[CANCER RESEARCH 63, 5159–5164, August 15, 2003]

Profiling Gene Expression Ratios of Paired Cancerous and Normal Tissue Predicts Relapse of Esophageal Squamous Cell Carcinoma

Yoshio Ishibashi,[1] Nobuyoshi Hanyu, Koji Nakada, Yutaka Suzuki, Takashi Yamamoto, Katsuhiko Yanaga, Kiyoshi Ohkawa, Noriko Hashimoto, Toshiharu Nakajima, Hirohisa Saito, Masato Matsushima, and Mitsuyoshi Urashima[1,2]

Department of Surgery [Y. I., N. Han., K. N., Y. S., T. Y., K. Y.], Department of Biochemistry I [K. O.], and Division of Clinical Research and Development [M. M., M. U.], The Jikei University School of Medicine, Minato-ku, Tokyo 105-8461, and National Center for Child Health and Development [N. Has., T. N., H. S.], Setagaya-ku, Tokyo 154-8567, Japan

ABSTRACT

Esophageal squamous cell carcinoma has heterogeneous clinical outcomes that cannot be predicted well using any existing clinical or molecular prognostic factors. Gene expression profiling may enable more precise prediction of the clinical outcome of these patients. We developed a new approach using gene expression ratios of paired cancerous and normal tissue specimens from the same patient to reduce the effects of variation among individuals. Using oligonucleotide microarrays, we analyzed total RNA expression levels corresponding to 12,600 transcript sequences in 24 paired cancerous and normal tissue operative specimens from 12 patients with esophageal squamous cell carcinoma. Hierarchical clustering analysis using gene expression ratios (cancer:normal) divided the 12 patients into two groups; all 7 patients in the first cluster survived without relapse (median follow-up, 483 days), whereas all 5 patients in the second cluster relapsed (median relapse-free survival time, 280 days; log-rank test, $P = 0.006$). In contrast, clustering either with cancerous tissue alone or with normal tissue alone did not show significant differences in the outcomes. The expressions of a variety of genes related with cell cycle, gene-repair, apoptosis and chemoradiotherpay resistance were up-regulated in the poor prognostic cluster. These results suggest that ratios of paired gene expression profiles may more efficiently predict relapse-free survival of esophageal squamous cell carcinoma than existing prognostic factors or than gene expression profiling with cancerous tissue alone.

INTRODUCTION

In general, squamous cell cancer in the esophagus is an aggressive tumor with poor prognosis, with or without chemoradiotherapy (1). However, the clinical outcomes of esophageal cancer are heterogeneous, and some patients not only with early stage but also with advanced stage cancer may survive longer than clinically expected, and *vice versa*. Yet, only the stage based on TNM classification is widely accepted as a prognostic factor at present (2), which is still far from an accurate predictor. Molecular analyses of esophageal cancer have largely focused on individual candidate genes, with particular emphasis on $p16^{INK4A}$ and *p53* (3), which are no better prognostic factors than existing clinical staging and TNM classification.

The recent development of microarray analysis provides the opportunity to take a genome-wide approach to predict clinical outcomes in a variety of cancers based on the molecular classification of a similar pathological group of cancers with different prognoses (4). However, unlike classical clinical studies, data on thousands of gene expression patterns from a much smaller number of samples must be examined. In addition, whether differential gene expression originates from variation among individuals or from clinical heterogeneity of a given cancer remains to be clarified.

We hypothesized that using the gene expression ratio of paired cancerous and normal tissue samples from the same patient for a given gene could reduce the effects of individuality and, consequently, increase the accuracy of predicting clinical outcomes compared with analysis of cancerous tissue alone.

MATERIALS AND METHODS

Patients. Paired esophageal cancer and normal tissue samples (total 24) were obtained from 12 patients with esophageal cancer after provision of written informed consent, and a protocol approved by the Ethics Committee for Biomedical Research of the Jikei Institutional Review Board, Jikei University School of Medicine, Japan. Eligibility requirements for this study included squamous cell carcinoma of the esophagus that was confirmed by more than two board pathologists and that was free of cancer invasion in the tumor margin and was clinically limited to the locoregional area (no distant metastasis in TNM classification; stage I to III). Tumor stage and grade were classified according to the fifth edition of the TNM classification of the International Union against Cancer (2). In the present study, primary tumors T_1 and $T_{2/3}$ in TNM classification were defined as early (E) and advanced (A), respectively. All patients had adequate general health including hepatic, renal, and bone marrow reserve, and could tolerate the planned surgical procedure. Patients were ineligible if they had Barrett's esophageal adenocarcinoma, stage IV based on TNM classification, presence of other primary cancers, inability of total resection, or prior chemotherapy and/or radiotherapy. On an outpatient basis, the 12 patients were periodically (1–4 months) examined to exclude relapse of the disease using standard examinations including blood examination of tumor markers, chest X-ray, endoscopy, and computed tomography of the chest and abdomen.

Samples. Each cancer tissue sample was divided into two specimens; one for pathological confirmation that the sample was composed of more than 80% cancer cells and the other for RNA extraction. Paired normal tissue was obtained from an area that was at least 4 cm away from any cancerous tissue. Normal tissues were also divided into two specimens: one for pathological confirmation that the sample did not contain either cancer cells or premalignant tissue morphologically. Total RNA was extracted using the RNeasy kit (Qiagen, Chuoh-ku, Tokyo, Japan), and biotinylated cRNAs were then generated. Samples were hybridized onto Affymetrix U95A Array oligonucleotide arrays (Affymetrix, Santa Clara, CA). Arrays were subsequently developed with phycoerythrin-conjugated streptavidin and were scanned to obtain quantitative gene expression levels. Paired normal and cancerous tissue specimens from each patient were handled simultaneously during RNA extraction and hybridization.

Statistics. Genes with expression levels ≤ 0 units in at least one sample were deleted, which resulted in a reduction in the number of genes from 12,626 to 5,843. Gene expression ratios of paired cancerous and normal tissues were simply calculated as unit levels of cancerous tissue (>0 units) divided by unit levels of cancerous tissue (>0 units), which were then transformed with log2:log(cancer/normal) = log(cancer) − log(normal). Next, these were normalized with the *z* value. Hierarchical clustering analysis was then performed using Spotfire software version 7.0 (Spotfire, Somerville, MA). Associations between categorical parameters were examined using Fisher's exact and χ^2 tests. The Kaplan-Meier product-limit method was used to estimate relapse-free survival. For comparison purposes, the Mantel log-rank test was per-

Received 11/25/02; accepted 6/603.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Y.I. and M.U. contributed equally to this work.

[2] To whom requests for reprints should be addressed, at Division of Clinical Research and Development, The Jikei University School of Medicine, 3-25-8 Nishi-shimbashi, Minato-ku, Tokyo 105-8461, Japan. Phone: 81-3-3433-1111, extension 2405; Fax: 81-3-5400-1250; E-mail: urashima@jikei.ac.jp.

formed. Differences between groups were considered significant if the P was <0.05.

RESULTS

Patient characteristics are shown in Table 1. Age ranged from 52 to 82 (mean, 66 ± 10) years, and the majority were males. Patients 8 and 9 were brothers. All of the patients underwent surgery between November 2, 2000, and January, 29, 2002, and were followed for 121–644 days.

The summary of raw data were as follows; mean of normal, 1016; quartile of normal, 140 (25)/312(50)/777(75); mean of cancer, 1102; quartile of cancer, 156(25)/369(50)/937(75). Correlation-of-law data between normal and cancer were 0.86 in 70,116 observations. SDs among 12 normal samples and 12 cancer samples in each gene were compared with show distributions of heterogeneity in normal and cancer expression profiles (Fig. 1). Before attempting to predict the clinical outcome by gene expression profiling, we examined whether hierarchical clustering can distinguish cancerous tissue from normal counterparts. Twenty-four samples containing 12 normal and 12 cancerous tissue samples were analyzed using hierarchical clustering, resulting in two major clusters: (*a*) 12 normal samples and 1 cancer and (*b*) 11 cancer samples. The profiles of the most prominent genes expressed at high levels in cancerous and normal tissue samples are shown in Fig. 2*A* and *B*, respectively. Of the 29 kinds of genes expressed at a higher level in cancerous than in normal tissue (similarity at 2.74 among the gene cluster), 19 (66%) kinds had been previously reported to associate with cellular proliferation and gene repair (Fig. 2*A*). In contrast, one-half of the 38 kinds of genes expressed at higher levels in normal than in cancer tissue (similarity at 2.41 among the gene cluster) have been associated with enzymatic proteins, and only two genes associated with cellular proliferation and gene repair were present (Fig. 2*B*).

Hierarchical clustering analyses were then performed for the 12 paired cancer:normal gene expression ratios, 12 cancerous tissue samples alone, and 12 normal tissue samples alone (Fig. 3), and compared with clinical information. Two distinct clusters were formed when the cancer:normal expression ratio was used (Fig. 3*A*); all of the patients in the first cluster (cluster I) survived without relapse (median follow-up, 483 days), whereas all patients included in the second cluster (cluster II) relapsed (median relapse-free survival time, 280 days; $P = 0.0013$). In contrast, hierarchical clustering using either cancerous tissue alone (Fig. 3*B*) or normal tissue alone (Fig. 3*C*) from the 12 patients was not associated with clinical stages, TNM classification, or relapse. Other pathological prognostic factors (pathological grade of differentiation and lymphovascular invasion) did not show significant association with any clusters obtained by analysis of cancer:normal ratio, cancerous tissue alone, or normal tissue alone, using the two-sided Fisher's exact test.

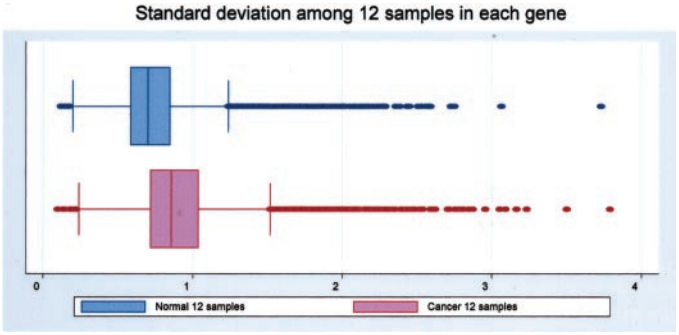

Fig. 1. SDs among 12 samples in each gene were compared between normal and cancer expression profiles. Raw data were transformed with log2, and then SDs in either normal or cancer profiling among 12 patients were calculated in each gene.

The two clusters based on gene expression ratios of cancer:normal tissue also showed a significant difference in relapse-free survival time (log-rank test, $P = 0.006$; Fig. 4*A*). There was also a significant difference in relapse-free survival time between presence (N_1) and absence (N_0) of lymph node metastasis ($P = 0.03$; Fig. 4*C*), but its difference was smaller than that of hierarchical clustering based on gene expression profiling. Other clinical prognostic factors, including stages based on TNM classification (Fig. 4*B*), invasiveness of primary tumor (early, T_1, or advanced, T2 and T3; Fig. 4*D*), grade of differentiation, and degree of lymphovascular invasion, did not make significant difference in relapse-free survival time, although the sample size was too small to conclude that these prognostic factors have no impact on predicting clinical outcomes.

The genes in cluster I with high paired cancer:normal expression ratios (Fig. 5*A*) were classified into several groups based on reported function: mucin and tumor antigen, chromosome (*ubiquitous Kruppel-like factor*, major centromere autoantigen: *CENP-B, histone H1, phosphoribosylpyrophosphate synthetase, CTF5*), signaling to direct cell fate (*Wnt-13, HOX2H, Bcl-2 interacting killer, DRAK2, SIM2, Smad3, RAIG1*), and others. In contrast, those in cluster II were mainly related to ubiquitin-related genes (*ubiquitin-conjugating enzyme, ubiquitin carrier protein, SUMO1*), viral oncogenes (*EBV small RNA-associated protein, 100 kDa EBNA coactivator, E6-AP, papillomavirus*; Fig. 5*B*), cell cycle, and gene repair [*Int-6, dynamin, CBP/p300-interacting transactivator* (*CITED2*), *c-myc binding protein* (*MBP-1*)], *serine/threonine protein kinase, SMARCA5, ckshs2, ku (p70/p80) subunit, MTPC1, hPMS1, RbAp48, PTTG1IP, CUTL1, replication protein 1,*

Table 1 *Characteristics and outcomes of patients*

No.	Age	Gender	Operation	TNM[a]	Stage[b]	Grade[c]	Inv[d]
1	69	Male	4/19/2001	$T_1N_0M_0$	I	II	0
2	70	Male	1/29/2002	$T_1N_1M_0$	IIB	I	1
3	55	Male	11/1/2001	$T_1N_0M_0$	I	I	0
4	66	Male	11/9/2000	$T_1N_0M_0$	I	II	2
5	76	Male	11/22/2001	$T_3N_0M_0$	IIA	II	2
6	58	Male	3/15/2001	$T_1N_1M_0$	IIB	II	0
7	75	Female	4/5/2001	$T_3N_1M_0$	III	I	1
8	79	Male	11/2/2000	$T_3N_1M_0$	III	III	2
9	82	Male	11/4/2000	$T_3N_1M_0$	III	III	2
10	52	Male	11/8/2001	$T_1N_1M_0$	IIB	I	2
11	59	Male	5/31/2001	$T_3N_1M_0$	III	I	2
12	56	Male	12/20/2001	$T_2N_1M_0$	IIB	II	2

[a] T, primary tumor; T_{is}, carcinoma *in situ*; T_1, invading the lamina propria or submucosa; T_2, invading the muscularis propria; T_3, invading the adventitia; T_4, invading adjacent structures; N_0, no regional lymph node metastasis; N_1, regional lymph node metastasis; M, metastasis; M_0, no metastasis; M_1, metastasis.
[b] Stage 0, $T_{is}N_0M_0$; stage I, $T_1N_0M_0$; stage IIA, $T_2N_0M_0$ or $T_3N_0M_0$; stage IIB, $T_1N_1M_0$ or $T_2N_1M_0$; stage III, $T_3N_1M_0$ or T_4 Any N M_0; stage IV, Any T Any N M_1.
[c] Grade I, well differentiated; grade II, moderately differentiated; grade III, poorly differentiated.
[d] Inv, lymphovascular invasion; 0, no invasion; 1, mild invasion; 2, more than moderate invasion.

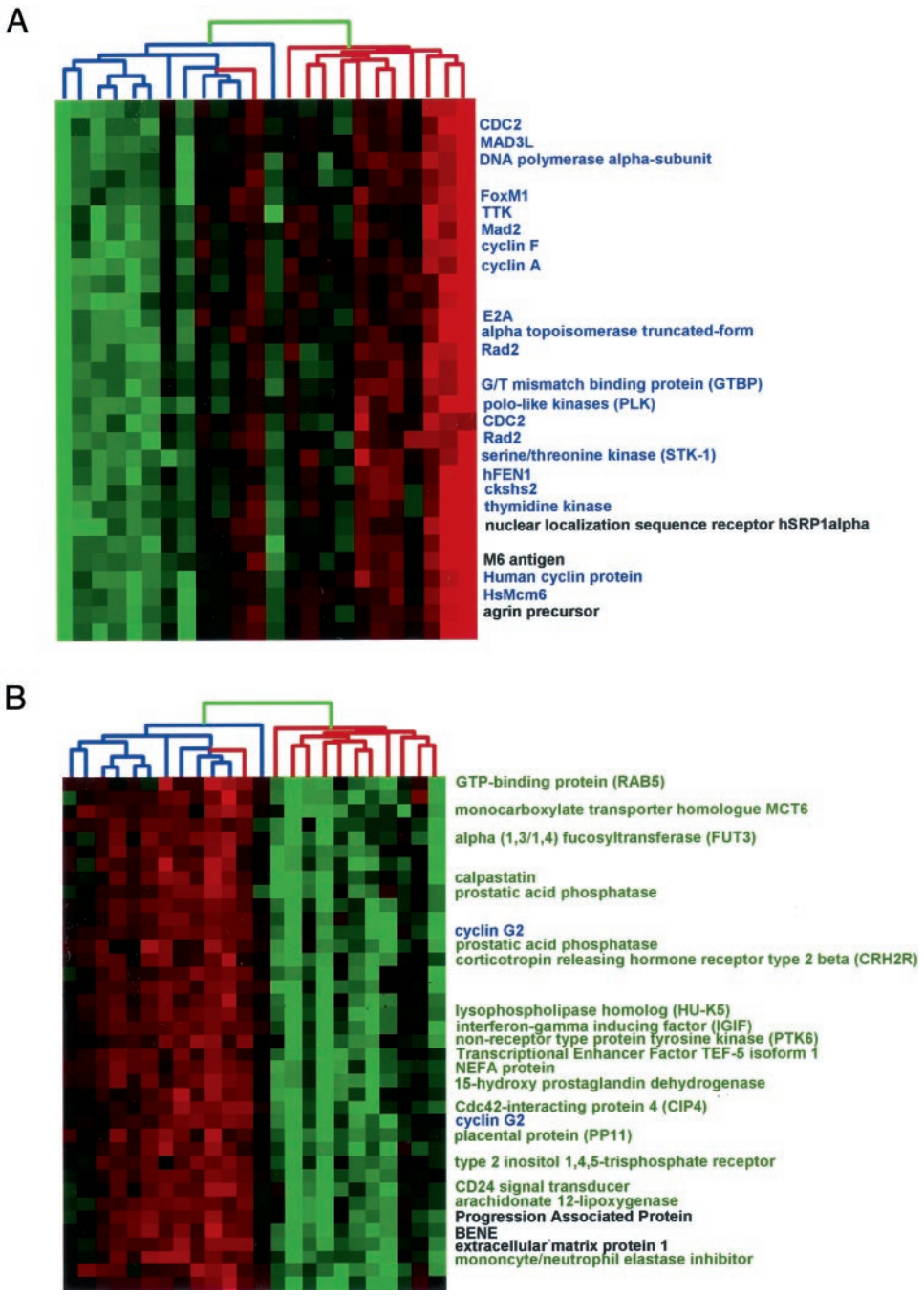

Fig. 2. Hierarchical clustering of 24 specimens composed of 12 pairs of cancer (*red lines in dendrogram at top*) and normal (*blue lines*) tissue samples. *Each column*, a single sample; *each row*, a single gene. *Red areas*, increased RNA expression; *green areas*, decreased RNA expression. *A*, a typical gene profile highly expressed in cancerous tissue (*red*); *B*, a typical gene profile highly expressed in normal esophageal tissue (*green*).

TTK, hMSH2, cAMP-dependent protein kinase, MZF1, TLE1, SPHAR, Centrin3, suppressin, CDK8, casein kinase II, CDC23; Fig. 5*C*). Genes related to resistance to chemoradiotherapy [*superoxide dismutase, MRP*,[3] *p53 binding protein, Rad6*] were also up-regulated in cluster II. Other gene sets [*NXP2, hnRNP A2, RNA polymerase II, Rho guanine nucleotide exchange factor 7, nucleotide binding protein, kinesin-related protein, arginine methyltransferase, histone-binding protein*] high in cluster II may contribute to nucleic acid metabolism. Metastasis-associated *mta1*, metalloproteinase-related collagenase, which was high in cluster II, may facilitate infiltration of tumor growth.

[3] The abbreviations used are: MRP, multidrug resistance-associated protein; HPV, human papillomavirus.

DISCUSSION

Profiling gene expression ratios of paired esophageal squamous cancerous and normal tissue specimens allowed us to predict relapse-free survival time more precisely than did any known clinical prognostic factors or the expression profiles of cancerous and normal tissues from each patient. Although the number of analyzed samples in this study was small and validation studies with larger numbers are required, the present findings do have statistically significant relationships (log-rank test in Kaplan-Meier survival curve, $P = 0.006$) that suggest less possibility of a type I error. Importantly, association with clinical outcomes was shown to be significant not by the gene expression profile of either cancerous tissue alone or normal tissue alone, but by the ratio of cancer to normal tissue specimens from the

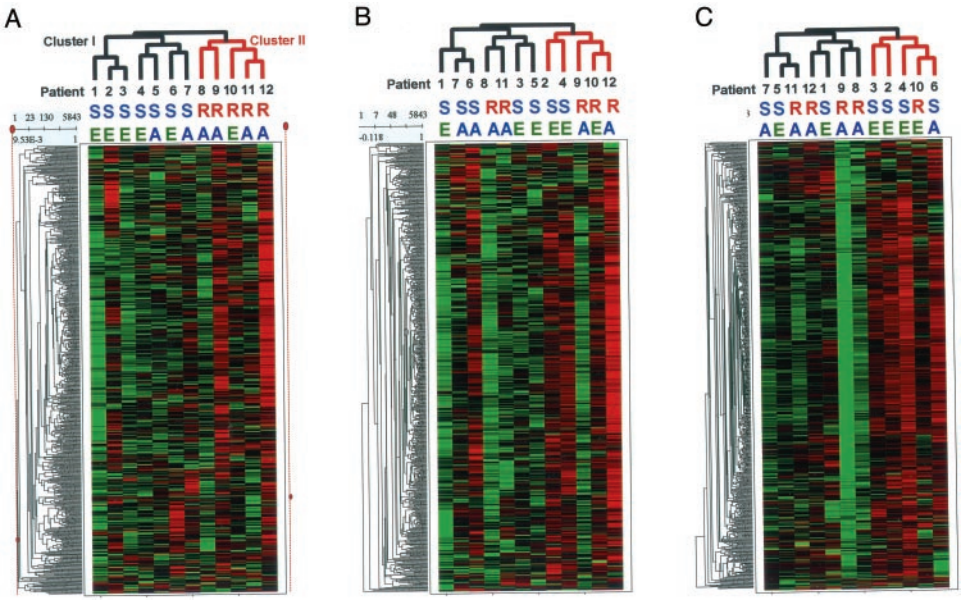

Fig. 3. Hierarchical clustering with dendrogram using the cancer:normal ratio (*A*) or using cancer alone (*B*) or using normal esophageal mucosa alone (*C*). *Branch length*, similarity of distances of samples as judged by their expression patterns. *Red or black*, major clusters. *R*, relapsed; *S*, surviving without relapse; *E*, early esophageal cancer (T_1); *A*, advanced esophageal cancer (T_2 and T_3).

same patient, suggesting that profiling gene expression ratios may provide more informative findings from a huge number of variables with a small number of samples, and reduce the effects of individuality. Comparisons of gene expression profiling between cancerous tissue and paired normal counterparts have been reported in previous studies (5, 6). However, they did not use the ratio of paired expression for analysis but merely used comparison. As shown in Fig. 2*C*, the gene expression profiles in normal esophageal tissue were not homogeneous, thus adjustment using normal counterparts is obligatory to produce an accurate comparison among patients. Differences in gene expression profiles may arise not only from the nature of cancer cell biology but also from individuality as well as from organ-specificity. Paired samples from the same patients share the same genomic DNA and same exposure of esophageal mucosa to hazardous environments such as heavy alcohol drinking and smoking in the same duration and amounts, thus the differences in gene expression ratios most likely originate from the biological nature of the esophageal carcinoma, canceling out individuality. Heterogeneity among individuals was high in both normal and cancer gene expression profiles as shown in Fig. 1. By matching original esophageal tissue, organ-specific RNA expression patterns could also be deleted. Therefore, the ratio of cancer:normal expression is thought to predict clinical outcome quite efficiently.

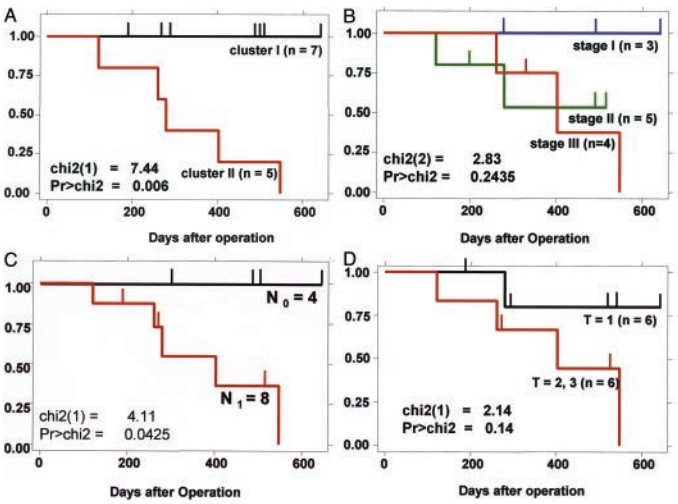

Fig. 4. Kaplan-Meier estimates of relapse-free survival among patients with squamous cell carcinoma of the esophagus days after operation by clusters calculated with gene expression ratios of cancer and normal tissues (*A*), by clinical stages (I, II, III; *B*), by presence of regional lymph node metastasis (N_0, N_1; *C*), and by invasion levels of primary tumor according to TNM classification (early, advanced; *D*).

Dhanasekaran *et al.* (7) examined over 700 prostate-cancer samples using a microarray to determine hepsin and pim-1 proteins as clinical prognostic markers. Bigger sample sizes enable us to stratify patients with known prognostic factors; however, the use of numerous arrays raises problems from an economical point of view. When a study population is narrowed depending on an etiologically identical group, sample size can be dramatically reduced. Hedenfalk *et al.* (8) targeted hereditary breast cancer with mutation at either *BRCA1* ($n = 7$) or *BRCA2* ($n = 7$) and demonstrated different gene expression profiles for the two groups. When investigators tried to abstract the truth from a large dataset obtained from gene expression profiles, they contrived new statistical approaches in addition to hierarchical clustering. Nielsen *et al.* (9) succeeded in molecular classification of soft-tissue tumors by singular value decomposition using 41 samples. For molecular classification of small round blue-cell tumors (SRBCTs), Khan *et al.* (10) applied artificial neural networks using 63 samples (23 biopsies and 40 different cell lines). Supervised machine learning has been used for analysis of gene expression profiling in diffuse large B-cell lymphoma and other kinds of carcinomas (11–13). Others applied epidemiological strategy in combination with hierarchical analysis (14, 15). However, to apply gene expression profiling for general clinical use, the predictability of the clinical outcome has to be superior to existing clinical prognostic factors. The present study has proven the efficiency of gene expression profiling for clinical oncology, as Zhang *et al.* (16) demonstrated with colon cancer using recursive partitioning for the analysis of gene expression profiles.

Many sorts of cell cycle-related genes were highly expressed in our cancerous tissue compared with our normal tissue, which is suspected to be a parallel phenomenon of cell proliferation or a consequence of the tumorigenic process rather than an etiological finding of esophageal carcinoma. Similar phenomena have also been reported in the gene expression profiling of other carcinomas (17, 18). In contrast, profiling gene expression ratios discriminated prognosis of patients efficiently, as described above. These profiles showed different hierarchical clustering from data sets of cancer tissue alone. The enhanced

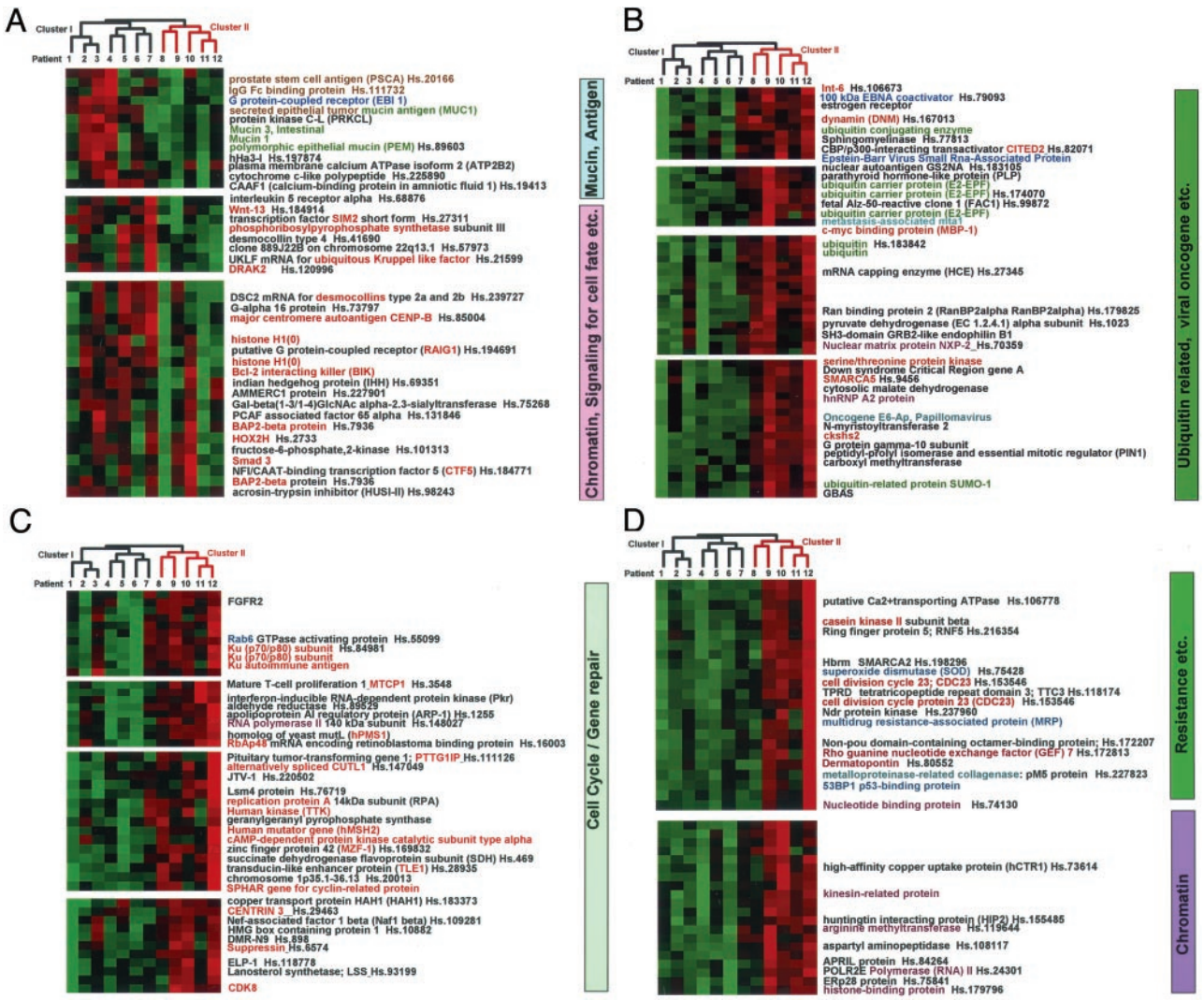

Fig. 5. Profiling gene expression ratios of paired cancerous and normal tissues from the same patients. The typical profiles of genes expressed higher in cluster I, in which no relapse cases were observed, are shown in *A*. The typical profiles of gene expressed higher in cluster II, in which all five cases relapsed, are shown in *B*, *C*, *D*).

expression of mucin-related proteins in the better prognostic group (cluster I) may reflect the existence of normal function as esophageal mucosa. Viral oncoproteins related to EBV and HPV E6-Ap were detected in the poor prognostic group (cluster II). No previous reports have demonstrated expression of the *EBV* gene in esophageal squamous cancer cells, even in a high-prevalence area in China (19). On the contrary, the high frequency of occurrence of HPV in esophageal squamous cell carcinoma implicated HPV as one of the possible etiological factors in this disease (20). However, its prevalence seemed to be lower than first reported, or else the contribution of HPV to the etiology of esophageal cancer differs among geographic areas, being low in European countries and high in Asian counties (21, 22).

Kihara *et al.* (23) applied microarray analysis to 20 esophageal cancer tissue samples to detect a 52-gene set that predicted sensitivity against chemotherapy. In our study, a variety of genes related to gene-repair, apoptosis, and chemoradiotherapy resistance were up-regulated in the poor prognostic group (cluster II). In particular, higher expression of *superoxide dismutase* (*SOD*), *MRP*, and *ubiquitin-conjugating* and *-carrier protein* may be consistent with previous reports (24, 25). *MRP* is well documented to play a role in the mechanism of chemoresistance, and, recently, VX-710 (biricodar; Incel), which restores drug sensitivity to P-glycoprotein (MDR1)- and MRP1-expressing cells, was shown to have efficacy in patients with advanced chemoresistant breast cancer in a Phase II study (26). Up-regulation of ku70/80 may also prevent DNA damage and apoptosis induced by chemoradiotherapy, as reported previously (27). The higher expression of MTA1 in the poor prognostic group (cluster II) in the present study is consistent with a previous report that an esophageal cancer overexpressing MTA1 showed significantly higher frequencies of adventitial invasion and lymph node metastasis (28).

In conclusion, gene expression profiling using the ratio of paired cancer and normal tissue samples from the same patient may more efficiently predict the clinical outcome of esophageal squamous cell cancer and delineate molecules that affect the prognosis of the patients.

REFERENCES

1. Kelsen, D. P., Ginsberg, R., Pajak, T. F., Sheahan, D. G., Gunderson, L., Mortimer, J., Estes, N., Haller, D. G., Ajani, J., Kocha, W., Minsky, B. D., and Roth, J. A. Chemotherapy followed by surgery compared with surgery alone for localized esophageal cancer. N., Engl., J., Med., *339:* 1979–1984, 1998.
2. Sobin, L. H., and Wittekind, C. TNM classification of malignant tumors, Ed. 5. New York: Wiley-Liss, 1997.
3. Sturm, I., Petrowsky, H., Volz, R., Lorenz, M., Radetzki, S., Hillebrand, T., Wolff, G., Hauptmann, S., Dorken, B., and Daniel, P. T. Analysis of p53/BAX/p16(ink4a/CDKN2) in esophageal squamous cell carcinoma: high BAX and p16(ink4a/CDKN2) identifies patients with good prognosis. J., Clin., Oncol., *19:* 2272–2281, 2001.

4. Ahr, A., Karn, T., Solbach, C., Seiter, T., Strebhardt, K., Holtrich, U., and Kaufmann, M. Identification of high risk breast-cancer patients by gene expression profiling. Lancet, *359:* 131–132, 2002.
5. Xu, X-R., Huang, J., Xu, Z-G., Qian, B. Z., Zhu, Z. D., Yan, Q., Cai, T., Zhang, X., Xiao, H. S., Qu, J., Liu, F., Huang, Q. H., Cheng, Z. H., Li, N. G., Du, J. J., Hu, W., Shen, K. T., Lu, G., Fu, G., Zhong, M., Xu, S. H., Gu, W. Y., Huang, W., Zhao, X. T., Hu, G. X., Gu, J. R., Chen, Z., and Han, Z. G. Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profile of hepatocellular carcinoma with those of corresponding noncancerous liver. Proc., Natl., Acad., Sci., USA, *18:* 15089–15094, 2001.
6. Moustafa, A-E., Alaoui-Jamali, M. A., Batist, G., Hernandez-Perez, M., Serruya, C., Alpert, L., Black, M. J., Sladek, R., and Foulkes, W. D. Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal and squamous carcinoma cells. Oncogene, *21:* 2634–2640, 2002.
7. Dhanasekaran, S. M., Barrette, T. R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K. J., Rubin, M. A., and Chinnaiyan, A. M. Delineation of prognostic biomarkers in prostate cancer. Nature (Lond.), *412:* 822–826, 2001.
8. Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O. P., Wilfond, B., Borg, A., and Trent, J. Gene-expression profiles in hereditary breast cancer. N. Engl. J. Med., *344:* 539–548, 2001.
9. Nielsen, T. O., West, R. B., Linn, S. C., Alter, O., Knowling, M. A., O'Connell, J. X., Zhu, S., Fero, M., Sherlock, G., Pollack, J. R., Brown, P. O., Botstein, D., and van de Rijn, M. Molecular characterization of soft tissue tumours: a gene expression study. Lancet, *359:* 1301–1307, 2002.
10. Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C. R., Peterson, C., and Meltzer, P. S. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat. Med., *7:* 673–679, 2001.
11. Shipp, M. A., Ross, K. N., Tamayo, P., Weng, A. P., Kutok, J. L., Aguiar, R. C., Gaasenbeek, M., Angelo, M., Reich, M., Pinkus, G. S., Ray, T. S., Koval, M. A., Last, K. W., Norton, A., Lister, T. A., Mesirov, J., Neuberg, D. S., Lander, E. S., Aster, J. C., and Golub, T. R. Diffuse large B-cell lymphoma outcome prediction by gene expression profiling and supervised machine learning. Nat. Med., *8:* 68–74, 2002.
12. Pomeroy, S. L., Tamayo, P., Gaasenbeek, M., Sturla, L. M., Angelo, M., McLaughlin, M. E., Kim, J. Y., Goumnerova, L. C., Black, P. M., Lau, C., Allen, J. C., Zagzag, D., Olson, J. M., Curran, T., Wetmore, C., Biegel, J. A., Poggio, T., Mukherjee, S., Rifkin, R., Califano, A., Stolovitzky, G., Louis, D. N., Mesirov, J. P., Lander, E. S., and Golub, T. R. Prediction of central nervous system embryonal tumour outcome based on gene expression. Nature (Lond.), *415:* 436–442, 2002.
13. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science (Wash. DC), *286:* 531–537, 1999.
14. van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A., Mao, M., Peterse, H. L., van der Kooy, K., Marton, M. J., Witteveen, A. T., Schreiber, G. J., Kerkhoven, R. M., Roberts, C., Linsley, P. S., Bernards, R., and Friend, S. H. Gene expression profiling predicts clinical outcome of breast cancer. Nature (Lond.), *415:* 530–536, 2002.
15. Rosenwald, A., Wright, G., Chan, W. C., Connors, J. M., Campo, E., Fisher, R. I., Gascoyne, R. D., Muller-Hermelink, H. K., Smeland, E. B., Giltnane, J. M., Hurt, E. M., Zhao, H., Averett, L., Yang, L., Wilson, W. H., Jaffe, E. S., Simon, R., Klausner, R. D., Powell, J., Duffey, P. L., Longo, D. L., Greiner, T. C., Weisenburger, D. D., Sanger, W. G., Dave, B. J., Lynch, J. C., Vose, J., Armitage, J. O., Montserrat. E., Lopez-Guillermo, A., Grogan, T. M., Miller, T. P., LeBlanc, M., Ott, G., Kvaloy, S., Delabie, J., Holte, H., Krajci, P., Stokke, T., and Staudt, L. M. Lymphoma/Leukemia Molecular Profiling Project. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N. Engl. J. Med., *346:* 1937–1947, 2002.
16. Zhang, H., Yu, C. Y., Singer, B., and Xiong, M. Recursive partitioning for tumor classification with gene expression microarray data. Proc. Natl. Acad. Sci. USA., *98:* 6730–6735, 2001.
17. Prakash, K., Pirozzi, G., Elashoff, M., Munger, W., Waga, I., Dhir, R., Kakehi, Y., and Getzenberg, R. H. Symptomatic and asymptomatic benign prostatic hyperplasia: molecular differentiation by using microarrays. Proc. Natl. Acad. Sci. USA, *99:* 7598–7603, 2002.
18. Bhattacharjee, A., Richards, W. G., Staunton, J., Li, C., Monti, S., Vasa, P., Ladd, C., Beheshti, J., Bueno, R., Gillette, M., Loda, M., Weber, G., Mark, E. J., Lander, E. S., Wong, W., Johnson, B. E., Golub, T. R., Sugarbaker, D. J., and Meyerson, M. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc. Natl. Acad. Sci. USA, *98:* 13790–13802, 2001.
19. Chang, F., Syrjanen, S., Shen, Q., Cintorino, M., Santopietro, R., Tosi, P., and Syrjanen, K. Evaluation of HPV, CMV, HSV and EBV in esophageal squamous cell carcinomas from a high-incidence area of China. Anticancer Res., *20:* 3935–3940, 2000.
20. Toh, Y., Kuwano, H., Tanaka, S., Baba, K., Matsuda, H., Sugimachi, K., and Mori, R. Detection of human papillomavirus DNA in esophageal carcinoma in Japan by polymerase chain reaction. Cancer (Phila.), *70:* 2234–2238, 1992.
21. Matsha, T., Erasmus, R., Kafuko, A. B., Mugwanya, D., Stepien, A., and Parker, M. I. Human papillomavirus associated with esophageal cancer. J. Clin. Pathol., *55:* 587–590, 2002.
22. Borysiewicz, L. K., Fiander, A., Nimako, M., Man, S., Wilkinson, G. W., Westmoreland, D., Evans, A. S., Adams, M., Stacey, S. N., Boursnell, M. E., Rutherford, E., Hickling, J. K., and Inglis, S. C. A recombinant vaccinia virus encoding human papillomavirus types 16 and 18, E6 and E7 proteins as immunotherapy for cervical cancer. Lancet, *347:* 1523–1527, 1996.
23. Kihara, C., Tsunoda, T., Tanaka, T., Yamana, H., Furukawa, Y., Ono, K., Kitahara, O., Zembutsu, H., Yanagawa, R., Hirata, K., Takagi, T., and Nakamura, Y. Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy. Cancer Res., *61:* 6474–6485, 2001.
24. Huang, P., Feng, L., Oldham, E. A., Keating, M. J., and Plunkett, W. Superoxide dismutase as a target for the selective killing of cancer cells. Nature (Lond.), *407:* 390–395, 2000.
25. Spataro, V., Norbury, C., and Harris, A. L. The ubiquitin-proteasome pathway in cancer. Br. J. Cancer, *77:* 448–455, 1998.
26. Toppmeyer, D., Seidman, A. D., Pollak, M., Russell, C., Tkaczuk, K., Verma, S., Overmoyer, B., Garg, V., Ette, E., Harding, M. W., and Demetri, G. D. Safety and efficacy of the multidrug resistance inhibitor Incel (biricodar; VX-710) in combination with paclitaxel for advanced breast cancer refractory to paclitaxel. Clin. Cancer Res., *8:* 670–678, 2002.
27. Zhao, H. J., Hosoi, Y., Miyachi, H., Ishii, K., Yoshida, M., Nemoto, K., Takai, Y., Yamada, S., Suzuki, N., and Ono, T. DNA-dependent protein kinase activity correlates with Ku70 expression and radiation sensitivity in esophageal cancer cell lines. Clin. Cancer. Res., *6:* 1073–1078, 2000.
28. Toh, Y., Kuwano, H., Mori, M., Nicolson, G. L., and Sigimachi, K. Overexpression of metastasis-associated MTA1 mRNA in invasive esophageal carcinomas. Br. J. Cancer, *79:* 1723–1726, 1999.

Profiling Gene Expression Ratios of Paired Cancerous and Normal Tissue Predicts Relapse of Esophageal Squamous Cell Carcinoma

Yoshio Ishibashi, Nobuyoshi Hanyu, Koji Nakada, et al.

Cancer Res 2003;63:5159-5164.