

Plasma proteome predicts chemotherapy response in osteosarcoma patients

YITING LI¹, TU ANH DANG³, JIANHE SHEN¹, JOHN HICKS², MURALI CHINTAGUMPALA¹,
CHING C. LAU¹ and TSZ-KWONG MAN¹

¹Texas Children's Cancer Center, Texas Children's Hospital, and Department of Pediatrics,

²Department of Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030;

³Cell Biosciences, Inc., 3040 Oakmead Village Drive, Santa Clara, CA 95051, USA

Received September 3, 2010; Accepted November 1, 2010

DOI: 10.3892/or.2010.1111

Abstract. Osteosarcoma is the most common malignant bone tumor that affects hundreds of children and young adults every year. The major prognostic factor in patients with localized osteosarcoma is the development of resistance towards pre-operative chemotherapy. However, modifications of post-operative chemotherapy based on the histological response have not significantly improved the outcome of patients. Thus, it would be of tremendous clinical value if the poor responders could be identified at the time of diagnosis, so that ineffective therapy can be prevented and intensified or alternative therapy could be provided to improve their outcome. We hypothesized that plasma proteomic profiles could be used to distinguish good from poor responders prior to the start of treatment. In order to test this hypothesis, we analyzed the proteomic profiles in two sets of plasma samples (n=54) from osteosarcoma patients collected before (n=27) and after (n=27) pre-operative chemotherapy. Using a linear support vector machine algorithm and external leave-one-out cross validation, we developed two classifiers that classified good and poor responders with an equal accuracy of 85% (p<0.01 after 5000 permutations) in both sets of plasma samples. In order to understand the biological basis of the classifiers, we further identified and validated two plasma proteins, serum amyloid protein A and transthyretin, in the classifiers. Our results suggest that plasma proteomic profiles can predict chemotherapy response before treatment as accurately as after

treatment. Our study could lead to the development of a simple blood test that can predict chemotherapy response in osteosarcoma patients. Since the two identified proteins are involved in innate immunity, our findings are corroborated by the notion that boosting the innate immunity in conjunction with chemotherapy, achieves a better anti-tumor activity, thus improving the overall survival of osteosarcoma patients.

Introduction

Osteosarcoma is a primary malignant bone tumor arising from primitive bone-forming mesenchymal cells and is characterized by the production of osteoid material (1). It is a common malignant bone tumor occurring in children and accounts for ~60% of malignant bone tumors which occur during the first two decades of life (2). Since the 1970s, the accepted standard of care for osteosarcoma involves four steps: Diagnosis by an initial biopsy, pre-operative chemotherapy, definitive surgery to resect the tumor, and post-operative chemotherapy (Fig. 1). Besides facilitating limb-sparing procedures by shrinking the tumor, pre-operative chemotherapy can theoretically be used as an *in vivo* indicator of the chemosensitivity of an individual tumor, thus providing a guide for the customization of post-operative chemotherapy. However, the strategy of pre-operative chemotherapy has not improved disease-free survival for patients with osteosarcoma. Despite a number of multi-institutional clinical trials of pre-operative chemotherapy in both Europe and North America, the overall survival rate has remained unchanged over the past 30 years (3).

One recognized approach for improving the outcome is to develop new and more effective therapeutic agents. However, at the present time, there are still no effective treatments against chemoresistant osteosarcoma. Development of risk-based stratification of patients will help future clinical trials to identify effective drugs for the treatment of poor responders. Another approach is to tailor existing therapy based on predictive factors. This could improve survival in patients who are predicted to have resistant disease by intensifying the use of the known effective agents. It could also allow for the reduction of dose-related toxicity while maintaining good survival rates for patients who are predicted to have a good prognosis. Both approaches are based on our ability to stratify

Correspondence to: Dr Tsz-Kwong Man, Texas Children's Cancer Center, Texas Children's Hospital, Dan L. Duncan Cancer Center, and Department of Pediatrics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
E-mail: ctman@txccc.org

Abbreviations: SAA, serum amyloid protein A; TTR, transthyretin; SELDI-TOF MS, surface-enhanced laser desorption/ionization-time of flight mass spectrometry; LOOCV, leave-one-out cross validation

Key words: osteosarcoma, surface-enhanced laser desorption/ionization, plasma proteomics, chemotherapy response, serum amyloid protein A, transthyretin

patients with either a good or poor response to pre-operative chemotherapy before the treatment is given, i.e. at the time of initial diagnosis.

In order to address this clinically important problem, we previously identified a chemoresistant signature that can predict the response of chemotherapy by using RNA expression profiling, suggesting that 'omics' approaches can be used to detect poor-responding osteosarcoma patients (4). However, the RNA expression profiling of tumor samples is not always possible due to a number of factors, such as RNA degradation, availability of tumor biopsies, and the involvement of invasive procedures for tumor collection. In this study, we used an alternative approach for identifying poor responders, known as surface-enhanced laser desorption/ionization-time of flight mass spectrometry (SELDI-TOF MS), which is a rapid and high-throughput proteomic technique for the discovery of biomarkers in various cancers (5,6). We previously demonstrated that a plasma-based proteomic signature generated by SELDI-TOF MS can be used to distinguish osteosarcoma from benign osteochondroma patients (7). However, the problem of classifying good and poor responders was not addressed in that study. Thus, we performed additional analyses of the previously generated plasma profiles of the 27 plasma samples from osteosarcoma patients collected at the time of diagnosis (pre-treatment) and generated plasma profiles of 27 additional plasma samples collected at the time of definitive surgery (post-treatment), that were not previously used before. Two support vector machine (SVM) classifiers were developed to test whether poor responders could be detected using the pre- or post-treatment samples. We also identified and validated two proteins in the classifiers that are related to the host response.

Materials and methods

Patients and samples. A total of 54 plasma specimens from 36 osteosarcoma patients were used in this study (Table I). All specimens were collected through IRB-approved protocols from three collaborative institutions, Texas Children's Hospital, Cook Children's Hospital in Forth Worth, Texas, and Oklahoma Children's Hospital, after informed consent was obtained from all patients. Plasma samples were collected in EDTA-containing tubes and centrifuged at 1,000 rpm for 10 min. The plasma supernatant was collected and divided into aliquots and stored at -80°C until use. As is shown in Fig. 1, 27 plasma samples were collected at the time of initial diagnosis (pre-treatment) and these were used as part of our previous study to distinguish osteosarcoma from benign osteochondroma (7). After diagnosis, all the osteosarcoma patients who had enrolled in this study received three-drug (cisplatin, doxorubicin and high-dose methotrexate) pre-operative chemotherapy. Then, patients underwent definitive surgery, at which time the osteosarcoma tissue was resected and the extent of tumor necrosis was evaluated by a pathologist (J.H.). Patients with tumor necrosis of $<90\%$ were classified as poor responders. For the poor responders, the treatment regimen was intensified by using high-dose cytoxan/melphalan as well as stem cell rescue during post-operative chemotherapy. For the good responders (tumor necrosis of $\geq 90\%$), the same pre-operative regimen was used

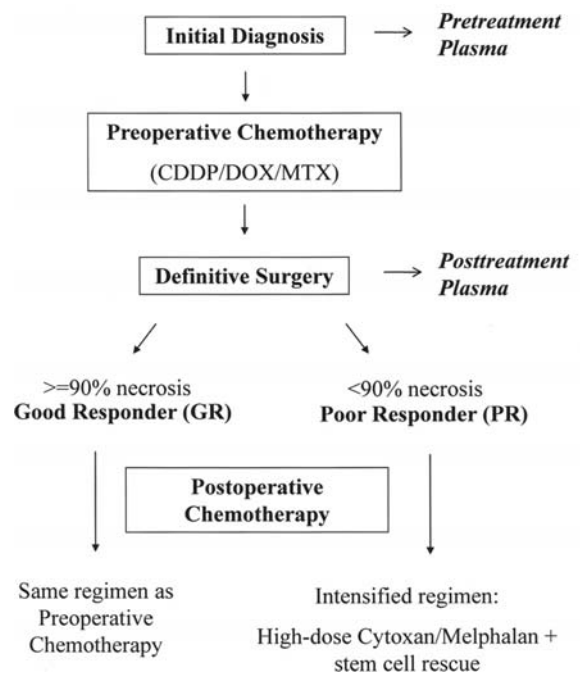


Figure 1. The schema of the treatment and sample collection protocol for this study. CDDP, cisplatin; DOX, doxorubicin; MTX, methotrexate.

for their post-operative chemotherapy. At the time of definitive surgery, 27 plasma samples were collected (post-treatment plasma). As shown in Table I, patients were divided into two groups, good and poor responders, with similar distributions of age, gender and primary tumor sites. All of the patients, except one from each of the pre- or post-treatment groups, were followed-up for at least 20 months. Their survival status is shown in Table I. Eighteen patients contributed both pre- and post-treatment plasma samples, which were used for paired analysis.

SELDI-TOF MS analysis. The SELDI-TOF MS analysis of the 27 pre-treatment plasma samples has been previously reported (7). For the post-treatment samples used in this study, a similar approach was carried out. In brief, 20 μl of each plasma sample were denatured and separated into six fractions by an anion-exchange Q HyperD F 96-well filter plate based on stepwise pH gradient (CIPHERGEN Biosystems, Fremont, CA; is now sold by Bio-Rad, Hercules, CA). The six fractions were either flow-through (fraction 1), pH 7 (fraction 2), pH 5 (fraction 3), pH 4 (fraction 4), pH 3 (fraction 5) or organic eluant fractions (fraction 6). The proteins in the fractions were then captured by the weak cation exchange (CM10) ProteinChip (CIPHERGEN). The fractionated samples were diluted in CM Low Stringency Buffer and sinapinic acid (CIPHERGEN) was added to facilitate the desorption and ionization of the proteins. Then the samples were randomly spotted in duplicates on chips in a 96-well format using the Biomek 2000 Robotic Station (Beckman Coulter, Fullerton, CA). The intensities of the protein peaks were measured by the Protein Biology System PBS IIc (CIPHERGEN) using three different laser intensities under the low, medium and high power settings, as previously reported (7). Mass calibration

Table I. Clinical information of the plasma samples and osteosarcoma patients used in this study.

Characteristics	Time of plasma collection			
	Pre-treatment (n=27)		Post-treatment (n=27)	
	GR	PR	GR	PR
Total number of samples	14	13	12	15
Age at diagnosis (years)				
Median (range)	13 (9-22)	13 (9-17)	13 (9-22)	13 (8-18)
≤10	2	2	2	2
11-17	10	11	9	12
≥18	2	0	1	1
Gender				
Male	9	8	9	8
Female	5	5	3	7
Primary site				
Extremities	13	13	12	15
Others	1 (palate)	0	0	0
Metastasis at diagnosis				
Yes	3	3	2	2
No	11	10	10	13
Survival				
Died from disease	3	7	4	8
Alive	10	6	8	6
Unknown	1 ^a	0	0	1 ^a

^aLost during the follow-up within 20 months after the initial diagnosis. GR, good responders; PR, poor responders.

was performed using the All-in-1 peptide and All-in-1 protein II molecular mass standards (CIPHERGEN).

Proteomic profiles were baseline subtracted and normalized based on total ion current using CIPHERGEN Express Software 3.0 (CIPHERGEN). Peaks were detected with the similar settings as previously reported (7). In this study, the m/z ranges for low-, medium- and high-molecular weight proteins were 2,000-10,000, 10,000-30,000 and 30,000-200,000, respectively. Peak intensities of duplicates from the same sample were averaged before the analysis.

Development of proteomic classifiers. The normalized intensity values of all protein peaks were log-2 transformed and imported into BRB ArrayTools 3.5.0 (8). Two criteria were used for the selection of informative protein peaks for classification: A univariate misclassification rate of <0.25 and a fold-change between the two responder groups of >1.2. A linear SVM algorithm was used to construct two separate multivariate classifiers for the classification of good and poor responders using the proteomic profiles generated from the pre- and post-treatment samples, respectively. SVM finds the most optimal hyperplane for the separation of data by projecting them into a high-dimensional space (9). The classification accuracy was measured by external leave-one-out cross validation (LOOCV), in which the feature selection and model building were performed at each cross validation

iteration to minimize the underestimation of classification error (10,11). Hierarchical clustering of the samples was performed using the informative protein peaks with Pearson's correlation and average linkage. Other statistical analyses were performed using SPSS 12.0 (SPSS Inc., Chicago, IL). The specificity, sensitivity, and positive and negative predictive values, were calculated as previously described (7).

Protein identification and validation. Four samples with a high-level expression of the biomarker were pooled as the positive control. Four samples from the same fraction of plasma with a low-level expression of the biomarker were also pooled as the negative control. The pooled samples were resolved by 4-20% Novex Tris-Glycine SDS-PAGE gel (Invitrogen, Carlsbad, CA). The bands of interest matching the expected sizes that were higher in the positive control and low in the negative control, were excised from the gel and were subjected to overnight tryptic digestion (proteomics sequencing grade, Sigma, St. Louis, MO). Subsequently, 3 μ l of the digestion mixture were analyzed on a reverse phase H4 ProteinChip Array with α -cyano-4-hydroxycinnamic acid (CIPHERGEN). Masses of the trypsin-digested peptides were then measured by PBS IIc. Background peaks were removed by referencing the spectrum of trypsin-digested gel without the protein. The spectrum was calibrated internally by using two protein standards: Human angiotensin I (1296.5 Da) and

Table II. The informative protein peaks used in the post-treatment classifier.

No.	Protein peak	Fraction	GR		PR		PR/GR Ratio
			Average	SD	Average	SD	
1	m/z 2,566	F1	1.62	2.58	0.63	2.25	0.39
2	m/z 2,868	F4	0.72	2.36	1.27	2.35	1.78
3	m/z 2,975	F1	0.78	1.73	1.05	1.17	1.36
4	m/z 3,159	F2	0.70	1.98	1.30	1.47	1.87
5	m/z 3,214	F2	1.17	3.22	0.52	2.9	0.45
6	m/z 3,391	F1	0.84	1.42	1.16	1.57	1.38
7	m/z 3,868	F5	0.61	2.69	1.30	1.79	2.15
8	m/z 4,303	F1	1.41	1.74	0.83	1.60	0.59
9	m/z 5,721	F1	1.26	2.74	0.77	2.29	0.61
10	m/z 5,826	F1	1.90	2.48	0.78	2.00	0.41
11	m/z 5,829	F4	1.59	1.97	0.82	1.48	0.52
12	m/z 5,851	F6	2.39	2.77	0.91	1.58	0.38
13	m/z 6,436	F3	0.78	1.31	1.05	1.82	1.35
14	m/z 6,438	F1	0.81	1.20	0.99	1.27	1.23
15	m/z 6,501	F1	0.80	1.38	1.08	1.19	1.34
16	m/z 6,661	F3	0.84	1.30	1.17	1.58	1.39
17	m/z 6,944	F2	0.78	1.57	1.10	1.30	1.41
18	m/z 8,587	F1	1.72	1.99	0.73	1.7	0.42
19	m/z 8,682	F1	1.29	1.75	0.83	1.49	0.64
20	m/z 8,907	F3	0.81	1.38	1.18	1.25	1.46
21	m/z 9,448	F6	0.90	1.40	1.09	1.32	1.21
22	m/z 9,683	F6	0.82	1.36	1.09	1.20	1.33
23	m/z 10,098	F2	0.88	1.65	1.27	1.44	1.44
24	m/z 10,735	F5	0.90	1.79	1.57	2.18	1.74
25	m/z 11,157	F1	1.31	3.68	0.53	3.67	0.40
26	m/z 11,419	F1	1.70	3.86	0.78	1.60	0.46
27	m/z 11,467 ^a	F6	3.88	5.06	0.97	1.84	0.25
28	m/z 11,472	F5	3.58	4.58	0.97	1.76	0.27
29	m/z 11,501	F4	5.27	7.55	0.62	3.08	0.12
30	m/z 11,530 ^a	F6	4.19	5.00	1.03	1.67	0.25
31	m/z 11,543	F5	2.38	5.62	0.81	1.64	0.34
32	m/z 11,565	F3	2.76	3.49	0.89	1.62	0.32
33	m/z 11,647	F5	3.62	4.77	0.98	2.06	0.27
34	m/z 11,680	F1	2.85	5.88	0.91	1.66	0.32
35	m/z 11,702	F4	7.00	7.82	1.08	1.96	0.15
36	m/z 11,713	F6	4.34	4.78	0.92	1.48	0.21
37	m/z 11,720	F5	2.47	3.95	0.74	1.71	0.30
38	m/z 11,757	F1	2.28	3.53	0.93	1.29	0.41
39	m/z 11,766	F3	1.82	2.25	0.82	1.44	0.45
40	m/z 11,897	F6	2.97	3.32	0.92	1.93	0.31
41	m/z 13,127	F1	0.71	1.68	1.06	1.32	1.50
42	m/z 13,256	F6	0.47	7.80	1.18	2.03	2.52
43	m/z 13,876	F1	0.70	2.02	1.09	1.29	1.55
44	m/z 15,296	F1	1.52	3.05	1.06	1.37	0.70
45	m/z 27,817	F1	1.32	1.83	0.94	1.55	0.71
46	m/z 30,193	F1	1.92	3.04	0.83	1.74	0.43
47	m/z 36,034	F6	1.38	2.06	0.84	1.95	0.61
48	m/z 38,720	F6	1.18	1.36	0.92	1.34	0.78
49	m/z 39,541	F6	1.54	1.86	0.76	1.85	0.49
50	m/z 39,915	F6	1.45	1.73	0.76	1.55	0.52

Table II. Continued.

No.	Protein peak	Fraction	GR		PR		PR/GR Ratio
			Average	SD	Average	SD	
51	m/z 55,474	F1	2.84	3.20	0.68	2.27	0.24
52	m/z 66,830	F2	0.72	2.17	1.11	1.3	1.53
53	m/z 67,447	F2	0.67	2.39	1.06	1.28	1.57
54	m/z 67,746	F2	0.67	2.42	1.04	1.26	1.54
55	m/z 67,822	F2	0.70	2.21	1.05	1.26	1.50
56	m/z 68,087	F2	0.71	2.34	1.06	1.26	1.50
57	m/z 69,631	F5	0.70	1.85	0.98	1.28	1.39
58	m/z 78,507	F4	1.44	2.58	0.70	2.02	0.49
59	m/z 78,522	F4	1.47	1.44	0.80	1.33	0.55
60	m/z 80,235	F5	1.38	1.93	0.71	1.92	0.52
61	m/z 116,619	F6	0.78	1.84	1.03	1.37	1.32
62	m/z 122,843	F6	0.74	1.62	1.18	2.69	1.60
63	m/z 133,362	F6	0.61	2.88	1.44	1.81	2.35
64	m/z 138,499	F2	0.82	1.66	0.99	1.29	1.21
65	m/z 196,316	F4	0.90	1.40	1.11	1.17	1.23

These peaks were selected based on a misclassification rate of <0.25 and a fold-change of >1.2. The peaks are arranged in an ascending order of their m/z values. SD, standard deviation; GR, good responders; PR, poor responders. *Peaks which were also identified in the pre-treatment classifier.

β -endorphin (3465.0 Da). Peptide masses were searched against the NCBI database using the search engine ProFound to identify the target protein (12).

Western blot analysis was performed to validate the differential abundance of the plasma proteins between good and poor responders. The protein concentration was measured using the DC Protein Assay (Bio-Rad). For validation, 20 μ g of total protein from each sample were separated in a 4-12% Bis-Tris SDS-PAGE gel (Invitrogen) and transferred to a PVDF membrane. The membrane was blocked with the Odyssey Blocking Buffer (LI-COR, Lincoln, NE) for 1 h at room temperature, and then incubated overnight at 4°C with mouse monoclonal anti-human serum amyloid A (SAA) antibody (Santa Cruz Biotechnology, Santa Cruz, CA) or rabbit anti-human transthyretin (TTR) antibody (Santa Cruz Biotechnology) at 1:200 dilution. The secondary antibody was IRDye 680 conjugated goat anti-mouse IgG or goat anti-rabbit IgG (LI-COR) at 1:2000 dilution. The signal was detected using the Odyssey Infrared Imaging System (LI-COR). The PVDF membrane was dried and stained with SimplyBlue Safe Stain (Invitrogen) after the protein detection. The staining of the albumin band (66 kDa) on the membrane was used as the loading control.

Results

The proteomic classifier of post-treatment plasma. In order to test whether plasma proteomic profiles can be used to detect which osteosarcoma patients would respond poorly to pre-operative chemotherapy, 54 plasma samples of osteosarcoma patients collected at the time of initial diagnosis (pre-treatment,

n=27) and during definitive surgery (post-treatment, n=27) were subjected to SELDI-TOF MS analysis using weak cationic protein arrays (CM10). All 783 protein peaks identified by SELDI-TOF MS from six different fractions and three laser power settings were combined to construct a multivariate classifier as previously described (7). As the histological response is measured clinically at the time of definitive surgery, we first tested whether we could identify a proteomic signature of chemotherapy response using the post-treatment plasma samples collected during definitive surgery. A good responder was defined as a patient who exhibited at least 90% tumor necrosis after pre-operative chemotherapy. Among the protein peaks, we identified 65 informative peaks based on both the misclassification rate and the fold-change criteria described in the Materials and methods (Table II). These 65 protein peaks constituted a post-treatment signature of chemotherapy response and contained information which could be used to separate the plasma samples into two clusters of good and poor responders (Fig. 2A). The levels of 29 protein peaks were higher and those of 36 protein peaks were lower in the plasma of poor responders relative to that of the good responders. We then used a supervised approach to test whether we could distinguish the good from the poor responders in a more precise manner. We developed a linear SVM classifier to classify the two responder groups using the plasma proteomic profiles of the post-treatment samples. This time the accuracy of the classification was measured using external LOOCV, which includes both feature selection and model building steps at each iteration (13), and hence provides a more accurate estimation of classification error (10). The results showed

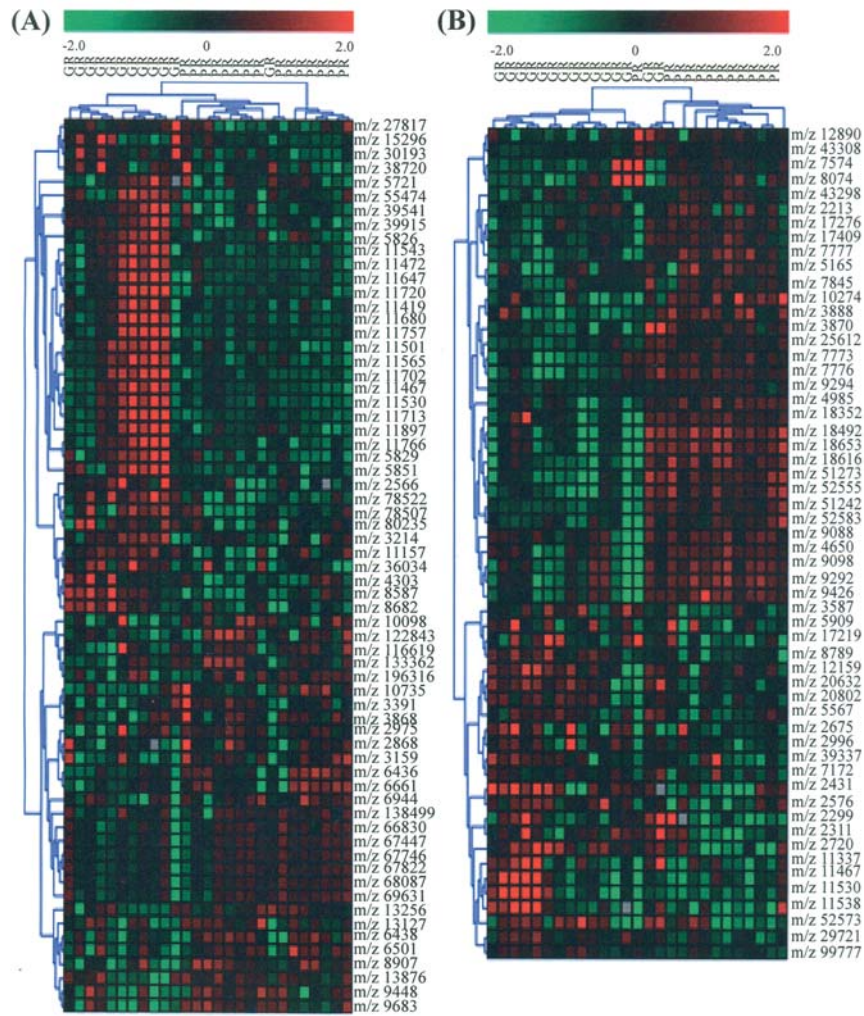


Figure 2. Hierarchical clustering of good and poor responders using (A) post- and (B) pre-treatment osteosarcoma plasma samples. Intensities of the protein peaks were mean centered. Pearson's correlation and average linkage were used for clustering. Red to green scale represents a log peak intensity from -2 to 2.

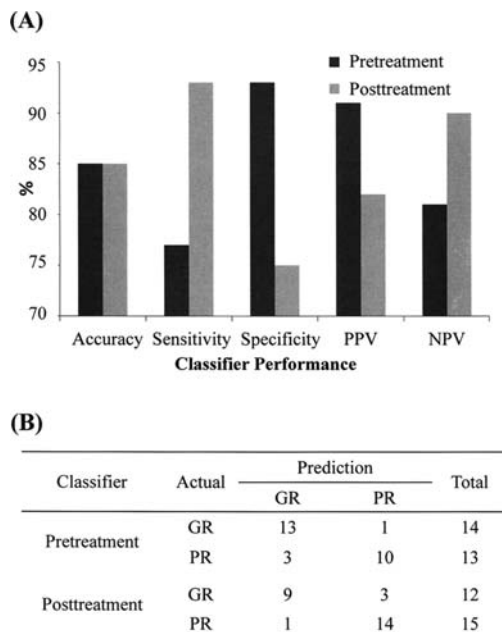


Figure 3. (A) The relative performance and (B) confusion matrix of the pre- and post-treatment proteomic classifiers developed in this study. PPV, positive predictive value; NPV, negative predictive value; GR, good responders; PR, poor responders.

that the post-treatment classifier correctly classified 85% of the samples (misclassified 4 from 27 cases) with 93% sensitivity and 75% specificity in identifying poor responders (Fig. 3A). One out of the 15 poor responders and three of the 12 good responders were misclassified by the post-treatment classifier (Fig. 3B). The positive and negative predictive values of detecting poor responders were 82 and 90%, respectively.

The proteomic classifier of pre-treatment plasma. However, it would be of much greater clinical value if we could predict the chemotherapy response before the start of treatment, so that alternative or personalized treatment could be offered to improve the outcome of the resistant patients. Therefore, we tested whether the proteomic profiles of the pre-treatment plasma collected at the time of initial diagnosis ($n=27$) could be used to detect the potential poor responders. Using the same selection criteria as the post-treatment samples, we identified a proteomic signature of chemotherapy response in the pre-treatment plasma samples, which consisted of 56 protein peaks (Table III). Similar to the post-treatment signature, hierarchical clustering of the pre-treatment plasma samples based on the pre-treatment signature formed two major clusters, which corresponded to good and poor responders (Fig. 2B). In the pre-treatment signature, 32 and

Table III. The informative protein peaks used in the pre-treatment classifier.

No.	Protein peak	Fraction	GR		PR		PR/GR Ratio
			Average	SD	Average	SD	
1	m/z 2,213	F1	0.94	1.33	1.15	1.98	1.22
2	m/z 2,299	F1	1.79	2.50	0.81	2.18	0.45
3	m/z 2,311	F1	1.40	2.27	0.60	2.32	0.43
4	m/z 2,431	F2	1.08	3.65	0.43	2.51	0.40
5	m/z 2,576	F2	1.30	1.43	0.76	2.36	0.59
6	m/z 2,675	F1	1.56	2.26	0.82	1.84	0.53
7	m/z 2,720	F2	1.43	1.63	0.45	3.52	0.31
8	m/z 2,996	F1	1.19	1.73	0.73	1.45	0.62
9	m/z 3,587	F2	1.36	1.61	0.85	1.61	0.63
10	m/z 3,870	F1	0.94	2.47	1.31	1.41	1.40
11	m/z 3,888	F1	0.84	1.82	1.62	1.85	1.93
12	m/z 4,650	F2	0.55	2.39	0.99	2.04	1.82
13	m/z 4,985	F5	0.59	2.36	1.24	1.46	2.11
14	m/z 5,165	F1	0.68	1.83	1.11	1.82	1.63
15	m/z 5,567	F1	1.10	1.44	0.66	1.60	0.59
16	m/z 5,909	F1	1.40	1.93	0.90	2.21	0.64
17	m/z 7,172	F6	1.13	1.35	0.82	1.58	0.73
18	m/z 7,574	F2	0.93	2.01	1.42	1.53	1.52
19	m/z 7,773	F6	0.67	1.83	1.13	1.29	1.70
20	m/z 7,776	F5	0.61	1.93	1.19	1.30	1.93
21	m/z 7,777	F2	0.72	1.42	1.42	1.66	1.97
22	m/z 7,845	F1	0.88	1.55	1.18	1.40	1.34
23	m/z 8,074	F2	0.63	3.32	1.26	1.78	1.98
24	m/z 8,789	F1	1.23	1.51	0.71	1.68	0.58
25	m/z 9,088	F2	0.57	4.03	0.94	2.42	1.65
26	m/z 9,198	F2	0.54	2.75	1.09	2.41	2.03
27	m/z 9,292	F2	0.45	3.65	1.19	2.09	2.61
28	m/z 9,294	F6	0.88	1.23	1.07	1.21	1.22
29	m/z 9,426	F2	0.52	2.94	1.10	2.40	2.12
30	m/z 10,274	F5	0.62	2.34	1.48	2.27	2.40
31	m/z 11,337	F6	1.86	2.32	0.83	1.75	0.45
32	m/z 11,467 ^a	F6	2.00	3.58	0.72	1.94	0.36
33	m/z 11,530 ^a	F6	2.48	3.78	0.83	1.46	0.33
34	m/z 11,538	F1	2.63	3.86	0.63	2.94	0.24
35	m/z 12,159	F1	1.06	2.24	0.50	3.40	0.47
36	m/z 12,890	F3	0.81	1.69	1.07	1.75	1.32
37	m/z 17,219	F1	1.57	2.04	0.89	2.14	0.57
38	m/z 17,276	F5	0.83	1.49	1.13	1.50	1.37
39	m/z 17,409	F5	0.84	1.48	1.15	1.59	1.37
40	m/z 18,352	F1	0.67	2.74	0.97	2.11	1.44
41	m/z 18,492	F1	0.44	2.95	1.04	3.29	2.35
42	m/z 18,616	F2	0.65	1.83	1.44	1.79	2.20
43	m/z 18,653	F1	0.48	2.48	1.10	2.47	2.28
44	m/z 20,632	F2	1.06	3.13	0.70	1.58	0.66
45	m/z 20,802	F2	1.20	1.56	0.78	1.49	0.65
46	m/z 25,612	F5	0.84	1.46	1.16	1.27	1.38
47	m/z 29,721	F5	1.29	1.40	0.82	1.41	0.64
48	m/z 39,337	F6	1.20	1.81	0.88	2.06	0.73
49	m/z 43,298	F5	0.81	1.41	1.15	1.32	1.42
50	m/z 43,308	F3	0.92	1.19	1.11	1.22	1.21

Table III. Continued.

No.	Protein peak	Fraction	GR		PR		PR/GR Ratio
			Average	SD	Average	SD	
51	m/z 51,242	F5	0.78	1.48	1.25	1.52	1.60
52	m/z 51,273	F6	0.59	2.12	1.43	1.48	2.41
53	m/z 52,555	F6	0.63	1.84	1.34	1.67	2.11
54	m/z 52,573	F3	0.95	2.14	0.47	2.11	0.50
55	m/z 52,583	F5	0.82	1.60	1.19	1.78	1.44
56	m/z 99,777	F5	1.21	1.30	0.92	1.25	0.76

These peaks were selected based on a misclassification rate of <0.25 and a fold-change of >1.2 . The peaks are arranged in an ascending order of their m/z values. SD, standard deviation; GR, good responders; PR, poor responders. ^aPeaks which were also identified in the post-treatment classifier.

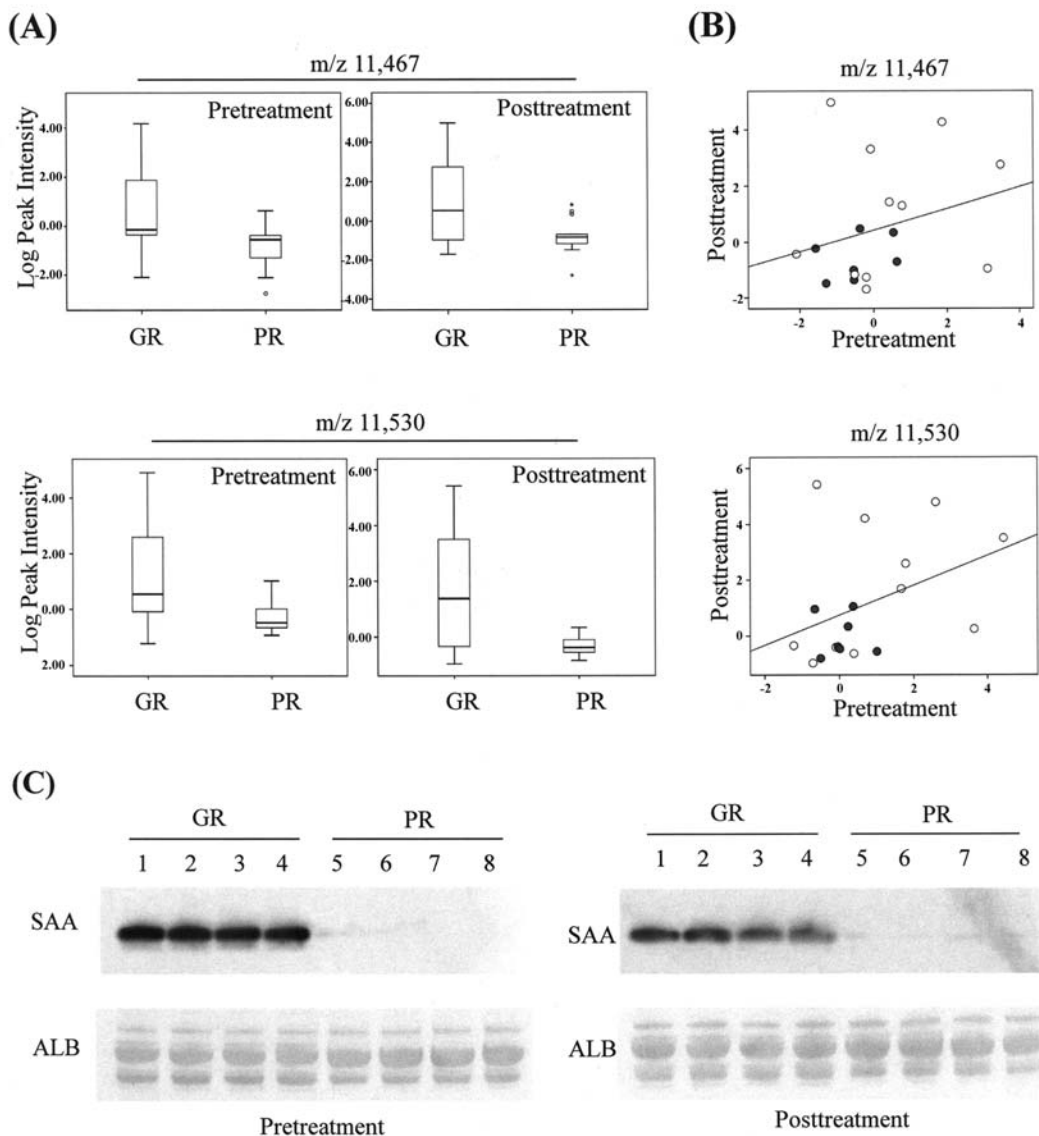


Figure 4. (A) Differential abundance of m/z 11,467 (des Arg Ser SAA) and m/z 11,530 (des 1-Arg SAA) in pre- and post-treatment plasma samples. (B) The correlation analysis of m/z 11,467 and 11,530 protein peaks in paired samples collected before and after chemotherapy. Empty and filled circles denote good and poor responders, respectively. The line represents the fitted line of the plot. (C) Western blot analysis validated the higher abundance of SAA in the pre- and post-treatment plasma samples of good responders. The upper panel shows the SAA levels (11.5 kDa) in the individual plasma samples. The lower panel shows the albumin (ALB) band (66 kDa) in each sample as the loading control. Lanes 1-4 represent good responder samples and lanes 5-8 poor responder samples. GR, good responders; PR, poor responders.

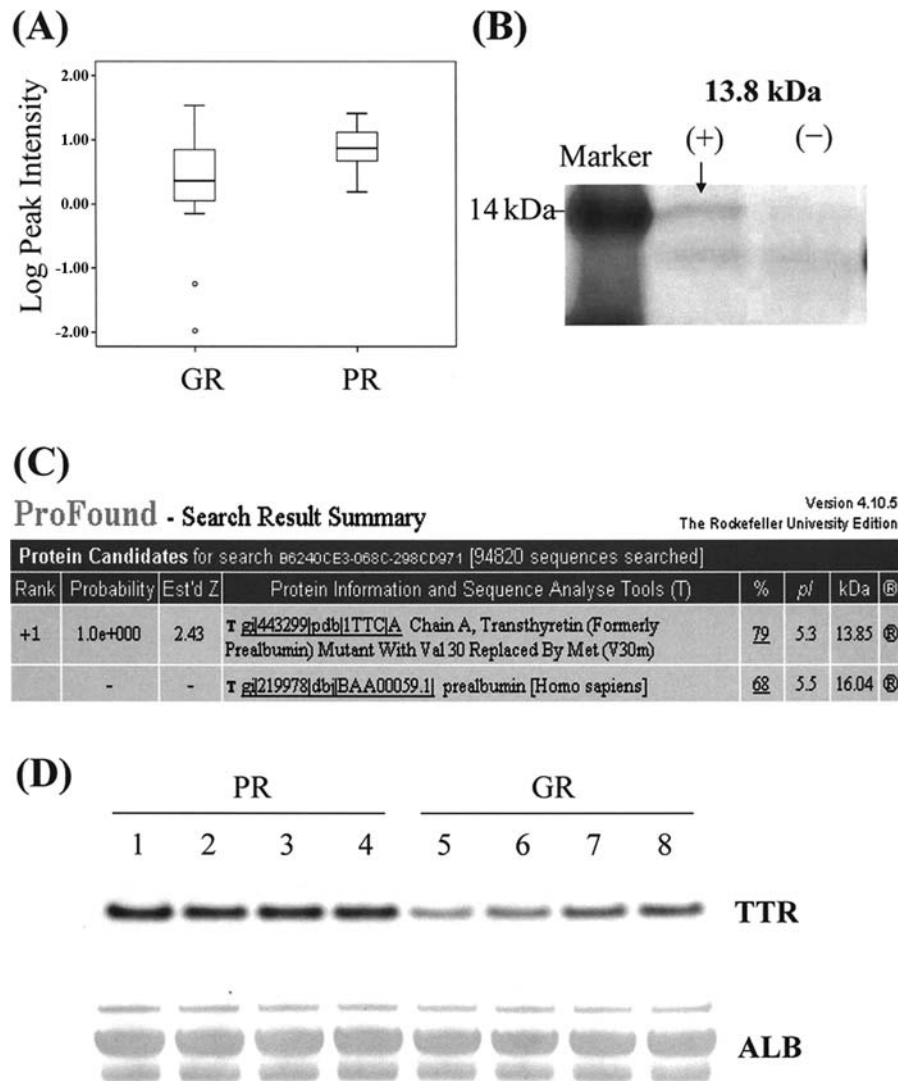


Figure 5. (A) Differential abundance of the m/z 13,876 protein peak in the post-treatment plasma samples of good and poor responders. (B) SDS-PAGE of pooled plasma samples with a high (+) or low (-) level of the 13.8 kDa protein peak. The arrow indicates the position of the candidate TTR band. (C) The ProFound database search result and the probability of matching the candidate 13.8 kDa protein with TTR. (D) Western blot analysis validation of a higher TTR level in the plasma samples of poor responders relative to that in the plasma of good responders. The upper panel shows TTR in the individual plasma samples. The lower panel shows the albumin (ALB) band (66 kDa) in each sample as the loading control. Lanes 1-4 represent poor responder samples and lanes 5-8 good responder samples. GR, good responders; PR, poor responders.

24 protein peaks in the plasma of the poor responders were higher and lower, respectively, than those in the plasma of the good responders. A linear SVM classifier was developed based on the pre-treatment samples and it achieved the same classification accuracy as the post-treatment classifier using LOOCV (85%, Fig. 3A). However, the pre-treatment classifier had a lower sensitivity (77%) but a higher specificity (93%) in recognizing poor responders when compared to the post-treatment classifier (Fig. 3A). Specifically, the pre-treatment classifier misclassified three of the 13 poor responders, and one of the 14 good responders (Fig. 3B). The positive and negative predictive values for the poor responders were 91 and 81%, respectively.

Identification and validation of two proteins in the classifiers. Although multivariate classifiers based on proteomic patterns are generally more accurate than a single biomarker, it is important to determine the identities of the discriminatory

proteins in order to understand the biological basis of the classification. We previously reported that SAA levels were higher in the pre-treatment plasma samples from the osteosarcoma patients relative to a benign bone disease (7). However, the correlation between SAA and chemotherapy response was not analyzed previously. In this study, we discovered that two protein peaks (m/z 11,467 and 11,530) were informative in distinguishing the poor from the good responders in both the pre- and post-treatment classifiers (Tables II and III; Fig. 4). The levels of these two protein peaks were significantly higher in the plasma of the good responders than in that of the poor responders ($p < 0.05$). The size of these two protein peaks (m/z 11,467 and 11,530) matched the size of two known SAA variants, des Arg Ser and des 1-Arg of SAA, respectively (14). Further analysis indicated that the intensities of these two peaks in the samples were highly correlated with each other ($r = 0.97$), suggesting that they could be variants of the same protein. A correlation

analysis of 18 paired patient samples collected before and after the chemotherapy indicated that the poor responders had a low level of SAA before and after the treatment. A similar phenomenon occurred in the good responders, except that they had a relatively higher level of SAA before and after the treatment (Fig. 4B). These results suggest that higher levels of SAA could have a beneficial effect on the patient response to chemotherapy. We then performed Western blot analysis to validate the MS results. Eight plasma samples with a high (good responder group) or low level of the ion peaks (poor responder group) were selected for the validation. The results from the Western blot analysis confirmed that SAA levels were higher in both the pre- and post-treatment plasma samples of the good responders (Fig. 4C).

Based on *m/z* ratios, we also discovered that the size of another protein peak (*m/z* 13,876) in the post-treatment classifier was similar to the reported size of the plasma protein, TTR (15) (Table II). In the post-treatment plasma, the levels of this peak in the plasma of the good responders were significantly lower than those in the plasma of the poor responders ($p < 0.05$, Fig. 5A). The peak with the same *m/z* ratio in the pre-treatment plasma was also lower in the plasma of the good responders. However, it did not reach statistical significance (data not shown). To confirm the identity of the *m/z* 13,876 protein peak, pooled plasma samples from the poor (increased TTR) and good responders (decreased TTR), were separated by electrophoresis. Peptide fingerprint analysis was performed on the gel slice with the molecular weight from the poor responder sample being close to 13.8 kDa (Fig. 5B). ProFound database searching of the peptides indicated that the *m/z* 13,876 protein peak was indeed TTR with a high probability ($z = 2.34$, $p = 1$) (Fig. 5C). In order to further validate this result, we performed Western blot analysis on individual plasma samples collected from the post-treatment sample set using a specific TTR antibody. Our result further confirmed that TTR was increased in the poor responders relative to the good responders (Fig. 5D). TTR is a negative host response protein, which decreases in reaction to inflammation. As both SAA and TTR are known to be involved in host response, the results from our study suggest that the host or innate immune response could be critical for the good response towards chemotherapy.

Discussion

Tumor resistance towards pre-operative chemotherapy is the major poor prognostic factor for patients with localized osteosarcoma (3). Despite efforts to identify prognostic markers in osteosarcoma, the survival of osteosarcoma patients has not improved significantly in the past three decades. The main reason is that the use of post-operative chemotherapy, which is based on the histological response of the tumor after pre-operative chemotherapy, does not improve the outcome of the poor responding patients. This suggests that these poor responders need to be identified at an earlier point in time, preferably at the time of initial diagnosis, before the resistant tumor cells have had a chance to evolve and develop. We, as well as others, have previously demonstrated the use of genomic classifiers for the prediction of chemotherapy response in patients with osteosarcoma (4,16,17). However,

genomic classifiers require tumor biopsies, and these are often difficult to obtain and not suitable for continuous monitoring. Clinically, it would be much more preferable if we could identify poor responders using peripheral blood samples from patients, such as plasma or sera. In this study, we used the proteomic profiles of the patient plasma samples to develop two SVM classifiers for the detection of chemotherapy response at the time of initial diagnosis (pre-treatment) and during definitive surgery (post-treatment). The classifier developed from the pre-treatment samples exhibited the same classification accuracy as that from the post-treatment samples, indicating that the poor chemotherapy response could be predicted at the time of diagnosis and before the start of treatment. Therefore, ineffective therapies could be avoided and alternative or more aggressive therapies could be offered to the patients who are likely to respond poorly to the standardized treatment. The early detection of chemotherapy response would also benefit future risk-based clinical trials, so that different drugs and treatment regimens could be evaluated at the time of diagnosis based on an expected response towards the standardized chemotherapy. To the best of our knowledge, this is the first study on the use of plasma proteome for the prediction of chemotherapy response in patients with osteosarcoma.

As osteosarcoma is a pediatric cancer with a relatively low annual incidence rate, we did not have enough samples to set aside for an independent validation set. Instead, we performed a well-accepted alternative approach, which was the external LOOCV. Our LOOCV includes feature selection and model building steps in each iteration, and hence minimizes the underestimation of classification error (10). A detailed examination of the classifier performances showed that the sensitivities and specificities of the pre- and post-treatment classifiers were slightly different. The pre-treatment classifier was more specific, while the post-treatment classifier was more sensitive in detecting poor responders. The two classifiers complement each other to a certain degree in terms of classification. Two of the four misclassified patients analyzed by the pre-treatment classifier were also analyzed by the post-treatment classifier. One of these two common patients was classified correctly by the post-treatment classifier. Similarly, three of the four misclassified patients in the post-treatment classifier were analyzed in the pre-treatment classifier. Two of them were correctly predicted by the pre-treatment classifier. We are fully aware of the relatively small sample size used in this study. Thus, the utility of the classifiers needs to be validated in a large patient cohort from different institutions. Nonetheless, our results show the potential of using plasma proteomic profiles or proteins to predict or monitor chemotherapy response in patients with osteosarcoma. Our findings could lead to the development of a simple blood test for detecting poor responders in osteosarcoma before the start of treatment.

Identification of the informative proteins in the classifiers is critically important for understanding the biological basis of the classification. The two informative proteins identified in this study were host response proteins. One is SAA, whose amount was higher in the plasma of good responders. SAA is a major component of the apolipoproteins in the high density lipoprotein particle (18). They act as chaperones in the trans-

port of high-density lipoprotein particles. SAA has been found to be overexpressed in a number of cancers, including renal and colorectal cancers (19,20). Previous studies have shown that the association of SAA with cancers may not merely be due to acute phase response (21,22). Evidence also suggests that SAA affects other cellular processes, such as calcium signaling and cell adhesion, which are important in drug resistance (23,24). In our previous study, we demonstrated that the amount of circulating SAA is higher in the plasma of patients with osteosarcoma relative to that in the plasma of patients with benign osteochondroma (7). Studies have also shown that SAA is expressed in bone tissues and osteosarcoma cell lines (25) and that the amounts of SAA in the sera are associated with the relapse of osteosarcoma (26). These results further implicate the importance of this protein in the biology of osteosarcoma. To the best of our knowledge, the role of SAA in chemotherapy response in patients with osteosarcoma has never been described. By combining the proteomic profiles of the pre- and post-treatment plasma samples, we found that two protein peaks (*m/z* 11,467 and 11,530) were higher in the good responders before and after chemotherapy. According to their size, we deduced that they were two known variants of SAA with the deduction of arginine and serine (des Arg Ser, *m/z* 11,467) or one arginine (des 1-Arg, *m/z* 11,530) from the intact SAA protein (14). The analysis based on the paired samples from 18 patients indicated that the expression of these peaks in most of the cases was not affected by the chemotherapy treatment. This could reflect an initial state of host response in the patients before receiving the treatment. Due to the lack of appropriate antibodies, we were not able to validate the results of the two specific SAA variants. However, the results from our Western blot analysis suggest that the total SAA level was higher in the good responders relative to that in the poor responders. We anticipate that the development of an immunoassay specific to either the des Arg Ser or des 1-Arg variant of SAA will further confirm our results.

Another discriminatory protein identified in this study is TTR. While two of the SAA variants exhibited differential abundance, only one variant of TTR (*m/z* 13,876) was significantly different between the good and poor responders. The size of this TTR variant matches the cysteinylated TTR reported by another SELDI study (15). TTR is involved in the transport of thyroxine and retinol-binding protein in the blood and cerebrospinal fluid. As a negative host response protein, the synthesis of TTR is decreased by the liver in reaction to inflammation. It has been reported that the down-regulation of TTR is correlated with the early stage of ovarian cancer, which suggests that TTR could be a reaction to the tumor instead of just other inflammations (27).

Regardless of the other potential roles of the two identified proteins in osteosarcoma, our results suggest that the acute phase response plays an important role in the chemotherapy response in patients with osteosarcoma. As the acute phase response is part of the innate immune system, our results further suggest that innate immunity is potentially activated in the good responders before chemotherapy. Innate immunity is the first immune defense mechanism of the body against pathogens, trauma and malignancy. The activation of innate immunity promotes the anti-tumor activity of chemotherapy.

In osteosarcoma, the use of a specific immune stimulating compound, muramyl tripeptide phosphatidylethanolamine (MTT-PE), has been shown to stimulate and promote anti-tumor activity in canines and humans with metastatic osteosarcoma (28,29). It has been shown that MTT-PE can improve the overall survival of osteosarcoma patients when used together with conventional chemotherapy (30). Our results corroborate these findings and further suggest that patients with heightened innate immunity could have higher anti-tumor activity, and hence a better histological response. In summary, the use of plasma proteomic profiles, as well as immunostimulating compounds and conventional chemotherapy, could improve the detection and treatment of the poor responding osteosarcoma patients.

Acknowledgements

We would like to thank Dr Tai-Tung Yip for technical consultation in this study. We also wish to thank Dr David Gorenstein (University of Texas Medical Branch, Galveston, Texas) for the use of the Biomek 2000. We are grateful to Carolyn Pena for her assistance in the manuscript preparation. This study was supported by grants from the Carl C. Anderson, Sr. and Marie Jo Anderson Charitable Foundation, the Sarcoma Foundation of America, and the Wendy Will Case Cancer Foundation (T.K.M.).

References

- Huvos A: Bone Tumors: Diagnosis, Treatment and Prognosis. 2nd edition. WB Saunders, Philadelphia, 1991.
- Dahlin DC and Unni K: Bone Tumors: General Aspects and Data on 8542 Cases. 4th edition. Charles C. Thomas, Springfield, 1986.
- Link MP, Gebhardt MC and Meyers PA: Osteosarcoma. In: Principles and Practice of Pediatric Oncology. Pizzo P and Poplack D (eds). 4th edition. Lippincott-Williams & Wilkins, Philadelphia, pp1051-1089, 2002.
- Man TK, Chintagumpala M, Visvanathan J, *et al*: Expression profiles of osteosarcoma that can predict response to chemotherapy. *Cancer Res* 65: 8142-8150, 2005.
- Yip TT and Lomas L: SELDI ProteinChip array in oncoproteomic research. *Technol Cancer Res Treat* 1: 273-280, 2002.
- Wiesner A: Detection of tumor markers with ProteinChip technology. *Curr Pharm Biotechnol* 5: 45-67, 2004.
- Li Y, Dang TA, Shen J, *et al*: Identification of a plasma proteomic signature to distinguish pediatric osteosarcoma from benign osteochondroma. *Proteomics* 6: 3426-3435, 2006.
- Simon R, Lam A, Li MC, Ngan M, Menendez S and Zhao Y: Analysis of Gene Expression Data Using BRB-Array Tools. *Cancer Inform* 3: 11-17, 2007.
- Vapnik V and Chapelle O: Bounds on error expectation for support vector machines. *Neural Comput* 12: 2013-2036, 2000.
- Radmacher MD, McShane LM and Simon R: A paradigm for class prediction using gene expression profiles. *J Comput Biol* 9: 505-511, 2002.
- Dang TA and Man TK: Classification of sarcomas using bioinformatics and molecular profiling. *Curr Pharm Biotechnol* 8: 83-91, 2007.
- Zhang W and Chait BT: ProFound: an expert system for protein identification using mass spectrometric peptide mapping information. *Anal Chem* 72: 2482-2489, 2000.
- Dudoit S and Friedman N: Classification in microarray experiments. In: *Statistical Analysis of Gene Expression Microarray Data*. Speed T (ed). Chapman & Hall/CRC, pp93-158, 2003.
- Tolson J, Bogumil R, Brunst E, *et al*: Serum protein profiling by SELDI mass spectrometry: detection of multiple variants of serum amyloid alpha in renal cancer patients. *Lab Invest* 84: 845-856, 2004.

15. Fung ET, Yip TT, Lomas L, *et al*: Classification of cancer types by measuring variants of host response proteins using SELDI serum assays. *Int J Cancer* 115: 783-789, 2005.
16. Ochi K, Daigo Y, Katagiri T, *et al*: Prediction of response to neoadjuvant chemotherapy for osteosarcoma by gene-expression profiles. *Int J Oncol* 24: 647-655, 2004.
17. Mintz MB, Sowers R, Brown KM, *et al*: An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. *Cancer Res* 65: 1748-1754, 2005.
18. Rienhoff HY Jr, Huang JH, Li XX and Liao WS: Molecular and cellular biology of serum amyloid A. *Mol Biol Med* 7: 287-298, 1990.
19. Kimura M, Tomita Y, Imai T, *et al*: Significance of serum amyloid A on the prognosis in patients with renal cell carcinoma. *Cancer* 92: 2072-2075, 2001.
20. Glojnaric I, Casl MT, Simic D and Lukac J: Serum amyloid A protein (SAA) in colorectal carcinoma. *Clin Chem Lab Med* 39: 129-133, 2001.
21. Cho WC, Yip TT, Yip C, *et al*: Identification of serum amyloid a protein as a potentially useful biomarker to monitor relapse of nasopharyngeal cancer by serum proteomic profiling. *Clin Cancer Res* 10: 43-52, 2004.
22. Rosenthal CJ and Sullivan LM: Serum amyloid A to monitor cancer dissemination. *Ann Intern Med* 91: 383-390, 1979.
23. Badolato R, Johnston JA, Wang JM, McVicar D, Xu LL, Oppenheim JJ and Kelvin DJ: Serum amyloid A induces calcium mobilization and chemotaxis of human monocytes by activating a pertussis toxin-sensitive signaling pathway. *J Immunol* 155: 4004-4010, 1995.
24. Xu L, Badolato R, Murphy WJ, *et al*: A novel biologic function of serum amyloid A. Induction of T lymphocyte migration and adhesion. *J Immunol* 155: 1184-1190, 1995.
25. Kovacevic A, Hammer A, Stadelmeyer E, *et al*: Expression of serum amyloid A transcripts in human bone tissues, differentiated osteoblast-like stem cells and human osteosarcoma cell lines. *J Cell Biochem* 103: 994-1004, 2008.
26. Jin S, Shen J-N, Guo Q-C, *et al*: 2-D DIGE and MALDI-TOF-MS analysis of the serum proteome in human osteosarcoma. *Proteomics Clin Appl* 1: 272-285, 2007.
27. 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.
28. Kurzman ID, Shi F, Vail DM and MacEwen EG: In vitro and in vivo enhancement of canine pulmonary alveolar macrophage cytotoxic activity against canine osteosarcoma cells. *Cancer Biother Radiopharm* 14: 121-128, 1999.
29. Kleinerman ES, Gano JB, Johnston DA, Benjamin RS and Jaffe N: Efficacy of liposomal muramyl tripeptide (CGP 19835A) in the treatment of relapsed osteosarcoma. *Am J Clin Oncol* 18: 93-99, 1995.
30. Chou AJ, Kleinerman ES, Krailo MD, *et al*: Addition of muramyl tripeptide to chemotherapy for patients with newly diagnosed metastatic osteosarcoma: a report from the Children's Oncology Group. *Cancer* 115: 5339-5348, 2009.