

Prediction of the response to chemotherapy in advanced esophageal cancer by gene expression profiling of biopsy samples

MASAAKI MOTOORI^{1,3}, ICHIRO TAKEMASA¹, MAKOTO YAMASAKI¹, TAKAMICHI KOMORI¹,
ATSUSHI TAKENO¹, HIROSHI MIYATA¹, SHUJI TAKIGUCHI¹, YOSHIYUKI FUJIWARA¹,
TAKUSHI YASUDA¹, MASAHIKO YANO³, NARIAKI MATSUURA², KENICHI MATSUBARA⁴,
MORITO MONDEN¹, MASAKI MORI¹ and YUICHIRO DOKI¹

Departments of ¹Gastroenterological Surgery and ²Functional Diagnostic Science, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871; ³Department of Surgery, Osaka Medical Center for Cancer and Cardiovascular Diseases, 1-3-3 Nakamichi, Higashinari-ku, Osaka, Osaka 537-8511; ⁴DNA Chip Research Inc., 1-1-43 Suehiro, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Received May 4, 2010; Accepted June 28, 2010

DOI: 10.3892/ijo_00000763

Abstract. To improve the prognosis of advanced esophageal cancer, neoadjuvant chemotherapy (NACT) followed by surgery is a promising treatment strategy. NACT has been shown to improve the prognosis of responders. However, non-responders not only suffer from side-effects, but they also lose precious time to take advantage of other possible treatments. Therefore, it is crucial to establish a reliable method that allows prediction of response before chemotherapy. A biopsy sample can provide valuable information on the biological characteristics of an individual esophageal cancer, which can affect chemosensitivity. Comprehensive gene expression profiling (GEP) using oligonucleotide microarray covering 30,000 human probes was performed in 50 pre-treatment endoscopic biopsy samples from 25 patients with esophageal squamous cell cancer (ESCC) who underwent cisplatin-based chemotherapy (two samples per patient). Chemotherapeutic responses were evaluated by the reduction rate of the tumor area on CT scans. Responders were defined as patients with reduction rates of $\geq 50\%$ and non-responders were defined as patients with $< 50\%$ decrease. The diagnostic system, that predicts responses to chemotherapy, was constructed with the 199 most informative genes, and showed 82% of accuracy. Furthermore, the predictive performance of this system was confirmed using an additional ten samples with an accuracy of 80%. This study shows that GEP of pre-treatment ESCC biopsy samples has the potential to predict responses to chemotherapy.

Introduction

Despite recent advances in surgical techniques and peri-operative management, surgery alone has not been satisfactorily able to improve the prognosis of advanced esophageal cancer. Even after curative resection by esophagectomy with extended lymphadenectomy, about 50% of patients show recurrence, (1) suggesting that micrometastasis may exist outside the surgical field at the time of diagnosis. Neoadjuvant chemotherapy (NACT), which is expected to eradicate systemic micrometastasis, followed by surgery is a promising treatment strategy that improves prognoses for advanced esophageal cancer patients. The significance of NACT is controversial, but there has been a consistent observation that survival is significantly prolonged in responders (2,3). While, non-responders not only suffer from side-effects but also lose a precious time to take advantage of other possible treatments, like chemoradiotherapy. Therefore, a reliable method that allows prediction of response before chemotherapy is considered to be crucial for the future use of NACT in treating advanced esophageal cancer.

It has not been possible to predict a patient's response to chemotherapy based on pretreatment clinical parameters. However, a tumor's biological characteristics are one of the most important factors that affect chemosensitivity. Molecular analyses of pretreatment biopsy samples, such as RT-PCR and immunohistochemistry, have been performed to understand biological characteristics of esophageal cancer (4-7). However, only one or a few genes have been addressed in these reports. Multiple genetic alterations are involved in the development and progression of esophageal cancer, and these aberrations may affect the expression of large number of genes (8,9). Moreover, numerous molecular pathways may contribute to the sensitivity of a tumor to chemotherapy or radiotherapy. Gene expression profiling (GEP) allows assessment of expression of thousands of genes simultaneously, making it one of the powerful tools for understanding the biological characteristics of a tumor. In fact, this approach has already been used to identify genes that could serve as

Correspondence to: Dr Ichiro Takemasa, Department of Gastroenterological Surgery, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
E-mail: alfa-t@sf6.so-net.ne.jp

Key words: esophageal cancer, chemosensitivity, biopsy sample, gene expression profiling

molecular markers of cancer classification and outcome prediction (10-14). Recently, GEP has been successfully performed not only using surgically resected samples but also biopsy samples (15-18). Our previous study found that biopsy esophageal squamous cell cancer (ESCC) samples were distinguishable from normal esophageal epithelia using GEP and intratumor heterogeneity of GEP was smaller than intertumor heterogeneity. This means that GEP of ESCC biopsy samples has the potential to represent the biological properties of the entire esophageal cancer (19). In this study, we performed comprehensive GEP of pretreatment biopsy samples to predict response to chemotherapy.

Patients and methods

Patients and clinical samples. We performed chemotherapy on ESCC patients with cN1 or cM1(lym) with any cT stage as neoadjuvant therapy or on patients with any cM1 other than lymph node metastasis. The first line chemotherapy protocol was cisplatin combined with adriamycin and 5-fluorouracil (FAP therapy) (20,21). The treatment regimen of FAP therapy was as follows: cisplatin (70 mg/m²) and doxorubicin hydrochloride (adriamycin, 35 mg/m²) were administered by drip infusion on day 1, and 5-fluorouracil (700 mg/m²) was administered by continuous infusion on days 1-7. Two cycles of chemotherapy were given, separated by a 3-week interval. Before treatment, we obtained a couple of endoscopic biopsy samples from each of the 25 patients and assayed them separately. We assigned these patients to an estimation set. Tissue specimens were disrupted in RNAlater (Ambion, Austin, TX) and stored at 4°C for 1-2 h, then at -80°C until use. For each biopsy specimen, an adjacent cancer tissue biopsy was given to a pathologist to assess the presence of cancer and its histology. Characteristics of the estimation set are listed in Table I. None of the patients received any treatment before the endoscopy.

As an independent validation set, we obtained endoscopic biopsy samples from 10 patients (one each) before chemotherapy. Characteristics of the validation set are also listed in Table I. The treatment regimen in the validation set was the same as in the estimation set. All aspects of our study protocol were performed according to the ethical guidelines set by the committee of the three Ministries of the Japanese Government and a signed consent form was obtained from each subject.

Evaluation of chemotherapeutic responses. Patients underwent CT scan before and two weeks after each cycle of chemotherapy. The largest tumor area of the primary tumor was measured bidimensionally, using the greatest diameter and the greatest perpendicular distance before and after chemotherapy. When the primary tumor could not be detected in CT scan, the largest area of the metastatic lymph node was measured. Responders were defined as patients with reduction rates of >50%; non-responders were defined as patients with <50% decrease.

Preparation of fluorescent-labeled aRNA targets and hybridization. Total RNA was purified from clinical samples utilizing TRIzol reagent (Invitrogen, San Diego, CA). For

control reference, 15 RNAs from normal esophageal epithelia were mixed. Extracted RNAs were amplified with T7 RNA polymerase using Amino Allyl MessageAmp™ aRNA kit (Ambion). Five μg of control- and experimental-aRNA samples were labeled with Cy3 and Cy5, respectively, mixed and hybridized on an oligonucleotide microarray covering 30,000 human probes (AceGene Human 30K; DNA Chip Research Inc. and Hitachi Software Engineering Co., Ltd., Yokohama, Japan). The experimental protocol is available at <http://www.dna-chip.co.jp/thesis/AceGeneProtocol.pdf>. Then, the microarrays were scanned using ScanArray 4000 (GSI Lumonics, Billerica, MA).

Analysis of microarray data. Signal values were calculated by DNASISArray software (Hitachi Software Inc. Tokyo, Japan). Following background subtraction, data with low signal intensities were excluded from additional investigation. In each sample, Cy5/Cy3 ratio values were log-transformed and global equalization to remove a deviation of the signal intensity between whole Cy3- and Cy5-fluorescence was performed by subtracting a median of all log(Cy5/Cy3) values from each log(Cy5/Cy3) value. Genes with missing values in >10% of samples were excluded from further analysis.

Hierarchical cluster analysis (HCA) with Euclidean distance as a similarity coefficient and Ward as a clustering algorithm was performed using GeneMath 2.0 software (Applied Maths, Inc., Austin, TX). To predict response in an unknown test sample, we adopted a weighted voting (WV) algorithm, which is a method for making classifications between two classes (10); in this work class 1 refers to the responder group and class 2 refers to the non-responder group. First, we calculated the signal-to-noise ratio (SNR), $S_i = (\mu_1 - \mu_2) / (\sigma_1 + \sigma_2)$ where μ and σ denote mean and standard deviation values of expression levels of a gene in each of the two classes. Each gene was assigned a 'vote', which was the weighted difference between the gene expression level in a test sample and the average of the two classes: $v_i = S_i \times (X_i - (\mu_1 + \mu_2) / 2)$. The ultimate vote for a particular class assignment was computed by summing all weighted votes made by the genes used in the class discrimination. The prediction strength (PS) was defined as $PS = (V_1 - |V_2|) / (V_1 + |V_2|)$, where V_1 is summed votes exceeding the threshold (here 0), and V_2 is those less than the threshold, respectively. When PS was ≥ 0 , we determined the test sample belonged to class 1. When PS was < 0 , we determined the test sample belonged to class 2. This model was evaluated by leave-one-out cross-validation, whereby one sample was withheld, a gene expression based model was trained using the remaining samples, and then the model was used to predict the class of the withheld sample (10). The process was repeated to cover the entire samples, and the cumulative accuracy was recorded. The leave-one-out cross-validation used in this study repeats the entire model-building process in each of the cross-validation sets without information leakage, including the selection of genes used in the class discrimination.

To select genes potentially contributing to chemosensitivity, random permutation test, which involves randomly permuting class labels to determine gene-class correlations, was used (10). The original score, $S_o = |\mu_1 - \mu_2|$, of each gene was calculated without permuting the labels (responder or

Table I. Characteristics of the estimation set and the validation set.

Estimation set								
Gender	Age	Cycles of chemotherapy	Tumor size (mm)	T	Lymph node metastasis	M (other than LYM)	Stage	Response
M	63	2	40	3	N1	0	III	R
M	67	2	49	3	N1	0	III	R
F	76	2	41	4	M1(LYM)	1	IV	R
M	65	2	20	3	M1(LYM)	0	IV	R
F	68	2	49	2	N1	1	IV	R
M	52	2	22	2	M1(LYM)	0	IV	R
M	75	2	16	1	M1(LYM)	0	IV	R
M	72	2	43	3	N1	0	III	R
M	66	2	36	3	N1	0	III	R
M	76	2	19	2	N1	0	IIB	R
M	68	2	63	4	N1	0	III	R
F	69	2	49	3	N1	0	III	NR
M	63	2	39	3	M1(LYM)	0	IV	NR
F	53	2	51	3	M1(LYM)	1	IV	NR
F	65	1	32	2	N1	0	IIB	NR
M	69	2	69	3	N1	0	III	NR
M	61	2	40	3	N1	1	IV	NR
F	50	2	36	2	M1(LYM)	0	IV	NR
M	55	1	47	3	M1(LYM)	0	IV	NR
M	74	2	48	3	N1	0	III	NR
M	73	1	30	3	N1	0	III	NR
F	62	2	37	4	M1(LYM)	0	IV	NR
M	72	2	48	4	M1(LYM)	0	IV	NR
M	60	1	52	4	M1(LYM)	1	IV	NR
M	67	1	40	4	M1(LYM)	0	IV	NR
Validation set								
Gender	Age	Cycles of chemotherapy	Tumor size (mm)	T	Lymph node metastasis	M (other than LYM)	Stage	Response
M	68	2	42	3	N1	0	III	R
M	58	2	26	3	M1(LYM)	1	IV	R
M	57	2	32	3	M1(LYM)	0	IV	R
M	62	2	31	3	N1	0	III	R
M	66	2	54	3	N1	1	IV	NR
M	58	1	34	2	N1	0	IIB	NR
M	58	1	29	2	N1	0	IIB	NR
F	63	2	37	3	N1	0	III	NR
M	59	2	40	3	M1(LYM)	0	IV	NR
M	56	1	34	3	M1(LYM)	0	IV	NR

Clinical diagnosis of tumor depth, lymph node metastasis, distant metastasis, and stage were classified according to TNM classification (6th edition). R, responder; NR, non-responder.

non-responder). The labels of all the samples were randomly permuted, and the scores were recalculated between two groups consisting of the new members. Repetition of this

permutation 10,000 times provides a score number (No) larger than the original score (So). For each gene, the permutation P-value was determined by $P = \text{No}/10,000$.

Statistical analyses were performed using StatView 5.0J software. Relationships between responses and variable clinicopathological characteristics were evaluated by the χ^2 test. Differences were considered to be significant at $P < 0.05$.

Real-time quantitative RT-PCR. To verify our microarray data, we randomly chose five genes (*PRDX6*, *SRF*, *PERP*, *DADI*, *SELPINB6*) out of predictive genes and performed real-time quantitative RT-PCR (qRT-PCR) on 20 biopsy ESCC specimens (responder: 10, non-responder: 10) and the control reference. Complementary DNA (cDNA) was generated using the Reverse Transcription System (Promega, Madison, WI). Quantification of mRNA expression was performed using a real-time thermal cycler, the LightCycler, and detection system (Roche Diagnostics, Mannheim, Germany). For detection of amplification products, Light-Cycler-DNA master SYBR green I (Boehringer-Mannheim, Mannheim, Germany) was used. Briefly, a 20- μ l reaction volume containing 2 μ l of cDNA and 0.2 μ mol/l of each primer was applied to a glass capillary. PCR conditions were as follows: one cycle of denaturing at 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, 62°C for 5 sec and 72°C for 10 sec. The fluorescence intensity was calculated at each cycle and the standard curve was constructed with 3-fold serial dilutions of cDNA of the control reference. Quantitative analysis of mRNA was performed using LightCycler analysis software (Roche Diagnostics). The relative expression level of a candidate gene was computed with respect to an internal standard glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) mRNA to normalize for variations in the amount of input cDNA.

Results

Clinicopathological characteristics. Characteristics of both the estimation and validation set are shown in Table I. In the estimation set, 20 patients received two cycles of chemotherapy and five patients received one cycle because tumor size reduction was not observed. Seventeen patients underwent surgery, and eight patients were not indicated for surgery because of cM1 other than lymph node metastasis or disease that had advanced too far. Eleven patients (44%) were classified as responders, and the remaining fourteen patients (56%) were classified as non-responders. In the validation set, seven patients received two cycles of chemotherapy, and three patients received one cycle because tumor size reduction was not observed. Six patients underwent surgery and four patients were not indicated for surgery because of cM1 other than lymph node metastasis or too far advanced disease. Four patients (40%) were classified as responders, and the remaining six patients (60%) were classified as non-responders. Clinicopathological characteristics of both the estimation and validation set are shown in Table II. There were no significant differences in any factors between the two groups.

Construction of a diagnostic system to predict response. After gene processing described previously, 19,166 genes were used for further analysis. Unsupervised HCA using all 19,166 genes is shown in Fig. 1. A couple of biopsy samples obtained from the same case were most closely clustered

Table II. Clinical characteristics between responders and non-responders.

Characteristics	Estimation set		P-value
	Responder (n=11)	Non-responder (n=14)	
Gender			
M	9	9	NS
F	2	5	
Age (mean)	65.6	64	NS
Tumor size			
<40 mm	5	5	NS
>40 mm	6	9	
T			
1, 2	4	2	NS
3, 4	7	12	
Lymph node metastasis			
N1	4	6	NS
M1(LYM)	7	8	
M (other than LYM)			
M0	9	11	NS
M1	2	3	
Stage			
II, III	6	5	NS
IV	5	9	
Characteristics	Validation set		P-value
	Responder (n=4)	Non-responder (n=6)	
Gender			
M	4	5	NS
F	0	1	
Age (mean)	60.5	58.5	NS
Tumor size			
<40 mm	3	5	NS
>40 mm	1	1	
T			
1, 2	0	2	NS
3, 4	4	4	
Lymph node metastasis			
N1	2	3	NS
M1(LYM)	2	2	
M (other than LYM)			
M0	3	5	NS
M1	1	1	
Stage			
II, III	2	3	NS
IV	2	3	

Clinical diagnosis of tumor depth, lymph node metastasis, distant metastasis, and stage were classified according to TNM classification (6th edition).

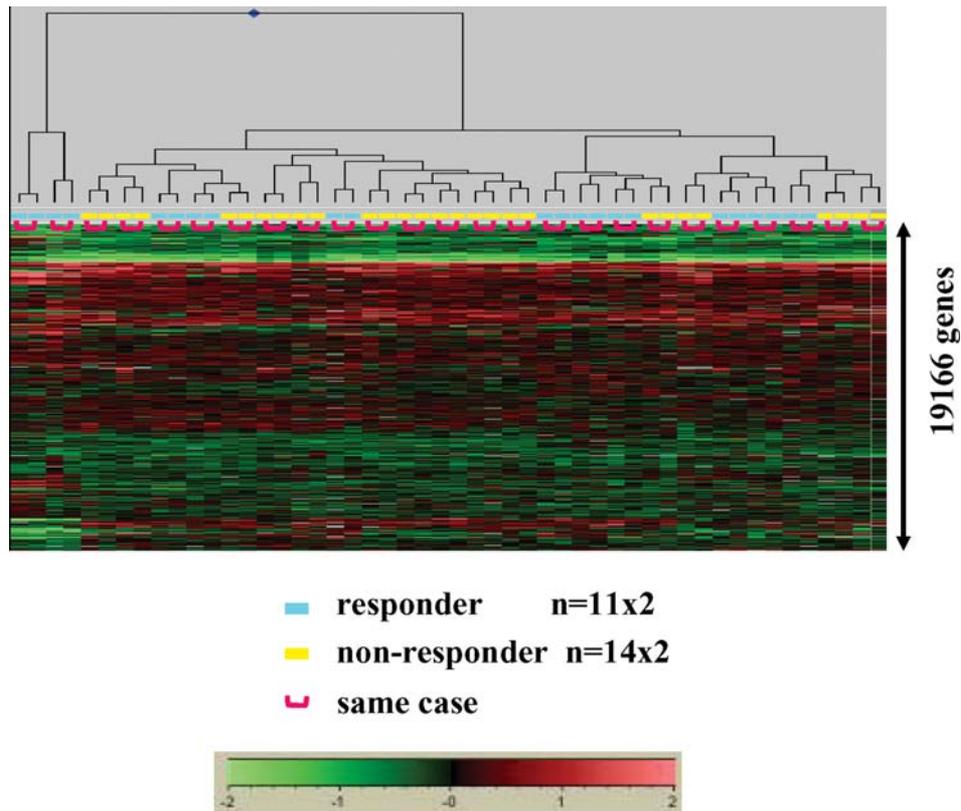


Figure 1. Hierarchical cluster analysis with whole genes in 50 biopsy ESCC samples from 25 patients of the estimation set. The rows and columns represent genes and samples, respectively. The color scale at the bottom indicates the relative expression levels in terms of standard deviations from the median. A couple of biopsy samples obtained from the same case were most closely clustered together.

together. When all 19,166 genes were used for prediction of response, the accuracy was 72% by the complete leave-one-out cross-validation. Then, we tried to reduce the number of genes used for prediction for clinical setting. First, we eliminated genes whose expression data differed by >1.5 -fold between pairs of samples from the same case in at least 20% of the cases, to remove genes that had a wide variation in the same case. Next, we defined significant genes as those with $P < 0.001$ by random permutation test. The leave-one-out cross-validation using significant genes showed the accuracy was 82%. The positive and negative predictive values were 88.2% (15/17) and 78.8% (26/33), respectively. We therefore defined these 199 genes with $P < 0.001$ between the 22 responder samples and 28 non-responder samples as predictive genes and constructed a diagnostic system using WV algorithm. Top 55 of predictive genes are listed in Table III. The expression patterns of 199 predictive genes showed distinct profiles between the two groups (Fig. 2).

Evaluation of the diagnostic system using an additional data set. Next, we validated the diagnostic system in the validation set to confirm its predictive performance. Our diagnostic system correctly predicted responses in 8 of 10 (80.0%) cases. The positive and negative predictive values were 75.0% (3/4) and 83.3% (5/6), respectively.

Real-time quantitative RT-PCR analysis. The relative expression values of *PERP* were lower in non-responders, and those of *PRDX6*, *DAD1*, *SELPINB6*, and *SRF* were

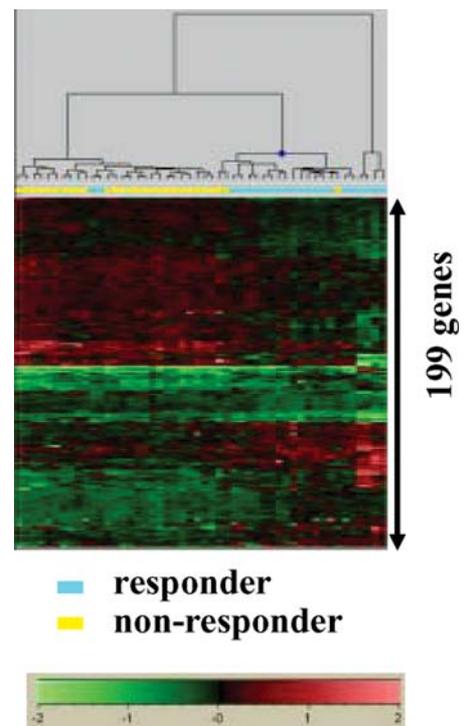


Figure 2. Hierarchical cluster analysis with predictive genes in 50 biopsy ESCC samples from 25 patients of the estimation set. The rows and columns represent genes and samples, respectively. The color scale at the bottom indicates the relative expression levels in terms of standard deviations from the median. The expression patterns of predictive genes showed distinct profiles between the two groups.

Table III. List of predictive genes (top 55).

Rank	AceGene Probe ID	RefSeq accession no.	Rep. gene symbol	P-value ^a	Up or down in non-responder
1	AGhsPSID01001654	NM_016258	YTHDF2	0	Down
2	AGhsPSID01001734	NM_024949	WWC2	0	Down
3	AGhsPSID01001773		FBXO31	0	Down
4	AGhsPSID01006730			0	Down
5	AGhsPSID01007687	NM_004568	SERPINB6	0	Up
6	AGhsPSID01010465	NR_002139	HCG4	0	Up
7	AGhsPSID01011257		PTF1A	0	Down
8	AGhsPSID01011870	XM_001720619		0	Up
9	AGhsPSID01015510	NM_006836	GCN1L1	0	Up
10	AGhsPSID01015629	XR_017723		0	Up
11	AGhsPSID01015781	NM_003196	TCEA3	0	Up
12	AGhsPSID01016693	NM_052925	LENG8	0	Up
13	AGhsPSID01018176	NM_199265	THSD3	0	Up
14	AGhsPSID01018214	NM_032569		0	Up
15	AGhsPSID01018230	XR_041827		0	Up
16	AGhsPSID01019416			0	Up
17	AGhsPSID01020366			0	Up
18	AGhsPSID01020754			0	Up
19	AGhsPSID01021895	NM_181784	SPRED2	0	Up
20	AGhsPSID01023047			0	Up
21	AGhsPSID01023680			0	Up
22	AGhsPSID01024546			0	Up
23	AGhsPSID01024757			0	Up
24	AGhsPSID01025093			0	Up
25	AGhsPSID01025691			0	Up
26	AGhsPSID01027780			0	Up
27	AGhsPSID01028294			0	Up
28	AGhsPSID01028399			0	Up
29	AGhsPSID01029206			0	Up
30	AGhsPSID01003661	NM_006869	CENTA1	0	Down
31	AGhsPSID01005792	NM_014300	SEC11A	0	Up
32	AGhsPSID01006241	NM_016615	SLC6A13	0	Down
33	AGhsPSID01007421	NM_004279	PMPCB	0	Down
34	AGhsPSID01009369	XR_019355	PRDX6	0	Up
35	AGhsPSID01015604	NM_024744	ALS2CR8	0	Up
36	AGhsPSID01015949			0	Up
37	AGhsPSID01019141	NM_018648	NOLA3	0	Up
38	AGhsPSID01020673	XR_042052	C2orf59	0	Up
39	AGhsPSID01021333			0	Up
40	AGhsPSID01022139	NM_001080458	EVX2	0	Up
41	AGhsPSID01024824			0	Up
42	AGhsPSID01026968			0	Down
43	AGhsPSID01008409	NM_001008390	CGGBP1	0	Down
44	AGhsPSID01013314	,NM_003145	SSR2	0	Up
45	AGhsPSID01015920	NM_014874	MFN2	0	Down
46	AGhsPSID01016024	NM_203477	RPL23AP7	0	Up
47	AGhsPSID01020259			0	Up
48	AGhsPSID01027568			0	Up
49	AGhsPSID01027956	NM_001002911	GPR139	0	Up
50	AGhsPSID01000407	NM_175609	ARFGAP1	0	Down
51	AGhsPSID01000584	NM_022873	IFI6	0	Down
52	AGhsPSID01004314	NM_018196	TMLHE	0	Down
53	AGhsPSID01014324	NM_173630	RTTN	0	Up
54	AGhsPSID01007495	NM_172057	KCNH2	0	Down
55	AGhsPSID01000660	NM_025087		0	Down

^aP-value was calculated by random permutation test.

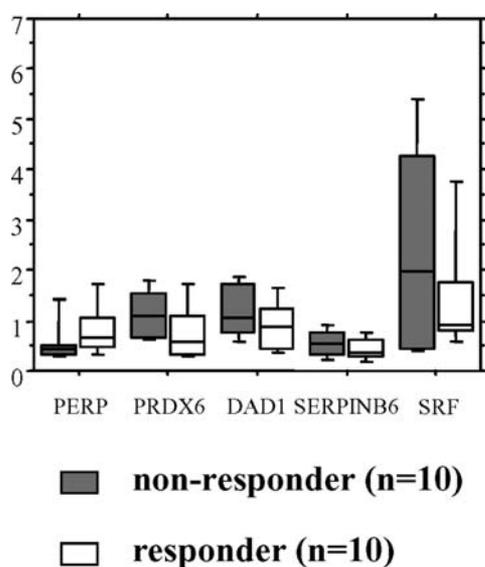


Figure 3. Real-time quantitative RT-PCR expression levels of five genes (*PERP*, *PRDX6*, *DAD1*, *SERPINB6*, and *SRF*) in 20 biopsy samples (10 non-responders and 10 responders).

higher in non-responders (Fig. 3). The expression patterns of qRT-PCR of these five genes were in agreement with their expressions on the microarray.

Discussion

Prediction of the response to chemotherapy or chemoradiotherapy is one of the most desirable achievements in the clinical field. There are several qualities of a tumor that can help to predict its response to therapy, such as its metabolic activity and biological characteristics. Recently, positron emission tomography (PET) has been introduced to assess tumor metabolic activity. In case of neoadjuvant therapy in esophageal cancer, several investigators reported that PET might be useful for predicting responses early during or after neoadjuvant therapy (22-24). Another approach for predicting responses to therapy is to analyze the biological characteristics of a tumor, which can be grasped before the administration of therapy by molecular analysis of pretreatment biopsy samples. Weber *et al* suggested that PET allows quantitative assessment of the entire tumor mass, whereas molecular analysis of pretherapeutic biopsies may not be representative for the entire tumor mass because of tumor heterogeneity (22). However, we previously found that the intratumor heterogeneity of GEP of ESCC biopsies is smaller than the intertumor heterogeneity and the GEP of ESCC biopsies may potentially represent the GEP of whole tumors (19). In contrast, since PET evaluates changes in metabolic activity before and after treatment, this approach is unable to predict a response prior to administration of therapy. Molecular analysis of pretherapeutic biopsies has this principal advantage over PET in that unnecessary therapy may be avoided.

Several studies regarding the predictions of treatment response by GEP in esophageal cancer have been reported. Luthra *et al* performed GEP of 19 pretreatment endoscopic

biopsy samples and found that HCA segregated the cancers into two groups that correlated with the response to chemoradiotherapy (17). Duong *et al* performed GEP of 46 pretreatment endoscopic biopsy samples and identified a 32-gene classifier that can be used to predict the response to chemoradiotherapy in ESCC (25). These studies showed the potential of GEP of biopsies for pretreatment prediction of chemoradiotherapy responses. In this study, we performed GEP of 60 ESCC biopsies, and successfully predicted the responses to chemotherapy. To our knowledge, this is the first report concerning the prediction of a chemotherapeutic response of esophageal cancer by GEP using biopsies.

There are several methods to assess the response to chemotherapy and to divide it into two response groups. In ESCC patients who underwent NACT and surgery, clinical responders, defined by >50% regression in the area of primary tumor, had a better prognosis than non-responders (21,26). Hence, in this study, patients with a reduction rate of >50% of the primary site were classified as responders, and patients with <50% decrease were classified as non-responders.

In the estimation set, although HCA with all 19,166 genes showed that a couple of biopsy samples obtained from the same case were most closely clustered together and genes that had a wide variation in the same case were removed, in three cases one of the sample pairs predicted the patient would respond while the other predicted the patient would not. This is partly because genes that had wide variation between a pair of samples were not able to be completely removed and the cut-off criteria of genes may be determined using a larger set of samples. In the validation set, the diagnostic system correctly predicted the response in eight of ten cases. Though further analysis should be done, these results would allow for individualized treatment for advanced ESCC, which is especially important for reducing the number of non-responders, whose prognosis seems to be worse than that of patients treated by surgery alone (2,27). This is partly because of therapy-induced side-effects, selection of chemotherapy-resistant, biologically more aggressive tumor cells, and delay of surgical treatment. Failure to respond to chemotherapy may be a marker of a biologically aggressive tumor, which is associated with a poor prognosis irrespective of the applied therapy.

Indeed, predictive 199 genes contained some genes that had previously been reported to be associated with drug sensitivity or malignant potential. *PERP* is a novel type of effector involved in p53-dependent apoptosis (28). In both the study of Luthra *et al* and our own, expression values of *PERP* were significantly higher in responders, even though the treatment regimens differed (17). *PRDX6* is one of the molecules involved in redox metabolism. Castagna *et al* reported that *PRDX6* was strongly up-regulated in cervix squamous cell carcinoma cell line after cisplatin exposure by proteomic approach (29). *SRF* encodes a ubiquitous nuclear protein that stimulates both cell proliferation and differentiation. Chai *et al* reported that *SRF* is a downstream mediator of VEGF signaling in endothelial cells and is critically required for VEGF-induced angiogenesis, which is essential for cancer growth (30). Defender against apoptotic cell death (*DAD1*) was initially identified as a negative

regulator of programmed cell death in a mutant hamster cell line. The expression of DAD1 was higher in hepatocellular carcinoma (HCC) tissues than in adjacent non-tumorous tissues by Northern blot analysis. High expression of DAD1 in HCC cells can block apoptosis, thereby enhancing tumor cell survival (31).

The expression patterns of five genes among the predictive 199 genes were confirmed by qRT-PCR, and were in agreement with their expressions on the microarray. This result indicates that combination of biomarkers analyzed by qRT-PCR which is less complicated in manner than microarray has the potential to prediction of chemosensitivity.

In conclusion, this study shows that GEP of pretreatment ESCC biopsy samples can classify tumors according to their future response to chemotherapy. The encouraging results of this study warrant validation of the predictive genes with a larger number of patients and eventually in a prospective trial to evaluate its utility.

Acknowledgements

We would like to thank Ms. Masayo Yamauchi for her valuable technical assistance. This study was partly supported by the Yasuda Medical Research Foundation.

References

- Akiyama H, Tsurumaru M, Udagawa H and Kajiyama Y: Radical lymph node dissection for cancer of the thoracic esophagus. *Ann Surg* 220: 364-372, 1994.
- Medical Research Council Oesophageal Cancer Working Group: Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. *Lancet* 359: 1727-1733, 2002.
- Kelsen DP, Ginsberg R, Pajak TF, *et al.*: Chemotherapy followed by surgery compared with surgery alone for localized esophageal cancer. *N Engl J Med* 339: 1979-1984, 1998.
- Shimada Y, Watanabe G, Yamasaki S, *et al.*: Histological response of cisplatin predicts patients' survival in oesophageal cancer and p53 protein accumulation in pretreatment biopsy is associated with cisplatin sensitivity. *Eur J Cancer* 36: 987-993, 2000.
- Miyata H, Doki Y, Shiozaki H, *et al.*: CDC25B and p53 are independently implicated in radiation sensitivity for human esophageal cancers. *Clin Cancer Res* 6: 4859-4865, 2000.
- Kishi K, Doki Y, Miyata H, Yano M, Yasuda T and Monden M: Prediction of the response to chemoradiation and prognosis in oesophageal squamous cancer. *Br J Surg* 89: 597-603, 2002.
- Langer R, Specht K, Becker K, *et al.*: Association of pretherapeutic expression of chemotherapy-related genes with response to neoadjuvant chemotherapy in Barrett carcinoma. *Clin Cancer Res* 11: 7462-7469, 2005.
- Ponder BAJ: Cancer genetics. *Nature* 411: 336-341, 2001.
- Ramaswamy S and Golub TR: DNA microarrays in clinical oncology. *J Clin Oncol* 20: 1932-1941, 2002.
- Golub TR, Slonim DK, Tamayo P, *et al.*: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286: 531-537, 1999.
- van't Veer LJ, Dai H, van de Vijver MJ, *et al.*: Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415: 530-536, 2002.
- Beer DG, Kardia SL, Huang CC, *et al.*: Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 8: 816-824, 2002.
- Iizuka N, Oka M, Yamada-Okabe H, *et al.*: Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. *Lancet* 361: 923-929, 2003.
- Motoori M, Takemasa I, Yano M, *et al.*: Prediction of recurrence in advanced gastric cancer patients after curative resection by gene expression profiling. *Int J Cancer* 14: 963-968, 2005.
- Ayers M, Symmans WF, Stec J, *et al.*: Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J Clin Oncol* 22: 2284-2293, 2004.
- Clarke PA, George ML, Easdale S, *et al.*: Molecular pharmacology of cancer therapy in human colorectal cancer by gene expression profiling. *Cancer Res* 63: 6855-6863, 2003.
- Luthra R, Wu TT, Luthra MG, *et al.*: Gene expression profiling of localized esophageal carcinomas: association with pathologic response to preoperative chemoradiation. *J Clin Oncol* 24: 259-267, 2006.
- Komori T, Takemasa I, Yamasaki M, *et al.*: Gene expression of colorectal cancer: preoperative genetic diagnosis using endoscopic biopsies. *Int J Oncol* 32: 367-375, 2008.
- Motoori M, Takemasa I, Yamasaki M, *et al.*: The feasibility of using biopsy samples from esophageal cancer for comprehensive gene expression profiling. *Int J Oncol* 35: 265-271, 2009.
- Akita H, Doki Y, Miyata H, *et al.*: Clinical significance of the second cycle response to cisplatin-based chemotherapy as preoperative treatment for esophageal squamous cell carcinoma. *J Surg Oncol* 93: 401-409, 2006.
- Matsuyama J, Doki Y, Yasuda T, *et al.*: The effect of neoadjuvant chemotherapy on lymph node micrometastases in squamous cell carcinomas of the thoracic esophagus. *Surgery* 141: 570-580, 2007.
- Weber WA, Ott K, Becker K, *et al.*: Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging. *J Clin Oncol* 19: 3058-3065, 2001.
- Wieder HA, Brücher BL, Zimmermann F, *et al.*: Time course of tumor metabolic activity during chemoradiotherapy of esophageal squamous cell carcinoma and response to treatment. *J Clin Oncol* 22: 900-908, 2004.
- Levine EA, Farmer MR, Clark P, *et al.*: Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography (18F-FDG-PET) in the identification of responders to chemoradiation therapy for the treatment of locally advanced esophageal cancer. *Ann Surg* 243: 472-478, 2006.
- Duong C, Greenawalt DM, Kowalczyk A, *et al.*: Pretreatment gene expression profiles can be used to predict response to neoadjuvant chemoradiotherapy in esophageal cancer. *Ann Surg Oncol* 14: 3602-3609, 2007.
- Darnton SJ, Archer VR, Stocken DD, Mulholland PJ, Casson AG and Ferry DR: Preoperative mitomycin, ifosfamide, and cisplatin followed by esophagectomy in squamous cell carcinoma of the esophagus: pathologic complete response induced by chemotherapy leads to long-term survival. *J Clin Oncol* 21: 4009-4015, 2003.
- Law S, Fok M, Chow S, Chu KM and Wong J: Preoperative chemotherapy versus surgical therapy alone for squamous cell carcinoma of the esophagus: a prospective randomized trial. *J Thorac Cardiovasc Surg* 114: 210-217, 1997.
- Ihrie RA, Reczek E, Horner JS, Khachatrian L, Sage J, Jacks T and Attardi LD: Perp is a mediator of p53-dependent apoptosis in diverse cell types. *Curr Biol* 13: 1985-1990, 2003.
- Castagna A, Antonioli P, Astner H, *et al.*: A proteomic approach to cisplatin resistance in the cervix squamous cell carcinoma cell line A431. *Proteomics* 4: 3246-3267, 2004.
- Chai J, Baatar D and Tarnawski A: Serum response factor promotes re-epithelialization and muscular structure restoration during gastric ulcer healing. *Gastroenterology* 126: 1809-1818, 2004.
- Tanaka K, Kondoh N, Shuda M, *et al.*: Enhanced expression of mRNAs of antisecretory factor-1, gp96, DAD1 and CDC34 in human hepatocellular carcinomas. *Biochim Biophys Acta* 1536: 1-12, 2001.