Forkhead box A1 transcriptional pathway in KRT7-expressing esophageal squamous cell carcinomas with extensive lymph node metastasis

MASAYUKI SANO[1,5], KAZUHIKO AOYAGI[1], HIRO TAKAHASHI[4,5,7], TAKESHI KAWAMURA[7], TOMOKO MABUCHI[1], HIROYASU IGAKI[2], YUJI TACHIMORI[2], HOICHI KATO[2], ATSUSHI OCHIAI[3], HIROYUKI HONDA[7], YUJI NIMURA[8], MASATO NAGINO[6], TERUHIKO YOSHIDA[1] and HIROKI SASAKI[1]

[1]Genetics Division and [2]Department of Surgery, National Cancer Center Research Institute and Central Hospital, Department of Surgery, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045; [3]Pathology Division, Research Cancer for Innovative Oncology, National Cancer Center Hospital East, 6-3-1 Kashiwanoha, Kashiwa, Chiba 277-8577; [4]College of Bioscience and Biotechnology, [5]Plant Biology Research Center, Chubu University, Matsumoto-cho 1200, Kasugai, Aichi 487-8501; [6]Division of Surgical Oncology, Department of Surgery, Nagoya University Graduate School of Medicine, Tsurumai, Showa-ku, [7]Department of Biotechnology, Nagoya University Graduate School of Engineering, Furo-cho, Chikusa-ku, Nagoya 464-8603; [8]Aichi Cancer Center, 1-1 Kakodono, Chikusa-ku, Nagoya 464-8681, Japan

Received September 7, 2009; Accepted October 18, 2009

DOI: 10.3892/ijo_00000503

Abstract. Prognosis of cancers with lymph node metastasis is known to be very poor; however, it is still controversial whether metastatic potential can be evaluated by expression profiles of primary tumors. Therefore, to address this issue, we compared gene expression profiles of 24 esophageal squamous cell carcinomas (ESCCs) with extensive lymph node metastasis and 11 ESCCs with no metastatic lymph node. However, there was no gene cluster distinguishing these two groups, suggesting that lymph node metastasis-associated genes are varied depending on cases or subgroups. Therefore, we applied a recently developed filtering method (S2N') to identify such genes, and successfully extracted 209 genes associated with node status. Among them, over-expression of *CALB1*, *KRT7*/*CK7*, *MUC1* and *CEA*/*CEACAM5* in poor prognostic cases with metastatic lymph nodes was confirmed in two sets of ESCCs by RT-PCR. Each often seemed to have glandular cell type-characteristics in both the gene expression and morphology. It was also revealed that *FOXA1* siRNA treatment of esophageal cancer cells reduced the mRNA level of both *KRT7* and a stabilizer of epithelial-mesenchymal transition (EMT) regulator *LOXL2*, and that both *FOXA1* and *LOXL2* siRNAs reduced invasion and migration of ESCC cells. In 15 *KRT7*-expressing ESCCs with metastatic lymph nodes, 60% expressed *FOXA1* and 33% expressed both *FOXA1* and *LOXL2*. These results suggest that FOXA1 induces not only *KRT7* but also *LOXL2* in a subset of poor prognostic ESCCs with metastatic lymph nodes, and it is also plausible, that other FOXA1 downstream genes could be therapeutic targets of poor prognostic ESCCs.

Correspondence to: Dr Hiroki Sasaki, Genetics Division, National Cancer Center Research Institute, 1-1, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
E-mail: hksasaki@ncc.go.jp

Key words: microarray, FOXA1, KRT7/CK7, LOXL2, esophageal cancer

Introduction

Cancer is a major cause of human deaths world-wide. Gene expression data from DNA microarrays are individualized and useful in the diagnosis and prognosis of diseases (1). Esophageal cancer is the eighth most common cancer and the sixth most common cause of cancer-related mortality (2). Esophageal cancer in East Asian countries and some parts of Europe consists mainly of squamous cell carcinomas. Chemoradiotherapy (CRT) followed by surgery is the standard therapy in Western countries. In Japan, CRT or neoadjuvant chemotherapy followed by surgery and definitive CRT are the standard therapies (3). For locally advanced esophageal cancers, surgery is still the standard therapy in Japan. A recent improvement in surgical resection following radical node dissection has been reported with 5-year survival rates of 31-55% (4). Yet also in this cancer, lymph node metastasis has been reported to be a most strong marker for poor prognosis in patients with the improved surgery (4), especially, those patients with >5 metastatic lymph nodes did very poorly compared with patients with no metastatic lymph nodes. Thus, lymph node metastasis is known to be tightly associated with a poor prognosis in many surgically resectable solid tumors. Identification of genes associated with lymph node metastasis is very important for

establishing a molecular diagnosis and also for understanding the malignant phenotype; however, it is still controversial whether metastatic potential can be evaluated by expression profiles of primary tumors. Esophageal cancer provides an opportunity to address such an important issue.

For a marker gene selection from mRNA expression profiles, the use of one of the many filtering methods is necessary, such as Mann-Whitney's U-test, Student's t-test, Welch's t-test, signal-to-noise (S2N) (5), significance analysis of microarrays (SAM) (6), and nearest shrunken centroids (NSC) (7). In our previous study, we developed the projective adaptive resonance theory (PART) filtering method (8) and reported that the PART filtering method showed a better performance than conventional methods such as S2N and NSC (9-14). By the PART method, the genes that have a low variance in the gene expression level in either of two classes or subgroups can be selected. We further developed another simple and practical filtering method, modified S2N (S2N) (12), based on the concept underlying the PART filtering method. The S2N' filtering method was statistically superior to the conventional methods such as Mann-Whitney's U-test, Student's t-test, Welch's t-test, S2N, SAM and NSC (12).

In the present study, we applied S2N' to genome-wide gene expression profiles to identify marker genes for poor prognostic esophageal squamous cell carcinomas (ESCCs) with lymph node metastasis, and successfully identified FOXA1 transcriptional pathways for cell invasion or migration in a subset of such ESCCs.

Materials and methods

Tissue samples. All cases of esophageal cancers examined in this study were diagnosed as squamous cell carcinoma. All esophageal squamous cell carcinoma (ESCC) patients underwent surgical resection following two- or three-field node dissection between 1994 and 2000 at the Central Hospital of the National Cancer Center. ESCC tissues were provided after obtaining informed consent from each patient and approval by the Center's Ethics Committee.

Microarray analysis. Gene expression profile data were obtained from 35 surgical specimens from ESCC patients: 24 patients with >5 metastatic lymph nodes (N5 group) and 11 patients with no metastatic lymph node (N0 group). For RNA extraction, trained pathologists carefully excised bulk tissue samples from the main tumor, leaving a clear margin from the surrounding normal tissue. Total RNAs extracted from the bulk tissue samples were biotin-labeled and hybridized to high-density oligonucleotide microarrays (Human Expression Array U95A version 2, Affymetrix, Santa Clara, CA, USA) according to the manufacturer's instructions. The scanned data of the arrays were processed by GeneChip Analysis Suite version 4.0, which scaled the average intensity of all the genes on each array to a target signal of 1,000 to reliably compare variable multiple arrays.

S2N' filtering of gene expression profiles. The S2N' filtering method for gene selection from microarray data was originally developed by us (12). This method allows us to extract as a marker gene even in the case that only a single gene (or a few genes) was expressed specifically in one of the two groups. By this method, we selected the top 209 marker gene candidates which were expressed preferentially in one of the two groups (N0 and N5) in this study.

S*urvival analysis and hierarchical clustering analysis*. The Kaplan-Meier metastases analysis plots were formulated using WinSTAT Statistics for Windows ver. 3.1 (Light Stone, Tokyo, Japan). The significance of the difference in survival rates was analyzed using a log-rank test (Mantel-Cox method). Hierarchical clustering is widely used as one of the unsupervised learning methods. In this study, hierarchical clustering of microarray data of 35 ESCCs was performed by the use of CLUSTER software (15).

Semi-quantitative and quantitative RT-PCR. Surgical specimens were snap-frozen in liquid nitrogen. Total RNA was isolated by suspending the cells in Isogen lysis buffer (Nippon Gene, Toyama, Japan) followed by precipitation with isopropanol. RT-PCR was carried out using primer sets designed for detecting the 3' side of cDNA of each gene: for *CALB1*, 5'-AATCAAAATGTGTGGGAAAG-3' and 5'-CCCAG CACGAGAATAAGAG-3', for *TFF3*, 5'-CTGAGGCACCT CCAGCTGCCCCCG-3' and 5'-GGAGCATGGGACCTT TTTCG-3', for *KRT7*, 5'-CTGAAGGCTTATTCCATCCG-3' and 5'-CCTCAGAATGAGGCTGCTTT-3', for *CLDN10*, 5'-ACGGCTAACTGGATAACTGA-3' and 5'-TGTCTCACT CACTCCTCACA-3', for *MUC1*, 5'-AATTGACTCTGG CCTTCCGA-3' and 5'-GCCACCATTACCTGCAGAAA-3', for *CEA*, 5'-AGACTCTGACCAGAGATCGA-3' and 5'-GGT GGACAGTTTCATGAAGC-3', for *PGC*, 5'-CCAGCTTGA CCTTCATCATC-3' and 5'-GCTGAATCCAGAGTGGA AAG-3', for *LIPF*, 5'-ACGAGTCGCTTGGATGTGTA-3' and 5'-CGTTCCACACTGCAATTGGT-3', for *LOXL2*, 5'-CACC ATGTGTCATCACAGAC-3' and 5'-CTCCTTAGATTGCTT CTCCC-3', for *FOXA1*, 5'-GGTCATGTCATTCTGAG GTC-3' and 5'-TAACACCATGTCCAACTGTG-3', for *ACTB* (ß-actin), 5'-TCATCACCATTGGCAATGAG-3' and 5'-CAC TGTGTTGGCGTACAGGT-3'.

For semi-quantitative RT-PCR, we showed data within linear range by performing 25-35 cycles of PCR. Quantitative real-time PCR was performed by a Bio-Rad iCycler with iQ Syber Green Supermix (Bio-Rad, Hercules, CA, USA) as directed by the manufacturer. The value of 1/2N (N: the number of PCR cycles corresponding to the onset of the linear amplification of each gene product) was calculated as a relative mRNA expression level of each gene normalized to *ACTB*. The data from 2 independent analyses for each sample were averaged.

Immunohistochemistry. All resected specimens were fixed in either 10% formalin or methanol and embedded in paraffin. Tissue sections of 4 μm in thickness were cut from a paraffin-embedded block including the most representative area of the tumor, and were used for immunohistochemical staining. The sections were deparaffinized in xylene, dehydrated in a graded series of ethanol, and immersed in methanol with 0.3% hydrogen peroxide for 15 min to inhibit endogenous peroxidase activity. The slides were heated at 95˚C for 20 min

in a microwave oven in a citrate buffer with pH 6.0 for antigen retrieval, and allowed to cool to room temperature. Non-specific binding was blocked by preincubation with 2% normal swine serum in phosphate-buffered saline (PBS) for 60 min at room temperature. The slides were then incubated overnight at 4˚C with respective primary antibodies. The slides were washed three times with PBS and incubated with EnVision (Dako, Carpinteria, CA, USA) for 1 h at room temperature. The sections were visualized using 3, 3'-diaminobenzidine tetrahydrochloride in 50 mM Tris-buffer (pH 7.6) containing 0.3% hydrogen peroxide as the chromogen, and counterstained by hematoxylin. Antibodies used in this study were anti-CK7 (mouse monoclonal, clone OV-TL, Dako, 1:50).

siRNA transfection. The esophageal cancer cell line TE3 was used in this study because among 6 cell lines (TE1, TE3, TE5, TE6, TE8 and TE10), TE3 has been reported to maintain the expression profile of primary esophageal cancers and to most mimick the basal cell (12). Three or four small interfering (si)RNA fragments were used for suppressing *FOXA1*, *KRT7* and *LOXL2* mRNA expression respectively, and the most effective one was selected by quantitative real-time RT-PCR analysis; for *FOXA1* mRNA target sequence: 5'-GGAGAGATAAGTTATAGGG-3' (siRNA ID: 107428, Ambion, Austin, TX, USA); for *KRT7*: 5'-GGCTGAGATC GACAACAT-3' (siRNA ID: 215171, Ambion); for *LOXL2*: 5'-CCCTCCAGTCTATTATAGT-3' (siRNA ID: 114218, Ambion). The siRNAs including control siRNA (1022076, Qiagen, Valencia, CA, USA) were introduced to TE3 using DharmaFECT (Dharmacon, La Fayette, CO, USA) following the procedure recommended by the manufacturer. The RT-PCR and Matrigel invasion assay were carried out after siRNA treatment of TE3 cells.

Matrigel invasion and wound healing assays. Invasion of the esophageal cancer TE3 cells *in vitro* was measured by BD BioCoat™ Matrigel™ Invasion Chamber (BD Biosciences, San Jose, CA, USA) according to the manufacturer's protocol. After siRNA transfection, the cells were trypsinized and transferred into triplicate wells. After 24-h incubation, the cells that passed through the filter into the lower wells were fixed, stained and counted. For the wound healing assay, TE3 cells were grown till they were 90% confluent. A lineal scratch wound was made on the plate by a plastic tip. Images were taken every 12 h for 2 days.

Results

Survival analysis. The Kaplan-Meier method was performed on two groups: the N0 group consisting of 11 esophageal squamous cell carcinoma (ESCC) patients with no metastatic lymph node and the N5 group consisting of 24 ESCC patients with more than 5 metastatic lymph nodes. All patients underwent surgical resection following two- or three-field node dissection. These two groups showed a significant difference ($p<0.001$) in survival rates after surgical resection (Fig. 1).

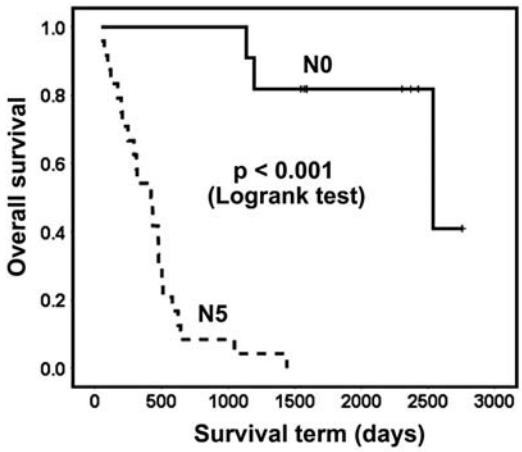

Figure 1. The Kaplan-Meier curve for each group. The N0 group with no metastatic lymph node consisted of 11 patients, and the N5 group with >5 metastatic lymph nodes consisted of 24 patients. The p-value was calculated by log-rank test.

Calculation of S2N' values and ranking of each gene probe. First, we obtained mRNA expression profiles of ESCC tissues from a total of 35 patients of the two groups (N0 and N5). By unsupervised hierarchical clustering using various types of data processing, the two groups were not separated clearly (data not shown). These results implied that no large gene clusters associated with each group were present or that subgroups were present in each group. In this situation, the conventional filtering methods, such as Mann-Whitney's U-test, Student's t-test and Welch's t-test, were thought to be ineffective. In fact, only a few marker genes were selected by these t- or u-test methods (data not shown). The number of selected gene was too small to show significance. Therefore, we applied our previously developed S2N' method for marker gene selection in the two groups of ESCCs. The S2N' values were calculated for each gene probe and we successfully extracted 209 genes which were expressed preferentially in one of two groups. The top 50 genes were shown in Table I. However, this method allows us to extract a marker gene even in the case that only a single gene was expressed specifically in one of the two groups (12). Therefore, we have to further select genes with regards to the frequency of specific expression in one of the two groups. Of the top 50 genes, we selected 8 genes (Rank 1: *CALB1*; Rank 2: *TFF3*; Rank 3: *KRT7*; Rank 9: *CLDN10*; Rank 13: *FOXA1*; Rank 15: *MAL*; Rank 20: *CEACAM5*, and Rank 21: *MUC1*) that were found to be expressed preferentially and frequently in the N5 group. Their signal levels are shown in Fig. 2. These results suggest that our S2N' is very useful as a new method of candidate gene selection for genome-wide gene expression profiles obtained from microarrays.

RT-PCR analysis of 6 candidate marker genes in two sets of ESCCs with more than 5 metastatic lymph nodes or with no metastatic lymph node. Of the 8 genes, 6 (*CALB1*, *TFF3*, *KRT7*/*CK7*, *CLDN10*, *MUC1* and *CEA*/*CEACAM5*) were first analyzed by semi-quantitative RT-PCR in two independent sets of ESCCs: in the first set are 42 samples (12 patients with no metastatic lymph node including 11 N0 patients, and 30 patients with more than 5 metastatic lymph nodes including 24 N5 patients); in the second set are 22 samples

Table I. The top 50 marker gene candidates for the N0 or N5 group selected by the S2N' filtering method.

Probe set ID	GenBank accession	Gene symbol	S2N' rank	S2N'	Marker type
36570_at	AF068862	CALB1	1	14.6	N5
31477_at	L08044	TFF3	2	8.2	N5
41294_at	AJ238246	KRT7	3	6.3	N5
39220_at	T92248	SCGB1A1	4	5.1	N5
38430_at	AA128249	FABP4	5	4.9	N0
174_s_at	U61167	ITSN2	6	4.5	N5
37149_s_at	U95626	LOC728320 /// LTF	7	4.3	N5
34637_f_at	M12963	ADH1A	8	4.0	N5
39579_at	U89916	CLDN10	9	3.7	N5
1497_at	L04270	LTBR	10	2.9	N0
38173_at	AB028999	SETD1B	11	2.7	N5
40193_at	X51956	ENO2	12	2.7	N0
37141_at	U39840	FOXA1	13	2.6	N5
35314_at	D63880	NCAPD2	14	2.6	N0
38051_at	X76220	MAL	15	2.5	N5
34873_at	Y16241	NEBL	16	2.5	N5
1890_at	AB000584	GDF15	17	2.3	N5
38648_at	U80760	ZNF384	18	2.3	N0
1389_at	J03779	MME	19	2.2	N5
1582_at	M29540	CEACAM5	20	2.1	N5
38784_g_at	J05581	MUC1	21	2.1	N5
41308_at	U37408	CTBP1	22	2.0	N5
37576_at	U52969	PCP4	23	2.0	N5
37687_i_at	M31932	FCGR2A	24	1.9	N5
1006_at	X07820	MMP10	25	1.9	N5
37242_at	U79260	FTO	26	1.9	N0
36123_at	D87292	TST	27	1.9	N5
1105_s_at	M12886	TRBV3-1 /// TRBV5-4 /// TRBV7-2	28	1.9	N5
1096_g_at	M28170	CD19	29	1.9	N5
36280_at	U26174	GZMK	30	1.9	N5
35425_at	AJ243512	BARX2	31	1.8	N5
37405_at	U29091	SELENBP1	32	1.8	N5
35227_at	U72066	RBBP8	33	1.7	N5
239_at	M63138	CTSD	34	1.7	N5
39249_at	AB001325	AQP3	35	1.7	N5
1020_s_at	U85611	CIB1	36	1.7	N5
41814_at	M29877	FUCA1	37	1.7	N5
37009_at	AL035079	CAT	38	1.7	N5
37272_at	X57206	ITPKB	39	1.7	N5
31637_s_at	X72631	NR1D1 /// THRA	40	1.7	N5
35693_at	AF070616	HPCAL1	41	1.7	N5
36804_at	M34455	INDO	42	1.7	N5
700_s_at	HG371-HT26388	-[a]	43	1.7	N5
32561_at	D63480	KIAA0146	44	1.6	N0
36023_at	AI864120	-	45	1.6	N0
38783_at	J05581	MUC1	46	1.6	N5
927_s_at	J05582	MUC1	47	1.6	N5
36851_g_at	U42360	TUSC3	48	1.6	N0
35337_at	AL050254	FBXO7	49	1.6	N5
40541_at	X01630	ASS1	50	1.6	N5

[a]An official gene name is not given.

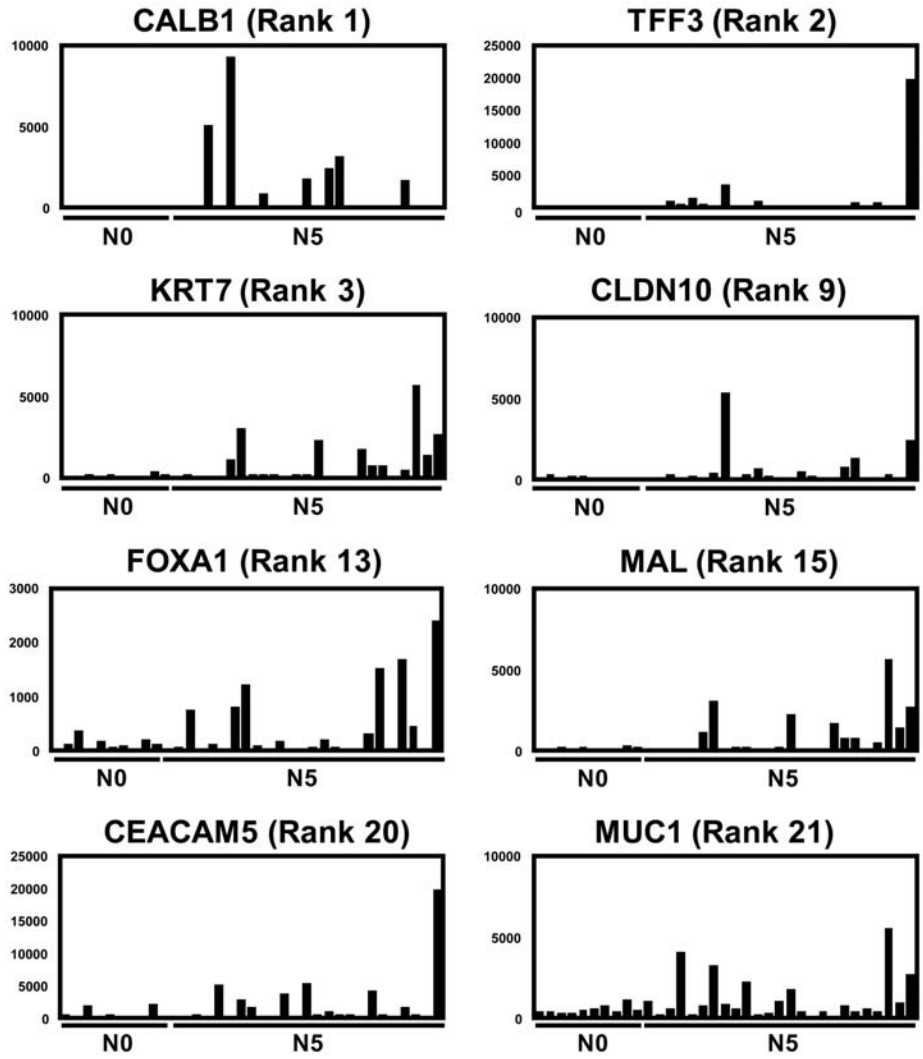

Figure 2. Each signal of microarray for 8 representative top genes whose level was highly correlated with N5 group, respectively.

(10 patients with no metastatic lymph node and 11 patients with >5 metastatic lymph nodes). Similar to the 11 N0 patients and the 24 N5 patients (Fig. 1), Kaplan-Meier analyses of the two groups in each set showed a significant difference in survival after surgical resection (data not shown). In correspondence with microarray data, preferentials expression of all the 6 genes in the N5 ESCC patients with >5 lymph nodes was confirmed in the first set (Fig. 3A). Among these 6 genes, 4 (*CALB1*, *KRT7*, *MUC1* and *CEA*) showed a reproducible expression pattern in the second set (Fig. 3B). These 4 genes preferentially expressed in patients with >5 metastatic lymph nodes in both of the two sets of samples. These results suggest that *CALB1*, *KRT7*, *MUC1* and *CEA* are potential molecular markers for ESCC with poor prognosis, and that the present gene list should provide other marker genes. Among the 6 candidate marker genes identified here, *KRT7* seemed to be one of the two best markers, *KRT7* and *CEA* for ESCC with poor prognosis, because it expressed in a few of the 22 patients with no metastatic lymph node, but in >50% of 42 patients with >5 metastatic lymph nodes (semi-quantitative RT-PCR in Fig. 3A and B). These results were confirmed by quantitative real-time RT-PCR (Fig. 3C and D).

Biological implications of the candidate marker genes for ESCC with poor prognosis. Among the 6 genes, *CEA* and *TFF3* were known not to be expressed in the squamous epithelium of the esophagus, skin and uterine cervix, but in the non-malignant or malignant glandular epithelium of the stomach and colon (17). Accordingly, a subset of poor prognostic ESCC is thought to express glandular epithelial cell markers. To confirm this hypothesis, we investigated the expression of *PGC*-encoding pepsinogen C and *LIPF*-encoding gastric lipase that are well-known as typical markers for glandular epithelium of the stomach. Three patients with >5 metastatic lymph nodes (2 in the first set and 1 in the second set) expressed both *PGC* and *LIPF*, clearly (Fig. 4A and B). The percentage in 42 patients with >5 metastatic lymph nodes was 7% (3/42). Interestingly, a patient positive with these two stomach markers (Fig. 4A, lane 42) had the most metastatic lymph nodes (49 positive nodes) among the 64 ESCC patients examined.

KRT7 has been reported to be expressed in the ducts of many tissues including the liver, kidney, pancreas and mammary gland (18); however, relating to esophagus, little is known. As shown in Fig. 4C, immunohistochemical analyses showed that KRT7 was expressed in the ducts, but not in the

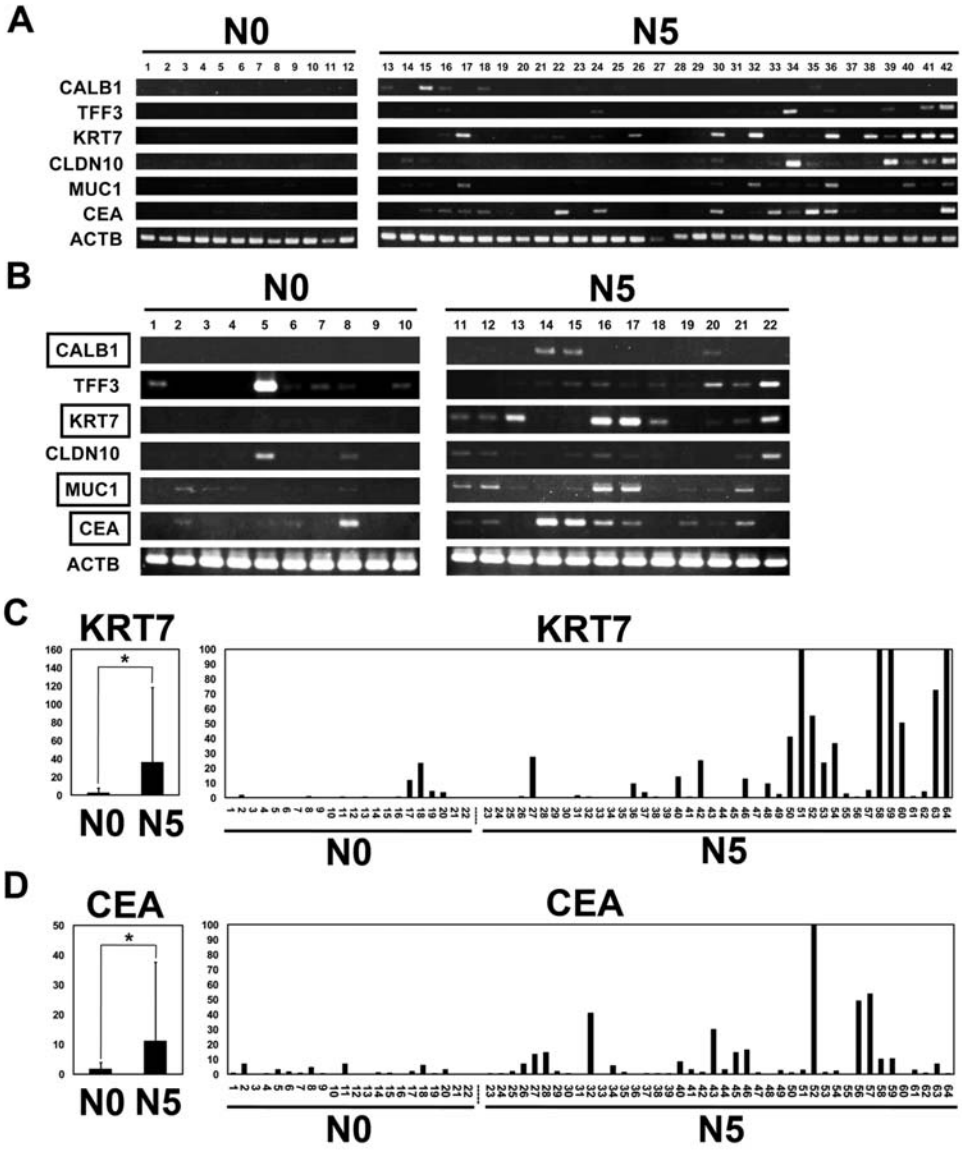

Figure 3. RT-PCR analyses of 6 candidate marker genes for poor prognostic ESCCs with >5 metastatic lymph nodes. (A) Semi-quantitative RT-PCR of 6 poor prognosis marker gene candidates (*CALB1*, *TFF3*, *KRT7*, *CLDN10*, *MUC1* and *CEA*, respectively) in 12 ESCCs with no metastatic lymph node (N0) and in 30 ESCCs with >5 metastatic lymph nodes (N5). (B) Semi-quantitative RT-PCR of these 6 genes in another set of ESCCs consisting of 10 ESCCs with no metastatic lymph node (N0) and in 12 ESCCs with >5 metastatic lymph nodes (N5). Three genes that showed preferential expression in patients with N5 in both the sets of ESCCs are indicated by boxes. (C) Quantitative real-time RT-PCR of *KRT7* in a total of 64 ESCCs consisting of 22 ESCCs with no metastatic lymph node (N0) and in 42 ESCCs with >5 metastatic lymph nodes (N5). (D) Quantitative real-time RT-PCR of *CEA* in these 64 ESCCs. The mean in all samples is indicated by a dotted line.

stratified epithelia of the esophagus. Therefore, KRT7 could be a marker of the esophageal ducts. Those results suggested that a small subset of ESCCs with poor prognosis expressed some markers of the glandular epithelial cell of the stomach, and that a large subset of the ESCCs with poor prognosis showed a feature of the ducts in the esophagus. We next carefully compared the morphological differences between 13 node negative cases and 19 node positive cases (>5 metastatic lymph nodes). In this morphological study, two expert pathologists examined the slides that were prepared by the other pathologist without information on the nodal status. Percentages of the cases with partial glandular structure in each group were averaged. Interestingly, in accordance with the above-mentioned glandular-epithelial cell marker expression in poor prognostic node-positive cases, the partial glandular structure was found preferentially in node-positive cases (74%) compared with node-negative cases (38%) (Fig. 4D).

Forkhead box A1 (FOXA1) is an upstream regulator of KRT7 and LOXL2 is expressed in a subset of poor prognostic ESCC patients. Although *KRT7* was found to be one of the useful markers for poor prognostic ESCCs, the reason why *KRT7*-overexpressing ESCCs are poor is still unknown. We next searched the upstream transcriptional regulator of *KRT7* in the gene list as shown in Table I and found FOXA1 as a transcriptional factor co-expressed with *KRT7* (Fig. 2). To investigate whether *FOXA1* is an upstream regulator for *KRT7*, we introduced siRNA of *FOXA1* and *KRT7* into an esophageal cancer cell line, TE3, that has been confirmed to

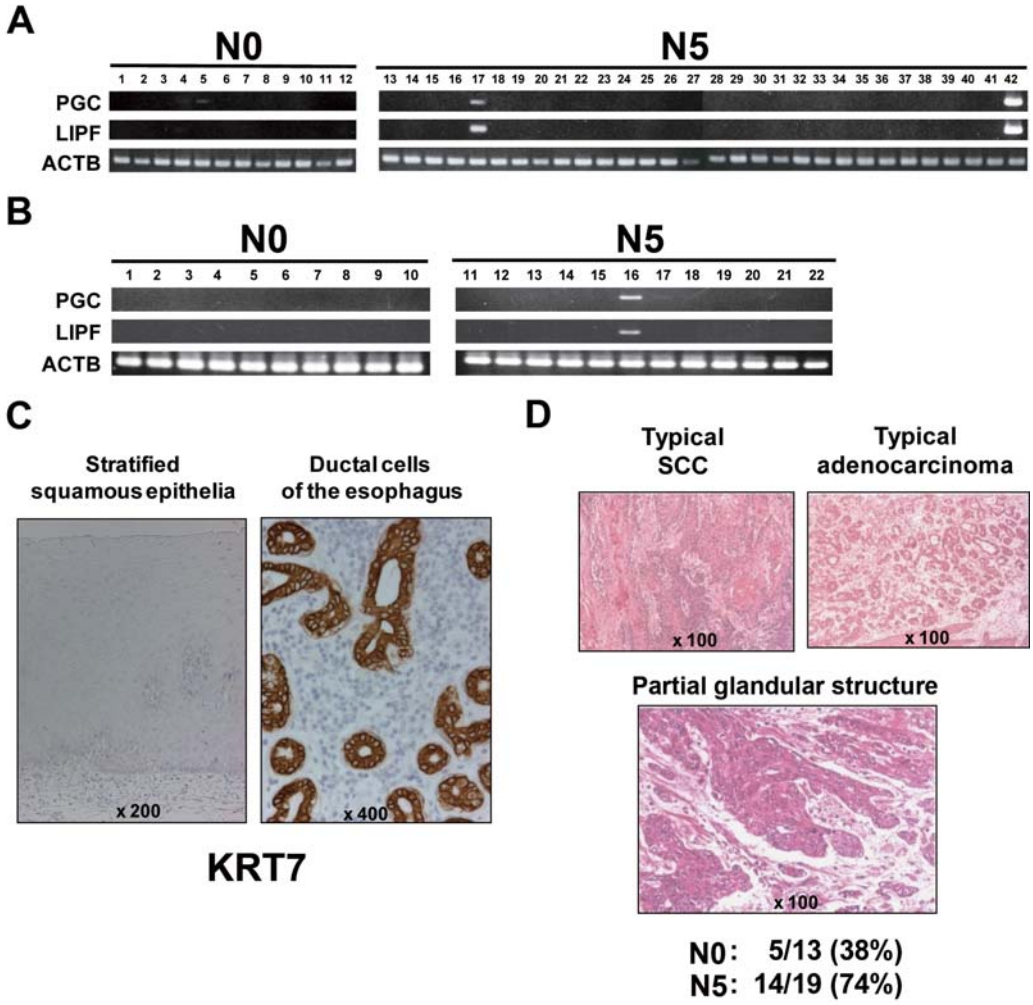

Figure 4. Poor prognostic ESCCs showed characteristics of glandular and/or ductal epithelial cells. (A) RT-PCR of 2 stomach glandular epithelial cell markers (*PGC* and *LIPF*) in the first set of ESCCs. (B) RT-PCR of these 2 stomach glandular epithelial cell markers in the second set of ESCCs. (C) Immunohistochemical staining of KRT7 in normal stratified squamous epithelia and ductal cells in the esophagus. KRT7 is expressed only in ductal cells of the esophagus. (D) Comparison of morphology between N0 and N5 cases. Partial glandular structure was observed preferentially in N0 cases (38%) compared with N5 cases (74%). A typical squamous cell carcinoma (SCC) and adenocarcinoma in the esophagus are shown as a reference (upper).

maintain the expression profile of primary ESCCs the most among 5 cell lines (16) and to express these two genes (data not shown). In this *in vitro* culture, *FOXA1* siRNA successfully reduced the mRNA level of both *FOXA1* and *KRT7*, while *KRT7* siRNA decreased only its own mRNA level (Fig. 5A). We next investigated whether *in vivo* co-expression between *FOXA1* and *KRT7* is observed in primary ESCCs by semi-quantitative RT-PCR. Nine (lanes 16, 17, 22, 34, 36, 39, 40, 41 and 42, respectively) (60%) out of 15 *KRT7* expressing ESCCs with >5 metastatic lymph nodes showed *FOXA1* expression (Fig. 5B). These results suggest that *FOXA1* is an upstream regulator for *KRT7* in a subset of poor prognostic ESCCs.

We next performed genome-wide screening of *FOXA1* downstream genes by microarray analysis of *FOXA1* siRNA-transfected TE3 cells, and found that reduced LOXL2 expression corresponded to a decrease of *FOXA1* mRNA. This result suggests that FOXA1 regulates *LOXL2* expression, which was recently reported as a new poor prognosis marker of laryngeal SCCs (19). To confirm the microarray result, we performed quantitative real-time RT-PCR after transfection of control and *FOXA1* siRNAs into TE3 cells. A significant decrease of *LOXL2* mRNA was found by *FOXA1* siRNA compared with control siRNA (Fig. 5C, left panel). In accordance with the initial *in vitro* study of breast cancer cells (20,21), *LOXL2* siRNA inhibited TE3 cell invasion in Matrigel (Fig. 5C, right panel). We examined the *LOXL2* mRNA expression in primary ESCCs by quantitative real-time RT-PCR. Overexpression of *LOXL2* was observed preferentially in ESCCs with >5 metastatic lymph nodes (Fig. 5D). Five (No. 17, 36, 40, 41, 42) (33%) out of 15 *KRT7* expressing ESCCs with more than 5 metastatic lymph nodes showed both *FOXA1* and LOXL2 expression (Fig. 5B and D). To this end, we investigated whether FOXA1 is also involved in cell migration by wound healing assay. As shown in Fig. 6, *FOXA1* siRNA strongly inhibited TE3 cell migration compared with control siRNA. This effect on of the *FOXA1* siRNA treatment was confirmed not to be dependent on growth inhibition of the treatment (data not shown).

Taken together, this study suggests that FOXA1 induces not only *KRT7* but also *LOXL2* in a subset of poor prognostic

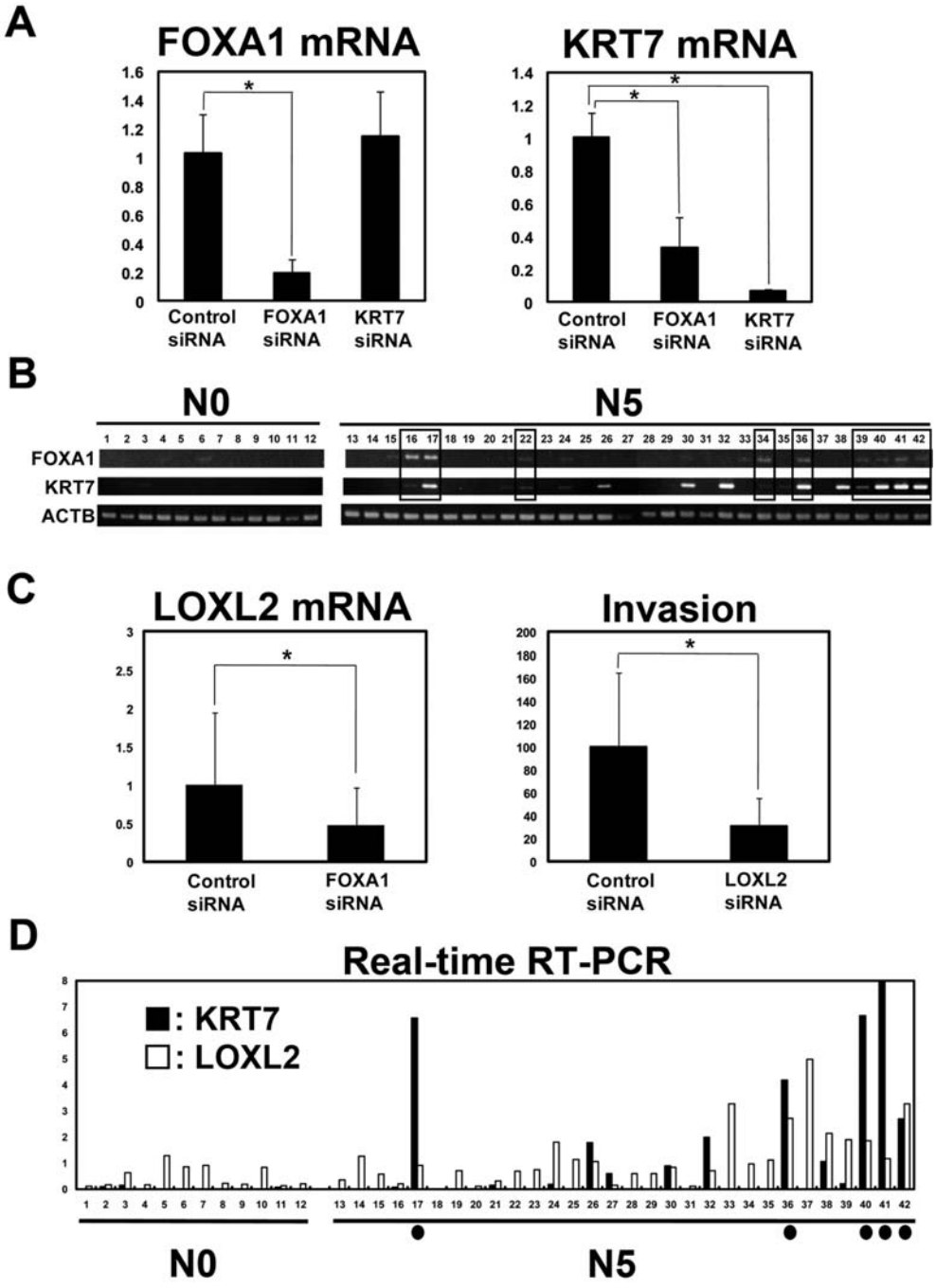

Figure 5. FOXA1 is an upstream regulator of *KRT7* and *LOXL2* in a subset of poor prognostic ESCCs. (A) *FOXA1* siRNA-transfected TE3 cells show reduction of both *FOXA1* and *KRT7* mRNA. Effects on *FOXA1* mRNA (left panel) and *KRT7* mRNA (right panel) after siRNA transfection. (B) Semi-quantitative RT-PCR for showing *in vivo* co-expression between *FOXA1* and *KRT7* in primary ESCCs with poor prognosis. Nine (60%) out 15 *KRT7*-expressing ESCCs with >5 metastatic lymph nodes show *FOXA1* expression (box). (C) *LOXL2* is another target of *FOXA1* in TE3 cells, and is involved in cell invasion. *FOXA1* siRNA-transfected TE3 cells show reduction of *LOXL2* mRNA (left panel). *LOXL2* siRNA treatment inhibits migration of TE3 cells compared with control siRNA (NC). (D) Overexpression of *LOXL2* in primary ESCCs with >5 metastatic lymph nodes is confirmed by quantitative real-time RT-PCR. Cases with co-expression between *KRT7* and *LOXL2* are indicated by closed circle.

ESCCs with metastatic lymph nodes, and that other FOXA1 downstream genes could be therapeutic targets of poor prognostic ESCCs.

Discussion

Based on our present results, we could divide *KRT7*-expressing ESCCs with >5 metastatic lymph nodes into three subgroups (Fig. 7). *KRT7* was found to be overexpressed in 15 (50%) of 30 poor prognostic ESCCs with >5 metastatic lymph nodes (Fig.3A). Out of the 15 *KRT7*-overexpressing ESCCs with poor prognosis, 9 cases (60%) showed *FOXA1* expression (Fig. 5B). Therefore, transcriptional factors other than *FOXA1* (as TFs shown in Fig. 7) must activate *KRT7* and metastasis-associated genes (as Xs shown in Fig. 7) in the 6 remaining cases with KRT7 over-

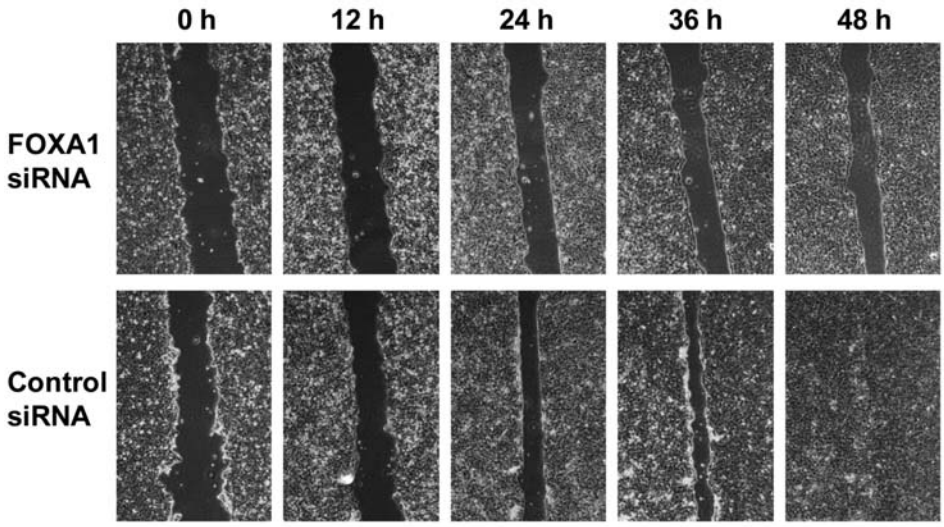

Figure 6. TE3 cell migration after *FOXA1* siRNA treatment. Scratch wound healing assays were performed on TE3 cells after *FOXA1* siRNA and control siRNA transfection. Phase contrast images (original magnification x40) of wound closure at 0, 12, 24, 36 and 48 h, respectively are shown.

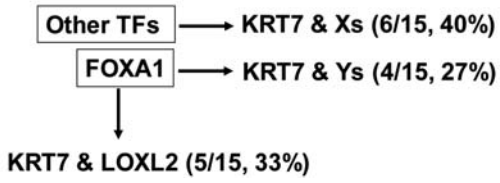

Figure 7. Three subgroups of *KRT7*-expressing ESCCs with poor prognosis. Fifteen of 30 ESCCs with >5 metastatic lymph nodes expressed *KRT7*. The *KRT7*-expressing ESCCs could be divided into three subgroups by the presence of different transcriptional cascades regulating KRT7. TFs, transcription factor regulating *KRT7* other than FOXA1, and Xs and Ys, metastasis-associated genes other than LOXL2.

expression (40%). In five (33%) of the 15 poor prognostic ESCCs, *FOXA1* may regulate both *KRT7* and *LOXL2* (Fig. 5D). Therefore, *FOXA1* must activate not only *KRT7* but also metastasis-associated genes other than *LOXL2* (as Ys shown in Fig. 7) because both FOXA1 and LOXL2 are shown to be involved in cell invasion and migration (Figs. 5 and 6).

In accordance with our results, immunohistochemical studies recently revealed that both LOXL2 and KRT7 are poor prognosis markers of squamous cell carcinomas including ESCCs (22,23). We also recently reported a map of crosstalk between Hedgehog and EMT signaling in ESCCs (16). LOXL2 has been reported to stabilize an EMT regulator SNAI1/SNAIL through physical interaction on the SLUG domain and Snail's lysine residues K98 and K137 (19). In ESCCs, the expression of this EMT regulator should be analyzed in the near future.

The presence of the variable transcriptional cascade in ESCCs may be one of the reasons why a robust gene set, which expresses in association with the prognosis of ESCC patients, could not be extracted by conventional t- or u-test. On this point, our introduced statistical method (S2N') could contribute to the extraction of some genes associated with poor prognosis. However, to understand the molecular mechanisms in lymph node metastasis and to develop a diagnostic method for all the ESCC patients with poor prognosis, the Xs and their upstream transcriptional factors and Ys should be identified. Chromatin immunoprecipitation (ChIP) -on-chip analysis is a potential tool for identifying *in vivo* direct interaction of the target gene promoter with FOXA1. Recently, genome-wide mapping of FOXA1 targets by ChIP-on-chip analysis revealed that FOXA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription and its binding consensus sequences in breast and prostate cancer (24). As shown in Fig. 3, poor prognostic ESCCs often seemed to have glandular cell-type characteristics in both the gene expression and morphology. This fact of transdifferentiation from squamous cell to glandular cell may correspond with cell lineage-specific functions of FOXA1 (24). We were able to find the binding sequences of FOXA1 at -2630 to -2635 (TAGTTTG) of *KRT7* and at -4530 to -4524 (TGTTTAC), -4304 to -4299 (TGTTTGT), -4300 to -4295 (TGTTTGG), and -2642 to -2636 (TGTTTAC) of *LOXL2*. Future studies on the ChIP-on-chip assay with an anti-FOXA1 antibody may contribute to an understanding of the malignancy of ESCCs and to the identification of therapeutic targets. In addition, we showed that aside from *KRT7*, other markers for ESCCs with poor prognosis could be *CALB*, *MUC1* and *CEA* (Fig. 3). Therefore, identification of their upstream transcriptional regulators also remains for future studies.

Acknowledgements

This study was supported in part by the program for promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation; a Grant-in-Aid for the Third Comprehensive 10-Year Strategy for Cancer Control and for Cancer Research (20-12) from the Ministry of Health, Labour and Welfare of Japan; Princess Takamatsu Cancer Research Fund, and Foundation for Promotion of Cancer Research (RRs: M.S. and T.M.).

References

1. Ando T, Suguro M, Hanai T, Kobayashi T, Honda H and Seto M: Fuzzy neural network applied to gene expression profiling for predicting the prognosis of diffuse large B-cell lymphoma. Jpn J Cancer Res 93: 1207-1212, 2002.
2. Parkin DM, Bray F, Ferlay J and Pisani P: Estimating the world cancer burden: Globocan 2000. Int J Cancer 94: 153-156, 2001.
3. Ishikura S, Nihei K, Ohtsu A, Boku N, Hironaka S, Mera K, Muto M, Ogino T and Yoshida S: Long-term toxicity after definitive chemoradiotherapy for squamous cell carcinoma of the thoracicesophagus. J Clin Oncol 21: 2697-2702, 2003.
4. Igaki H, Kato H, Tachimori Y, Sato H, Daiko H and Nakanishi Y: Prognostic evaluation for squamous cell carcinomas of the lower thoracic esophagus treated with three-field lymph node dissection. Eur J Cardiothorac Surg 19: 887-893, 2001.
5. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD and Lander ES: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286: 531-537, 1999.
6. Tusher VG, Tibshirani R and Chu G: Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 98: 5116-5121, 2001.
7. Tibshirani R, Hastie T, Narasimhan B and Chu G: Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci USA 99: 6567-6572, 2002.
8. Takahashi H, Kobayashi T and Honda H: Construction of robust prognostic predictors by using projective adaptive resonance theory as a gene filtering method. Bioinformatics 21: 179-186, 2005.
9. Takahashi H, Aoyagi K, Nakanishi Y, Sasaki H, Yoshida T and Honda H: Classification of intramural metastases and lymph-node metastases of esophageal cancer from gene expression based on boosting and projective adaptive resonance theory. J Biosci Bioeng 102: 46-52, 2006.
10. Takahashi H and Honda H: Lymphoma prognostication from expression profiling using a combination method of boosting and projective adaptive resonance theory. J Chem Eng Jpn 39: 767-71, 2006.
11. Takahashi H, Nemoto T, Yoshida T, Honda H and Hasegawa T: Cancer diagnosis marker extraction for soft tissue sarcomas based on gene expression profiling data by using projective adaptive resonance theory (PART) filtering method. BMC Bioinformatics 7: 399, 2006.
12. Takahashi H and Honda H: Modified signal-to-noise: a new simple and practical gene-filtering approach based on the concept of projective adaptive resonance theory (PART) filtering method. Bioinformatics 22: 1662-1664, 2006.
13. Takahashi H, Murase Y, Kobayashi T and Honda H: New cancer diagnosis modeling using boosting and projective adaptive resonance theory with improved reliable index. Biochem Eng J 33: 100-109, 2007.
14. Kawamura T, Takahashi H and Honda H: Proposal of new gene filtering method, BagPART, for gene expression analysis with small sample. J Biosci Bioeng 105: 81-84, 2008.
15. Eisen MB, Spellman PT, Brown PO and Botstein D: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95: 14863-14868, 1998.
16. Isohata N, Aoyagi K, Mabuchi T, Daiko H, Fukaya M, Ohta H, Ogawa K, Yoshida T and Sasaki H: Hedgehog and epithelial-mesenchymal trasition signaling in normal and malignant epithelial cells of the esophagus. Int J Cancer 125: 1212-1221, 2009.
17. Leung WK, Yu J, Chan FK, To KF, Chan MW, Ebert MP, Ng EK, Chung SC, Malfertheiner P and Sung JJ: Expression of trefoil peptides (TFF1, TFF2, and TFF3) in gastric carcinomas, intestinal metaplasia, and non-neoplastic gastric tissues. J Pathol 197: 582-588, 2002.
18. Smith FJD, Porter RM, Corden LD, Lunny DP, Lane EB and McLean WHI: Cloning of human, murin, and marsupial keratin 7 and a survey of K7 expression in the mouse. Biochem Biophys Res Commun 297: 818-827, 2002.
19. Peinado H, del Carmen M, La Cruz I-D, Olmeda D, Csiszar K, Fong KSK, Vega S, Nieto MA, Cano A and Portillo F: A molecular role for lysyl oxidase-like 2 enzyme in Snail regulation and tumor progression. EMBO J 24: 3446-3458, 2005.
20. Akiri G, Sabo E, Dafni H, Vadasz Z, Kartvelishvily Y, Gan N, Kessler O, Cohen T, Resnick M, Neeman M and Neufeld G: Lysyl oxidase-related protein-1 promotes fibrosis and tumor progression in vivo. Cancer Res 63: 1657-1666, 2003.
21. Kirschmann DA, Seftor EA, Fong SF, Nieva DR, Sullivan CM, Edwards EM, Sommer P, Csiszar K and Hendrix MJ: A molecular role for lysyl oxidase in breast cancer invasion. Cancer Res 62: 4478-4483, 2002.
22. Peinado H, Moreno-Bueno G, Hardisson D, Perez-Gomez E, Santos V, Mendiola M, de Diego JI, Nistal M, Quintanilla M, Portillo F and Cano A: Lysyl oxidase-like 2 as a new poor prognosis marker of squamous cell carcinomas. Cancer Res 68: 4541-4550, 2008.
23. Yamada A, Sasaki H, Aoyagi K, Sano M, Fujii S, Daiko H, Nishimura M, Yoshida T, Chiba T and Ochiai A: Expression of cytokeratin 7 predicts survival in stage I/IIA/IIB squamous cell carcinoma of the esophagus. Oncol Rep 20: 1021-1027, 2008.
24. Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, Carroll JS, Liu XS and Brown M: FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell 132: 958-970, 2008.