that differentiate between NC and CRC patients; ii) identify peptides that predict response to chemotherapy; iii) confirm our previous observation of proteins that differentiate between NC and CRC patients; and iv) identify peptides that change differentially over time in chemotherapy responsive and non-responsive patients. Serum of patients with advanced CRC was prospectively collected, prior to and during first-line chemotherapy with oxaliplatin and capecitabine, and analyzed with SELDI-TOF mass spectrometry simultaneously with serum from normal control subjects. Subsequently, we compared baseline (t=0) serum proteomic profiles of CRC patients with those of NC and investigated whether the baseline proteomic profiles of CRC patients were different between treatment responders and non-responders.

Materials and methods

Subject characteristics. We prospectively collected serum samples from chemotherapy naive patients with histologically confirmed advanced colorectal cancer (CRC) that were eligible for standard first-line chemotherapy. The study was approved by our local medical ethics committee and all patients gave written informed consent. Only patients with performance score WHO ≤2, measurable disease according to RECIST criteria (11) and acceptable hematologic, renal, and liver function: neutrophils ≥1.5x10^9/l, platelets ≥100x10^9/l, creatinine ≤130 μmol/l, liver AST and ALT less than twice the ULN (≤5 x ULN in case of liver metastases) and bilirubinemia <1.5 times the ULN, were included. Patients treated with pre-operative chemoradiotherapy for locally advanced rectal cancer with a low dose of a fluoropyrimidine at least 6 months prior to study entry were allowed to participate. The control group consisted of healthy subjects that were selected based on a short questionnaire and matched for age, gender and time period of blood donation.

Treatment. All patients received first line chemotherapy with oxaliplatin 130 mg/m^2 given intravenously in 120 min on day 1, followed by capecitabine tablets 1000 mg/m^2, twice daily orally, on day 1-14, in a cycle of 3 weeks. Oxaliplatin was given in 500 ml glucose 5% with standard pre-medications and anti-emetics. Tumor response was assessed every other cycle by computer tomography scan.

Sample collection. Whole blood samples were obtained at regular predefined time-intervals starting within 2 weeks prior to the start of chemotherapy and immediately prior to each chemotherapy cycle. All blood samples were collected during a standardized drawing and handling procedure in standard tubes (BD Vacutainer™ SST II 8.5 ml, BD Co., Franklin Lakes, NJ, USA). Samples were allowed to clot during 15 min and centrifuged within 1 h at 3000 rpm for 10 min at room temperature. Subsequently, the serum was transferred in equal aliquots to 5 polypropylene tubes (1.4 ml) and stored at -30˚C until analysis. All serum samples originate from the Netherlands Cancer Institute serum bank. The primary analysis was the comparison of the proteomic profiles of colorectal cancer patients (CRC) and normal subjects (NC). For the subsequent analysis of proteomic profile differences between responding patients and non-responders the patients were divided in two groups according to response: i) responders; patients developing complete response, partial response and stable disease of >6 months duration, respectively; and ii) non-responders; patients developing stable disease of <6 months duration or progressive disease, respectively. In the search for prognostic proteomic profile predicting survival we divided the patients according to shorter/equal to or longer than 6 months survival.

SELDI-TOF analysis. Protein profiling was performed using SELDI-TOF mass spectrometry (Bio-Rad Laboratories, Hercules, CA, USA). Previously, we screened different chromatographic and binding conditions in patients with colorectal cancer (10). The CM 10 chip is a weak cation exchange chip containing anionic carboxylate groups that bind positively charged proteins in serum. A binding buffer of 20 mM sodium phosphate + 0.1% Triton X-100 (Sigma, St. Louis, MO, USA) (pH 5.0) and a 100% solution of sinapinic acid (SPA; Bio-Rad Laboratories) in 50% acetonitrile (ACN) + 0.5% trifluoracetic acid (TFA) as energy absorbing matrix gave the most discriminating *m/z* values between CRC patients and NC (10). All samples were thawed only once for analysis. After thawing the serum samples were denatured by adding 180 μl of a solution containing 9 M urea, 2% CHAPS, 1% DTT (all from Sigma) to 20 μl of serum. CM 10 chips were assembled in 96-well format bioprocessors (Bio-Rad Laboratories). During all steps of the protocol, the bioprocessor was placed on a platform shaker at 350 rpm. Chips were equilibrated twice with 200 μl of binding buffer for 5 min. Subsequently, 180 μl of binding buffer and 20 μl of denatured sample were applied to the chip surface. Sample allocation was at random for comparison of CRC vs. control sera. For the analysis of serial CRC sera all samples from the same patient were analyzed on the same chip whenever possible and remaining samples were allocated at random. For quality control a separate sample from a normal subject was used and spotted on remaining locations (4-6 spots) across the bioprocessor. Incubation was set to 30 min. After binding, the chips were washed twice for 5 min with

buffer, followed by two 5-min washes with binding thout Triton X-100. Finally, chips were rinsed with de-ionized water; air dried and finished with two 1-μl SPA applications to the sample spots.

The reproducibility of the applied methodology was previously validated by our group (10). Protein chips were analyzed using the PBS-IIC ProteinChip Reader (Bio-Rad Laboratories). Data were collected between 0 and 200,000 Da). Data collection was optimized for detection of discriminating peaks, resulting in an average of 65 laser shots per spectrum at laser intensity 150, detector sensitivity 8 and laser focusing at 3000 Da. *M/z* values for the detected proteins were calibrated externally with a standard peptide mixture (Bio-Rad Laboratories) containing vasopressin (1084.3 Da), somatostatin (1637.9 Da), dynorphine (2147.5 Da), ACTH (2933.5 Da), insulin ß-chain (bovine; 3495.5 Da), insulin (human recombinant; 5807.7 kDa) and hirudin (7033.6 Da).

Bioinformatics. Clinical data processing and overall survival estimation was performed by SPSS 13 edition. (SPSS Inc., Chicago, IL, USA) and using the Kaplan-Meier method. Proteomic data were analyzed with ProteinChip Software package version 3.1 (Bio-Rad Laboratories).

Initially we compared serum samples from CRC patients collected prior (T=0) to start of chemotherapy with matched NC. For this analysis all acquired spectra were compiled and analyzed as a whole. Next we used the T=0 proteomic profiles from all evaluable CRC patients and analyzed in accordance with chemotherapy response. For these analyses all spectra were baseline subtracted and normalized to the total ion current between 1.500 and 200.00 Da. Peak differences between CRC responders and non-responders and NC were calculated with the Biomarker Wizard (BMW) software application, comparing intensities of all detected peaks with non-parametric statistical tests. Peak clustering settings were: signal-to-noise ratio ≥5, appearance in ≥30% of spectra, second pass signal-to-noise ratio of ≥2, and a 0.5% cluster mass window. P-values <0.01 were considered statistically significant (10,12-14). Based on the quality control sample the mean coefficient of variation (CV) over all analyses was calculated for each significant peak. We assumed that the identified potential biomarkers of CRC, peaks with *m/z* values that differentiate between CRC patients and NC, were most likely to change in intensity according to response to chemotherapy, acting as predictive biomarkers.

Finally, we analyzed serum samples, which were serially collected throughout the chemotherapy treatment, investigating changes in proteomic spectra during chemotherapy.

For pre-processing the spectra were re-sampled to a common m/z vector and the baseline was corrected using the PROcess R-package. Furthermore, the intensity of the spectra was normalised to the total ion current to reduce noise variance between replicate measurements (15). To correct for small deviations in the m/z values due to the calibration the alignment algorithm by Jeffries was implemented in tbimass and applied (16). For classification the support vector machine implementation within the MCRestimate R-package was applied. For variable selection a variable filtering procedure based on the relative intensity variance was used for classification. To assess the classification accuracy a 10-fold repetition of 10-fold cross validation with a nested 3-fold parameter optimisation loop was conducted. The number of variables used for classification was reduced in each classification by recursive feature elimination (17).

Peaks that were identified in less than four patients of the serial analysis were judged as not to be candidate biomarkers of response and therefore excluded from further analysis. Furthermore, we included peaks with *m/z* values previously shown to correlate with malignancies (10,18-20) and all peaks that were prominently present in 4 or more spectra of each individual patient. Changes in peak intensity of identified peaks (*m/z* values) were followed over time throughout chemotherapy. The peak selection was independently reviewed by two of the authors (H.H. Helgason and J.Y.M.N. Engwegen). Response assessments were according to protocol less frequent than blood sampling. Therefore, we correlated response to chemotherapy with proteomic profile obtained in serum collected at similar time point. Assuming that possibly predictive profiles would be different in responding patients compared with not responding patients we excluded patients with stable disease as best response. Serum collections from patients that had been assessed as having PD according to RECIST were classified as PD observations.

For each peak a linear mixed effects model (using the 'lme' package in S-plus (S-plus, v. 6.2 professional; Insightful Inc., Seattle, WA, USA) was constructed with the log transform of peak intensity as the response variable. To account for the skewedness of the peak intensity data log transforms were applied prior to the analysis. Time (weeks), disease state, chemotherapy and the interaction term between week and disease state, were taken as fixed effect variables. Patient and its interaction with time (week) were taken as random effect variables. The Wald F-test (21) was used to select significant fixed variables. A relationship was considered significant if in the final model either disease state or its interaction with time (week) had a p-value <0.05. A sensitivity analysis was performed by repeating the analysis on the subset of the data in which the PD observations that lie between PR and PD assessments were excluded.

Due to the explorative nature of the analysis we consider the analysis as a hypothesis creative and consequently do not adjust the final p-values for multiple testing. Instead, we interpret the p-values as the strength of evidence for each particular relationship by calculating the lower bound of the type I error probability conditional on each observed p-value (22).

Results

Clinical outcome. A total of 63 patients with advanced colorectal cancer were screened. Fifty-five patients were eligible for study objectives. The reason for exclusion were previous (neo-) adjuvant chemotherapy (n=4), chemoradiotherapy (n=2), previous liver perfusion (n=1) or prior hyperthermal intra-peritoneal chemotherapy (n=1). In addition 13 patients were excluded from proteomic analysis because of protocol violation regarding the baseline serum sample collection. In total 42 patients with mean age of 57 years (range, 37-73) were treated with first-line oxaliplatin and

Table I. Patient characteristics.

Characteristic	Patients	%
Number	42	100
Age (years)		
Mean (range)	57 (37-73)	
Gender		
Male	26	62
Female	16	38
Prior treatment		
Chemo-radiotherapy (>6 months previously)	8	19
Sites of metastases		
Liver	30	71
Abdominal cavity	10	24
Lungs	15	36
Lymph nodes	14	33
Abdominal wall	2	5

Table II. Objective response rate in all evaluable patients.

Response	Patients	%
Partial response	20	50
Stable disease	10	25
Progressive disease	10	25

capecitabine chemotherapy in a 3-week cycle according to protocol. Further clinical characteristics are shown in Table I. Twenty-six patients (62%) had two or more metastatic sites. No patient had only peritoneal involvement. The mean follow-up was 13.5 months and the 42 patients completed in total 233 (mean 6; range 1-10) chemotherapy cycles. Two patients were not assessable for response, one because of localized stage III disease and one because of clinical progression and deterioration after only one cycle. There were no complete responses, 20 patients developed a partial response (response rate: 50%) and 10 patients had stable disease. The maximal response was usually seen after 2-4 cycles of chemotherapy followed by stabilization. In case of a favorable tumor response and reasonable tolerability, it is common practice at our institute to maximize treatment by 6-8 cycles followed by watchful waiting. Ten patients were chemotherapy resistant from the start of treatment and showed progressive disease at first evaluation. Median overall survival was 19.5 months (95% CI: 16-23) (Table II). The median time to progression was 6.2 months (95% CI: 5.6-6.7), median progression-free survival was 6 months 95% CI: 5.4-6.5) and progression-free survival was 10.8 months (95% CI: 9.5-12.1), respectively. In case of progression the most common metastatic sites were the liver and lung.

Proteomic profiling of colorectal cancer patients and normal controls. Serum obtained immediately prior to start of chemotherapy in 40 eligible and evaluable patients with advanced CRC was analyzed by SELDI-TOF MS and compared to the proteomic profile of 40 NC who were matched for age, gender and time period of serum collection. By using the protein chip analysis and applying the BMW software application we

Table III. Peaks and their tentative identities, significantly discriminating between colorectal cancer patients (CRC) and normal control subjects (NC) in order of significance.

M/z value	Mean intensity CRC	Mean intensity NC	P-value	Mean CV in quality-control sample (%)	Identification
14060	2.09	3.10	<0.0001	39.1	Glutathionylated transthyretin (23)
28100	2.99	4.16	<0.0001	41.1	Apolipoprotein A-I (10)
15948	10.30	5.83	<0.001	30.8	Haptoglobin α-2 or hemoglobin α SPA adduct (24)
31948	0.39	0.14	<0.001	63.0	Tissue factor pathway inhibitor
8078	4.45	2.27	<0.001	27.1	
16329	2.05	1.04	<0.001	21.6	
16140	5.44	3.00	<0.001	19.3	
7970	9.46	5.22	<0.001	23.4	
5912	5.05	3.17	<0.01	43.2	Fibrinogen α-E chain fragment
8037	1.70	1.13	<0.01	18.9	
2022	2.18	3.32	<0.01	67.3	Hemoglobuline α-chain fragment (36)
12861	0.02	0.03	<0.01	24.6	Transthyretin fragment (37)
2129	1.26	2.39	<0.01	78.5	

CV, coefficient of variation.

. *M/z* values that differentiated between responding and non-responding patients, possibly predicting chemotherapy

M/z	P-value	Mean peak intensity PD	Mean peak intensity PR	Mean CV in quality-control sample (%)
3330	0.007	1.86	4.83	42.5
2756	0.016	2.47	4.64	65.5
6847	0.018	3.99	6.12	23.7
3893	0.023	3.02	4.80	74.9
3978	0.026	2.21	5.30	55.8
2055	0.030	2.39	0.38	41.9

detected 38 peaks that significantly differentiated between CRC patients and NC. From these we excluded low mass-to-charge ratios (*m/z* values <2 kDa) as they cannot be reliably attributed to real proteins due to interference by the sinapinic acid (SPA) matrix up to this *m/z*. Furthermore, peak integrity of all significant peaks was checked visually and peaks not visually discernable from noise were excluded (14). This approach resulted in the identification of 13 proteins that were significantly differentially expressed in colorectal cancer patients and matched NC (Table III). The two proteins, most significantly differentially expressed, had *m/z* values of 14060 Da, previously identified as glutathionylated transthyretin (23) and 28100 Da which was previously detected by our group as a potential biomarker of CRC (10). This protein was identified as apolipoprotein A-I (10). Furthermore, the *m/z* mass 15948 Da corresponds to that of the haptoglobin α-2 chain (24) and the *m/z* 16140 and 16329 Da are possibly SPA adducts of this protein [+ n x 206 Da] (25). The other 8 candidate biomarkers for CRC, their tentative identities and the mean coefficient of variation (CV) of these peaks, in the quality-control samples, are shown in Table III. Although all these peaks could be identified as significantly different, some of them (*m/z* 31.9, 2.0 and 2.1 kDa) showed very high technical variability as shown by their mean CV. Therefore, the results for these potential biomarkers should be interpreted with caution.

Proteomic profile and response prediction. Proteomic profiles of serum obtained from CRC patients immediately prior to start of chemotherapy were analyzed according to response. Response evaluation was determined prior to start and after every second cycle, according to protocol, in 40 patients eligible for response evaluation according to RECIST criteria. Patients were divided in two groups according to maximal response and survival. Responders were 20 patients with a partial response and a mean overall survival 20 months and non-responders 10 patients with primary progressive disease and a mean overall survival of 10 months ($p<0.05$). In the pre-processing normalization procedure 1 serum sample from the CRC non-responders was categorized as outlier and excluded from further analysis (26). By applying Mann-Whitney U test a positive correlation was observed between 6 proteomic peaks. One of these proteins with a *m/z* value of 3330 Da reached significance at $p<0.01$ (Table IV), serving as a potential predictive biomarker for chemotherapy response in CRC patients.

Proteomic profile changes according to response. Finally, we selected patients of whom serial blood samples had been collected during chemotherapy for identification of peptides that significantly changed during treatment. Patients were selected according to response to chemotherapy, number of blood sample collections (≥4 during chemotherapy) and time of blood sample collection, which was predefined at immediately prior to chemotherapy (<2 weeks), 6, 12 and 18 (±2) weeks, respectively. Fifteen patients (11 with PR and 4 with PD as best response, respectively) fulfilled all criteria and were analyzed. Serum samples from 7 patients of the 11 responsive patients were collected during subsequent progression. Therefore, we were able to build two sets of eleven serial blood sample collections during PR and PD disease state. The patients had received 3-9 (mean 6) chemotherapy cycles and had given 4-12 (mean 8) blood samples at different time points. By proteomic profile analysis of these serum samples we were able to identify a number of peaks that changed prominently during treatment and according to response (Fig. 1). Two of the peaks, *m/z* 2022 and 28100 Da, showed a significant relationship between presence of CRC and peak intensity as analyzed by the previously described model. As previously described a linear mixed effect (lme) model was constructed for each peak with peak intensity as response variable. For peak 2022 the intensity was the only variable found to change significantly ($p<0.05$) according to CRC, response or non-response. The p-value resulting from dropping CRC from the lme model was 0.014, which corresponds to a lower bound on the conditional type I error probability (CEP) of 0.12. An estimate for the mean intensities (and 95% C.I.s) for patients with PR was 7.5 (95% CI: 6.1-9.2), and for patients with PD the estimate was 5.3 (95% CI: 4.2-6.8) as shown in Fig. 2. During the treatment period the interaction term between disease state and time (week) were determined to be significant for peak 28100 Da. The effect of removing disease state from this model was $p=0.008$, which has a corresponding CEP of at least 0.10. The effect of chemotherapy was to decrease the mean intensities by 25% (95% CI: 2%, 53%). There was no change over time of the mean intensities of patients with PD, however, patients with PR experienced a drop in mean intensities of 2% per week (95% CI: 0.2%, 4%) (Fig. 3).

While the results for peaks 2022 and 28100 show promising correlations between changes in peak intensities and disease state, response or non-response, it must be noted that the conditional type I error rates for these results were

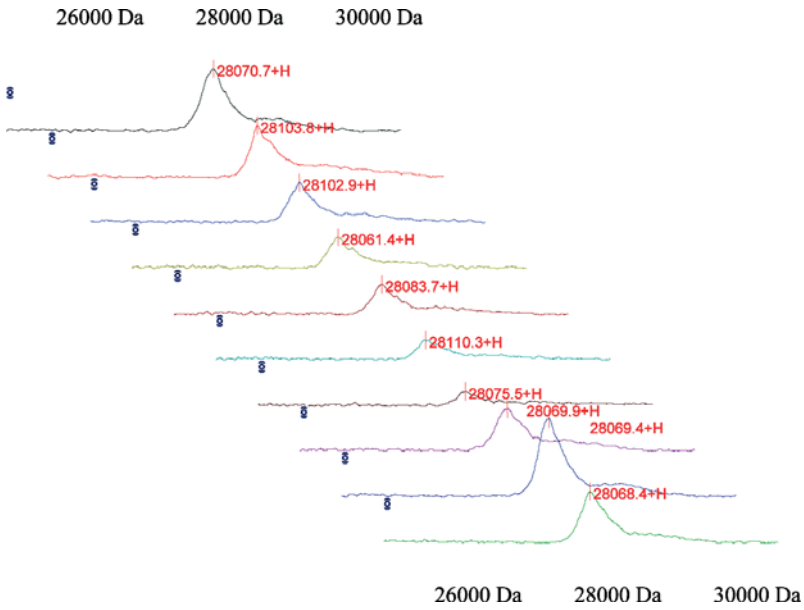

Figure 1. Intensity change of peak 28100 during chemotherapy in one of the patients who developed a partial remission (PR) followed later by progressive disease (PD) [PR at week 12, followed by chemotherapy rest at week 20 (6th cycle) and PD at 39 weeks].

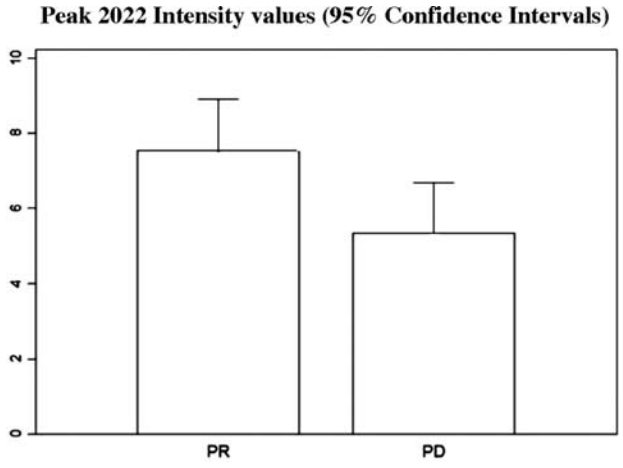

Figure 2. Intensity estimates and confidence intervals of peak 2022 in serial analysis of CRC patients who developed a partial remission (PR), or progressive disease (PD). All values are transformed back into the original units.

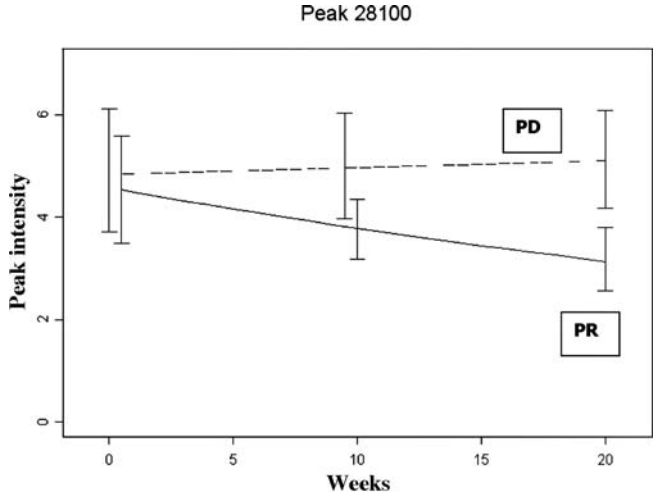

Figure 3. Mean peak intensity of m/z 28100 for patients with respectively partial remission (PR) and progressive disease (PD) (PR solid line and PD dashed line) during chemotherapy treatment. Confidence intervals (95%) are presented for weeks 0, 10 and 20. All values are transformed back into the original units.

both at least 10%. Given that the accepted upper bound of type I errors is 5%, further research must be conducted to confirm these results.

Discussion

In this single institutional phase II study we describe detection of significantly different proteomic patterns in CRC patients vs. NC subjects serving as potential biomarkers of colorectal cancer. Protein analysis with SELDI-TOF mass spectrometry of serum from cancer patients is a promising tool for the detection and identification of novel proteins or proteomic profiles that are disease specific. Individual proteins could possibly serve as specific biomarkers for disease detection (screening), prognosis (prognostic) or treatment response (predictive). In our study we found that, in prospectively collected serum samples, 13 proteins were differently expressed in advanced CRC patients compared to NC subjects. Of these, we identified two proteins, with *m/z* value of 14060 and 28100 Da that significantly ($p<0.0001$) differentiated between CRC patients and NC acting as candidate biomarkers of CRC. Subsequent analysis identified *m/z* value of 3330 kDa as a potential predictive marker of therapeutic response ($p<0.01$) and two other proteins, with *m/z* value of 2022 and 28100 Da, as potential biomarkers for therapeutic monitoring.

in with *m/z* value of 14060 Da, discriminating CRC and NC subject, has previously been identified as glutathionylated transthyretin (23,27,28). Zhang *et al* (29) has identified transthyretin as a potential biomarker of ovarian cancer. In their validation set of 20 CRC patients transthyretin was shown to be significantly less abundant, than in healthy controls, albeit to a lesser degree than in ovarian cancer patients. Fung *et al* (13) suggested that measuring different forms of serum transthyretin resulted in higher diagnostic accuracy than measuring total transthyretin alone. Furthermore, they stated that the lower level of total transthyretin seen in CRC patients is due to down-regulation of truncated and unmodified forms of transthyretin, but that the cysteinylated and glutathionylated forms are only significantly decreased in ovarian cancer. Our results, showing reduced expression of glutathionylated transthyretin in CRC patients in comparison to NC subjects, are thus in accordance with this theory. Moore e*t al* (30) confirmed later the role of transthyretin, in combination with apolipoprotein A-I and CA-125, as potential biomarker of ovarian cancer (29,30).

The second protein, with *m/z* 28.1 kDa, was previously identified by us and others as apolipoprotein A-I (Apo A-I) and that this protein differentiated between CRC patients and NC subjects (10). The serum level of Apo A-I was significantly lower in CRC patients than in NC suggesting lower activity of the protease necessary for its production. Apolipoprotein A-I is synthesised as a pre-pro-peptide of which the mature peptide is generated by N-terminal cleavage of six amino acids. The responsible enzyme is metal-dependent and insensitive to serine-protease inhibitors (31). The expression of Apo A-I has been shown to be decreased in several malignancies including cholangiocarcinoma (32) and ovarian cancer (29,33,34). Furthermore, increased expression of Apo A-I has been observed in liver metastases as well as, although to a lesser extent, in primary tumors of colorectal origin (35). The precise role of Apo A-I in colorectal cancer has to be determined but, taken into account the lower expression in other malignancies (29,30), we acknowledge that it is not disease specific and thus unlikely to be a selective biomarker with high specificity for colorectal cancer.

Among the other proteins that significantly differentiated between CRC patients and NC subjects, the 32 kDa protein ($p<0.001$) is of greatest interest. This *m/z* value belongs most likely to tissue factor pathway inhibitor 2 (TFPI-2) with molecular mass of 32 kDa (24) although not confirmed in our study. This protein had a low intensity but with serum level significantly higher in CRC patients compared with NC subjects.

Our second aim was to identify predictive biomarkers of response by comparing serum obtained prior to chemotherapy, according to response, in evaluable patients. Six candidate predictive biomarkers were identified. Of these only the *m/z* value of 3.3 kDa was significantly differently expressed in patients that developed a partial response (PR) in comparison to patients that were progressive (PD) at first evaluation ($p<0.01$). Serum levels of the 3.3 kDa protein at baseline were significantly lower in patients with PD compared with patients that developed PR (Table IV), which were again lower than in NC subjects. This correlates with our previous analysis in which the 3.3 kDa protein was markedly reduced in CRC patients compared with NC subjects (10), making the 3.3 kDa protein another important discriminating protein between CRC patients and NC. Furthermore, in our previous analysis the 3.3 kDa peak value was highly correlated with apolipoprotein C-I, which has a theoretical mass of 6630.58 Da (10). The 3.3-kDa peak is considered a double charged artefact of the 6.6-kDa protein although not confirmed. These results should be interpreted with caution because of the limited number of patients analysed.

Searching for potential biomarkers of therapy monitoring we discovered two proteins with, *m/z* 2.0 and 28.1 kDa (Apo A-1), that had different expression according to chemotherapy response by serial analysis of serum obtained during chemotherapy. The decline in the intensity of both proteins was more pronounced in responding patients compared with non-responding patients. This may indicate that the change in serum levels of these proteins correlated with response, which could be useful to monitor antitumor activity of chemotherapy (Figs. 1-3). The 2.0 kDa protein is probably a hemoglobulin α-chain fragment (amino acids 110-128, NH2-AAHLPAEFTPAVHASLDKF-COOH) although this was demonstrated in analysis of cervicovaginal fluid (36). Considering the large variability of this protein, as shown by the high CV, its usefulness to monitor chemotherapy response should be interpreted cautiously. Apo A-I was found to discriminate between CRC and NC with higher serum levels in NC, but it was not predictive of chemotherapy response in samples taken prior to start chemotherapy. Declining levels of the acute phase protein Apo A-I coinciding with treatment response might reflect a decrease in systemic inflammation associated with advanced CRC and not directly related to shrinkage in tumor volume. Regardless of this the expression of apolipoprotein A-I might play a future role in therapy monitoring of first-line oxaliplatin-capecitabine chemotherapy in patients with advanced colorectal cancer.

In this study, we were able to reproduce our previously identified biomarkers for CRC patients in an independent prospective data set, which supports the use of the applied CM 10 chip and our previously described material selection and analysis methods. In addition, our previously described collection and storage at -30˚C does not appear to cause significant degradation of the proteome profile for up to 2 years. Although not conclusive there were some differences seen in responding and non-responding patients.

Study limitations are the relatively low number of patients which was caused by the introduction of bevacizumab to the standard first-line chemotherapy for CRC patients. Although we were able to reproduce our previous results and methodology the current study should be repeated in a larger cohort of patients to confirm the obtained results. Our ability to reproduce the previously detected peak at *m/z* 28.1 kDa supports the applied experimental conditions for proteomic analysis in patients with advanced colorectal cancer.

In conclusion, we identified 13 proteins that significantly differentiated between CRC and normal subjects of which two proteins, with *m/z* values of 14060 and 28100 Da respectively, may serve as candidate biomarkers of CRC but previously we identified the 28.1 kDa protein as Apo A-I, a candidate biomarker of CRC.

Acknowledgements

We would like to thank Dr J.M.G. Bonfrer and Ms. Tiny Korse for sample collection and storage assistance.

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J and Thun MJ: Cancer statistics 2009. CA Cancer J Clin 59: 225-249, 2009.
2. Fearon ER and Vogelstein B: A genetic model for colorectal tumorigenesis. Cell 61: 759-767, 1990.
3. Young GP and Le Leu RK: Preventing cancer: dietary lifestyle or clinical intervention? Asia Pac J Clin Nutr 11 (Suppl. 3): S618-S631, 2002.
4. Young GP, St John DJ, Winawer SJ and Rozen P; WHO and OMED: Choice of fecal occult blood tests for colorectal cancer screening: recommendations based on performance characteristics in population studies: a WHO (World Health Organization) and OMED (World Organization for Digestive Endoscopy) report. Am J Gastroenterol 97: 2499-2507, 2002.
5. Cotterchio M, Manno M, Klar N, McLaughlin J and Gallinger S: Colorectal screening is associated with reduced colorectal cancer risk: a case-control study within the population-based Ontario Familial Colorectal Cancer Registry. Cancer Causes Control 16: 865-875, 2005.
6. Winawer SJ and Zauber AG: The advanced adenoma as the primary target of screening. Gastrointest Endosc Clin N Am 12: 1-9, v, 2002.
7. Segnan N, Senore C, Andreoni B, Aste H, Bonelli L, Crosta C, Ferraris R, Gasperoni S, Penna A, Risio M, Rossini FP, Sciallero S, Zappa M, Atkin WS; SCORE Working Group Italy: Baseline findings of the Italian multicenter randomized controlled trial of 'once-only sigmoidoscopy' SCORE. J Natl Cancer Inst 94: 1763-1772, 2002.
8. Dong SM, Traverso G, Johnson C, Geng L, Favis R, Boynton K, Hibi K, Goodman SN, D'Allession M, Paty P, Hamilton SR, Sidransky D, Barany F, Levin B, Shuber A, Kinzler KW, Vogelstein B and Jen J: Detecting colorectal cancer in stool with the use of multiple genetic targets. J Natl Cancer Inst 93: 858-865, 2001.
9. Mulhall BP, Veerappan GR and Jackson JL: Meta-analysis: computed tomographic colonography. Ann Intern Med 142: 635-650, 2005.
10. Engwegen JY, Helgason HH, Cats A, Harris N, Bonfrer JM, Schellens JH and Beijnen JH: Identification of serum proteins discriminating colorectal cancer patients and healthy controls using surface-enhanced laser desorption ionisation-time of flight mass spectrometry. World J Gastroenterol 12: 1536-1544, 2006.
11. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabbeke M, Van Oosterom AT, Christian MC and Gwyther SG: New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst 92: 205-216, 2000.
12. Adam BL, Qu Y, Davis JW, Ward MD, Clements MA, Cazares LH, Semmes OJ, Schellhammer PF, Yasui Y, Feng Z and Wright GL Jr: Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. Cancer Res 62: 3609-3614, 2002.
13. Fung ET, Yip TT, Lomas L, Wang Z, Yip C, Meng XY, Lin S, Zhang F, Zhang Z, Chan DW and Weinberger SR: Classification of cancer types by measuring variants of host response proteins using SELDI serum assays. Int J Cancer 115: 783-789, 2005.
14. Petricoin EF and Liotta LA: SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. Curr Opin Biotechnol 15: 24-30, 2004.
15. Meuleman W, Engwegen JY, Gast MC, Beijnen JH, Reinders MJ and Wessels LF: Comparison of normalisation methods for surface-enhanced laser desorption and ionisation (SELDI) time-of-flight (TOF) mass spectrometry data. BMC Bioinformatics 9: 88, 2008.
16. Jeffries N: Algorithms for alignment of mass spectrometry proteomic data. Bioinformatics 21: 3066-3073, 2005.
17. Guyon I, Weston J and Barnhill S: Gene selection for cancer classification using support vector machines. Machine Learning 46: 389-422, 2002.
18. Li J, Zhang Z, Rosenzweig J, Wang YY and Chan DW: Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clin Chem 48: 1296-1304, 2002.
19. Ward DG, Cheng Y, N'kontchou G, Thar TT, Barget N, Wei W, Martin A, Beaugrand M and Johnson PJ: Preclinical and post-treatment changes in the HCC-associated serum proteome. Br J Cancer 95: 1379-1383, 2006.
20. Ward DG, Suggett N, Cheng Y, Wei W, Johnson H, Billingham LJ, Ismail T, Wakelam MJ, Johnson PJ and Martin A: Identification of serum biomarkers for colon cancer by proteomic analysis. Br J Cancer 94: 1898-1905, 2006.
21. Greenland S: Tests for interaction in epidemiologic studies: a review and a study of power. Stat Med 2: 243-251, 1983.
22. Selke BaB: Calibration of p-values for testing precise null hypotheses. Am Stat 55: 61-72, 2001.
23. Schweigert FJ, Wirth K and Raila J: Characterization of the microheterogeneity of transthyretin in plasma and urine using SELDI-TOF-MS immunoassay. Proteome Sci 2: 5, 2004.
24. ExPAsy Proteomics Server; entry: P00738, 2006.
25. Dijkstra M, Vonk RJ and Jansen RC: SELDI-TOF mass spectra: A view on sources of variation. J Chromatogr B Analyt Technol Biomed Life Sci 847: 12-23, 2006.
26. Engwegen JY, Gast MC, Schellens JH and Beijnen JH: Clinical proteomics: searching for better tumour markers with SELDI-TOF mass spectrometry. Trends Pharmacol Sci 27: 251-259, 2006.
27. Kishikawa M, Nakanishi T, Miyazaki A and Shimizu A: Enhanced amyloidogenicity of sulfonated transthyretin in vitro, a hypothetical etiology of senile amyloidosis. Amyloid 6: 183-186, 1999.
28. Theberge R, Connors L, Skinner M, Skare J and Costello CE: Characterization of transthyretin mutants from serum using immunoprecipitation, HPLC/electrospray ionization and matrix-assisted laser desorption/ionization mass spectrometry. Anal Chem 71: 452-459, 1999.
29. Zhang Z, Bast RC Jr, Yu Y, Li J, Sokoll LJ, Rai AJ, Rosenzweig JM, Cameron B, Wang YY, Meng XY, Berchuck A, Van Haaften-Day C, Hacker NF, De Bruijn HW, van der Zee AG, Jacobs IJ, Fung ET and Chan DW: Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. Cancer Res 64: 5882-5890, 2004.
30. Moore LE, Fung ET, McGuire M, Rabkin CC, Molinaro A, Wang Z, Zhang F, Wang J, Yip C, Meng XY and Pfeiffer RM: Evaluation of apolipoprotein A1 and posttranslationally modified forms of transthyretin as biomarkers for ovarian cancer detection in an independent study population. Cancer Epidemiol Biomarkers Prev 15: 1641-1646, 2006.
31. Scanu AM, Edelstein C and Gordon JI: Apolipoproteins of human plasma high density lipoproteins. Biology, biochemistry and clinical significance. Clin Physiol Biochem 2: 111-122, 1984.
32. Wang X, Dai S, Zhang Z, Liu L, Wang J, Xiao X, He D and Liu B: Characterization of apolipoprotein A-I as a potential biomarker for cholangiocarcinoma. Eur J Cancer Care 18: 625-635, 2009.
33. Wang Z, Yip C, Ying Y, Wang J, Meng XY, Lomas L, Yip TT and Fung ET: Mass spectrometric analysis of protein markers for ovarian cancer. Clin Chem 50: 1939-1942, 2004.
34. Nosov V, Su F, Amneus M, Birrer M, Robins T, Kotlerman J, Reddy S and Farias-Eisner R: Validation of serum biomarkers for detection of early-stage ovarian cancer. Am J Obstet Gynecol 200: 639-e1-5, 2009.
35. Tachibana M, Ohkura Y, Kobayashi Y, Sakamoto H, Tanaka Y, Watanabe J, Amikura K, Nishimura Y and Akagi K: Expression of apolipoprotein A1 in colonic adenocarcinoma. Anticancer Res 23: 4161-4167, 2003.
36. Brown AG, Leite RS, Engler AJ, Discher DE and Strauss JF III: A hemoglobin fragment found in cervicovaginal fluid from women in labor potentiates the action of agents that promote contraction of smooth muscle cells. Peptides 27: 1794-1800, 2006.
37. Kozak KR, Su F, Whitelegge JP, Faull K, Reddy S and Farias-Eisner R: Characterization of serum biomarkers for detection of early stage ovarian cancer. Proteomics 5: 4589-4596, 2005.