Table 4 *Association between clinicopathological features and mRNA expression levels of genes involved in tumor progression and metastasis obtained by serial analysis of gene expression*

Gene name		Case number	mRNA expression level Mean ± SE	P^a
CDH17				
T grade[b]	1/2	20	0.065 ± 0.022	0.0060
	3/4	26	0.275 ± 0.068	
N grade[b]	0	11	0.127 ± 0.054	0.8367
	1/2/3	35	0.201 ± 0.053	
Stage[b]	I/II	18	0.102 ± 0.036	0.2035
	III/IV	28	0.235 ± 0.064	
FUS				
T grade	1/2	20	0.050 ± 0.017	0.5714
	3/4	26	0.062 ± 0.014	
N grade	0	11	0.027 ± 0.016	0.0416
	1/2/3	35	0.066 ± 0.013	
Stage	I/II	18	0.033 ± 0.012	0.0414
	III/IV	28	0.072 ± 0.015	
COL1A1				
T grade	1/2	20	6.84 ± 1.34	0.1407
	3/4	26	20.66 ± 8.14	
N grade	0	11	5.57 ± 1.67	0.1048
	1/2/3	35	17.50 ± 6.11	
Stage	I/II	18	5.32 ± 1.14	0.0156
	III/IV	28	20.64 ± 7.54	
COL1A2				
T grade	1/2	20	10.91 ± 1.86	0.1377
	3/4	26	27.90 ± 9.14	
N grade	0	11	10.45 ± 2.30	0.2572
	1/2/3	35	23.67 ± 6.91	
Stage	I/II	18	9.45 ± 1.54	0.0395
	III/IV	28	27.62 ± 8.49	
APOE				
T grade	1/2	20	6.17 ± 2.90	0.0139
	3/4	26	11.29 ± 2.98	
N grade	0	11	1.36 ± 0.22	0.0006
	1/2/3	35	11.49 ± 2.66	
Stage	I/II	18	6.33 ± 3.22	0.0125
	III/IV	28	10.83 ± 2.78	

[a] Mann-Whitney *U* test.
[b] Tumor staging of gastric carcinoma were done according to the Tumor-Node-Metastasis Stage Grouping (15).

pressed in GC. Quantitative RT-PCR analysis of 27 selected genes showed that *COL1A1*, *CDH17*, *APOC1*, *COL1A2*, *YF13H12*, *CEACAM6*, *APOE*, *REGIV*, *FUS*, and *S100A11* were overexpressed in 40–80% of the 46 GC samples analyzed. Among them, *TFF3*, *REGIV*, and *S100 calcium-binding proteins* have been reported to be commonly up-regulated in GC by other SAGE studies (10, 12).

Among the 27 selected genes, only *COL1A1* and *CDH17* were overexpressed in >70% of the 46 GC samples. *COL1A1* was most frequently overexpressed, and *COL1A2* was also frequently overexpressed as determined by quantitative RT-PCR. Although *COL1A1* expression has been demonstrated in tumor cells and tumor-associated stromal cells in multiple cancers (23, 24), *COL1A1* and *COL1A2* have been reported to be elevated in tumor endothelium as compared with normal endothelium (25), suggesting that they play an important role in angiogenesis and the formation of desmoplasia in GC. In fact, we found a significant association between tumor stage and mRNA expression level for both genes. *CDH17* is a structurally unique member of the cadherin superfamily and is expressed in intestinal epithelial cells (26) and in intestinal metaplasia of the stomach (27). Although overexpression of *CDH17* has been reported in intestinal type GC (27), the association between *CDH17* and tumor invasion has not been examined. In the present study, we showed that the high level of *CDH17* expression was associated with advanced T grade, indicating that *CDH17* is a candidate marker gene for tumor progression. However, a recent study of pancreatic cancer reported that high CDH17 expression correlates with good survival (28). Thus, the significance of the association of high *CDH17* expression and advanced tumor invasion remains unclear. Organ specificity of CDH17 expression may be involved in tumor invasion and progression.

Frequently overexpressed genes in this study included 2 apolipoproteins. *APOC1* was commonly up-regulated in GC, and *APOE* was a candidate marker for tumor metastasis. Although the expression status of these genes has not been previously examined in GC, it has been reported in certain cancers. *APOC1* gene expression localizes to tumor-associated macrophages in breast carcinoma (24). In colorectal carcinoma, intense apolipoprotein E expression has been identified in macrophages surrounding the tumor area (29), suggesting that overexpression of these 2 apolipoproteins occur in tumor-associated macrophages. Macrophages appear to play a pivotal role in tumor angiogenesis, and in our previous observation, macrophage infiltration is significantly associated with tumor vessel density in GC (30). In addition, we found that a high level of *APOE* expression was associated with advanced T grade, N grade, and stage. Apolipoprotein E produced by tumor-associated macrophages may play an important role in tumor progression. Because *APOE* mRNA expression in lymph node metastasis tended to be higher than that in primary GC, *APOE* expression may be up-regulated in GC cells. In prostate cancer, apolipoprotein E expression was identified in cancer cells and correlated directly with Gleason grade (31). Whether GC cells or tumor-associated macrophages express apolipoprotein E remains unclear. Immunohistochemical analysis will be required to answer this question.

S100 calcium-binding proteins (*S100A4*, *S100A9*, *S100A10*, and *S100A11*) were among the 20 up-regulated genes. *S100A4* is commonly up-regulated in GC. In fact, S100A4 expression was detected in 51 of 92 primary GC samples (55%; Ref. 32). Previous SAGE analysis of moderately differentiated GC indicated that 5 calcium-binding proteins (*S100A2*, *S100A7*, *S100A8*, *S100A9*, and *S100A10*) are overexpressed (10). *S100A11* is potentially involved in tumor metastasis. However, no obvious up-regulation of *S100A11* was identified in lymph node metastasis of GC. *S100A11* may be important for stomach carcinogenesis, and overexpression of S100 calcium-binding proteins may be a common alteration in GC.

CEACAM6 is a member of the immunoglobulin superfamily (33) and functions as an intercellular adhesion protein (34). *CEACAM6* overexpression independently predicts poor overall survival and disease-free survival in colorectal carcinoma (35). In GC, although frequent overexpression of *CEACAM6* was identified in the present study, we found no association between the expression levels of *CEACAM6* and clinicopathological features.

Overexpression of 2 genes related to wound-healing was identified in the present study. *TFF3* functions in the maintenance and repair of the intestinal mucosa (36). *TFF3* was commonly up-regulated in GC, and overexpression of *TFF3* in GC has been reported previously (37). *REGIV* was a candidate gene specifically expressed in GC. *REGIV* is a member of the *Reg* gene family, which includes 3 other genes (22). *REGIV* expression is restricted to the gastrointestinal tract and pancreas and is up-regulated in response to mucosal injury in active Crohn's disease and ulcerative colitis (22). It has been reported that *REGIV* expression is increased in most colorectal cancers compared with normal tissues (38). Although overexpression of *REGIV* has been reported by conventional RT-PCR in 6 GC samples (12), the specificity of *REGIV* expression has not been investigated. In our study, Virtual Northern and quantitative RT-PCR analysis showed *REGIV* expression to be narrowly restricted. We performed additional quantitative RT-PCR analysis of 10 colorectal cancers, 10 lung cancers, and 10 breast cancers (data not shown). Although *REGIV* expression was identified in all 10 colorectal cancers, the levels of *REGIV* expression were <100 arbitrary units. We also confirmed that the expression levels of *REGIV* in all 10 colorectal cancers were higher than those in normal colon. No *REGIV* expression was identified in lung or breast cancers. These results are consistent with the Virtual

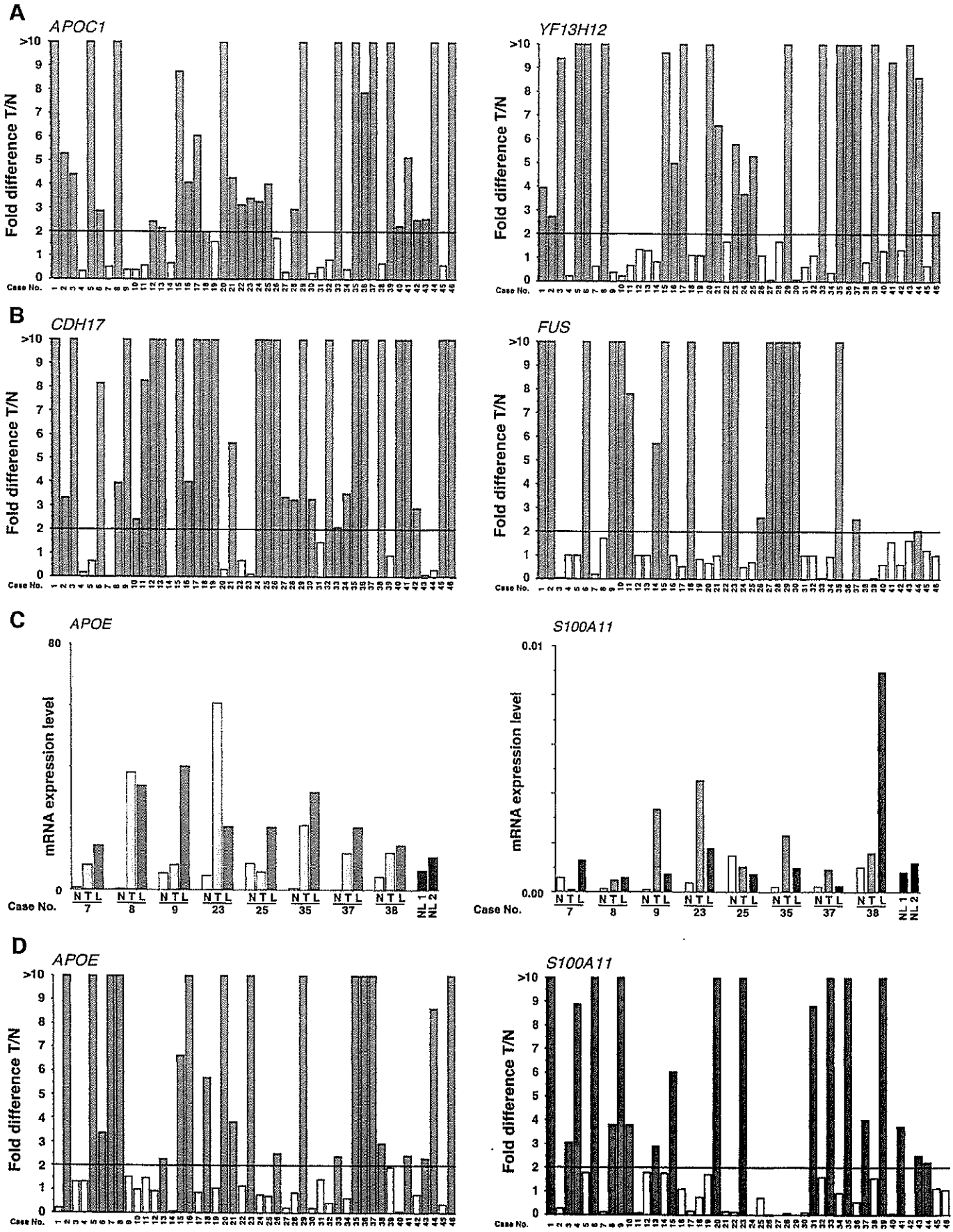

Fig. 3. Validation of serial analysis of gene expression (SAGE) data by quantitative reverse transcription-PCR (RT-PCR). Fold change indicates the ratio of target gene mRNA level in gastric carcinoma (GC) to that in corresponding nonneoplastic mucosa. *A*, quantitative RT-PCR analysis of genes commonly up-regulated according to SAGE analysis. Of the 46

A

Library name	Tags per million	Tag counts	Total tags
SAGE Panc 96 6252 pancreas epithelium ductal adenocarcinoma non-normalized bulk EST	27	1	35750
SAGE SW837 adenocarcinoma colon SAGE CGAP non-normalized SAGE library method cell line	16	1	61061
PC3 AS2	24	1	40768
PC3 Mock	25	1	38819
SAGEHEXO1	56	1	17809
SAGE Hiroshima GC P208T	259	3	11582
SAGE Tu102 adenocarcinoma colon SAGE CGAP non-normalized SAGE library method bulk	69	4	57686
SAGE Hiroshima GC W226T	113	5	43908
SAGE NC1 epithelium normal colon SAGE CGAP non-normalized SAGE library method bulk	199	10	50179
SAGE NC2 epithelium normal colon SAGE CGAP non-normalized SAGE library method bulk	201	10	49593
SAGE Hiroshima GC P208L	1105	17	15382
SAGE Hiroshima GC W246T	714	23	32174
SAGE Hiroshima GC S219T	1933	67	34660
SAGE gastric cancer G189 gastrointestinal tract carcinoma CGAP non-normalized SAGE library method bulk	1482	95	64102

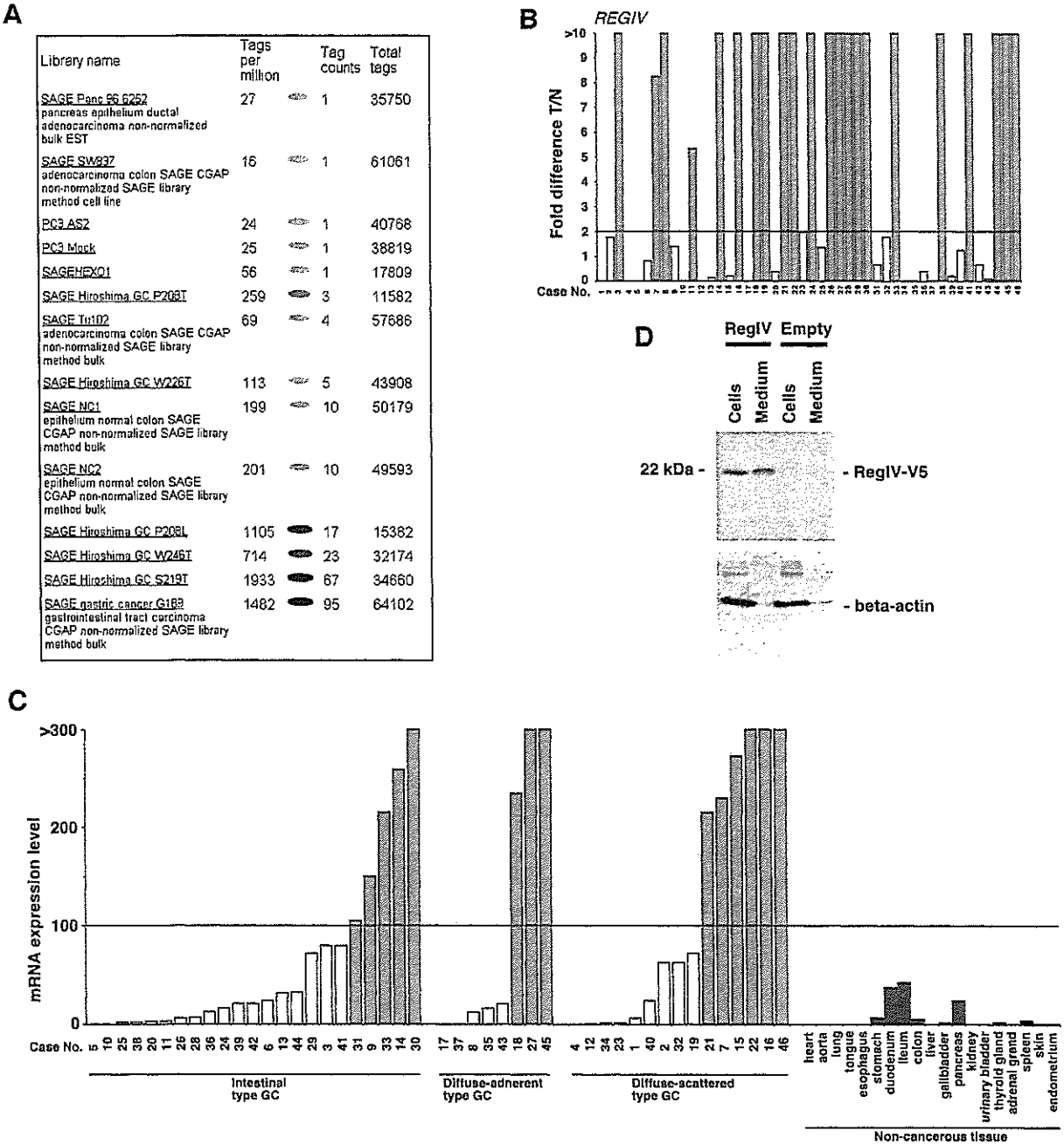

Fig. 4. *A*, Virtual Northern analysis shows *REGIV* expression to be narrowly restricted. Poorly differentiated gastric cancer (GC G189) showed strong expression of *REGIV*, whereas moderately differentiated GC (GC G234) did not. In our SAGE analysis, *REGIV* expression was detected as follows: 1933 in S219T; 714 in W246T; 1105 in P208L; 113 in W226T; and 259 in P208T. *B*, quantitative RT-PCR analysis of *REGIV* in primary GC and corresponding nonneoplastic mucosa. Fold change indicates the ratio of *REGIV* mRNA level in GC to that in corresponding nonneoplastic mucosa. Of the 46 GC samples, overexpression ($T/n > 2$) of *REGIV* was identified in 22 (47.8%). *C*, quantitative RT-PCR analysis of *REGIV* in primary GC samples and various noncancerous tissues. In GC, a high level of *REGIV* expression (>100 arbitrary units) was identified in samples 31, 9, 33, 14, 30, 18, 27, 45, 21, 7, 15, 22, 16, and 46. In various noncancerous tissues, a high level of *REGIV* expression was not identified. *REGIV* expression was found in noncancerous stomach, duodenum, ileum, colon, and pancreas. The units are arbitrary, and we calculated *REGIV* mRNA expression by standardization to 1.0 μg of total RNA from HSC-39 as 1.0. *D*, anti-V5 Western blot assay of V5 epitope- tagged RegIV protein. Cells and media from MKN-28 cells transfected with pcDNARegIV-V5 (RegIV) or pcDNA 3.1 (empty) constructs were lysed, resolved by SDS-PAGE, and immunoblotted with monoclonal mouse anti-V5 antibody. We confirmed by anti-β-actin Western blot that contamination of cells in culture medium was minimal.

Northern analysis. Furthermore, we showed that the RegIV protein is secreted, suggesting that RegIV may serve as a serum tumor marker. The number of samples we studied was small, and serum RegIV protein levels have not been examined. Additional investigation will clarify whether the RegIV protein can serve as a serum tumor marker. The role of *REGIV* gene overexpression in stomach carcinogenesis

GC samples, overexpression ($T/n > 2$) was detected at the following frequencies: 31 (67.4%) for *APOC1* and 24 (52.2%) for *YF13H12*. *B*, quantitative RT-PCR analysis of genes potentially involved in tumor progression according to SAGE analysis. Of the 46 GC samples, overexpression ($T/n > 2$) was detected at the following frequencies: 34 (73.9%) for *CDH17* and 19 (41.3%) for *FUS*. *C* and *D*, quantitative RT-PCR analysis of genes potentially involved in tumor metastasis according to SAGE analysis. *C*, mRNA expression levels of indicated genes in nonneoplastic mucosa, tumor, and lymph node metastasis. The units are arbitrary, and we calculated the target mRNA expression level by standardization to 1.0 μg of total RNA from HSC-39 as 1.0. T, tumor; n = nonneoplastic mucosa; L, lymph node metastasis; NL, normal lymph node from autopsy. *D*, mRNA expression levels of indicated genes in 46 GC samples. Of the 46 GC samples, overexpression ($T/n > 2$) was detected at the following frequencies: 23 (50%) for *APOE* and 19 (41.3%) for *S100A11*.

Table 5 *The 10 most up-regulated tags in lymph node metastasis of gastric carcinoma in comparison with primary gastric carcinoma*

Tag sequence	Tags per million		UniGene ID	Symbol	Description
	P208T	P208L			
ATCGGGCCCG	0[a] (0)[b]	1105 (17)	Hs.274411	*SCAND1*	SCAN domain containing 1
TATGAGGGTA	0 (0)	975 (15)	Hs.24950	*RGS5*	Regulator of G-protein signalling 5
CAGGCCCCAC	0 (0)	780 (12)	Hs.417004	*S100A11*	S100 calcium binding protein A11 (calgizzarin)
			Hs.145696	*RNPC2*	RNA-binding region (RNP1, RRM) containing 2
CGACCCCACG	0 (0)	780 (12)	Hs.169401	*APOE*	Apolipoprotein E
GCCCAGGTCA	86 (1)	1560 (24)	Hs.10499	*FLJ10815*	Hypothetical protein FLJ10815
TTAACCCCTC	86 (1)	1430 (22)	Hs.78224	*RNASE1*	Ribonuclease, RNase A family, 1 (pancreatic)
			Hs.393660	*H3F3B*	H3 histone, family 3B (H3.3B)
CAAGCAGGAC	0 (0)	650 (10)	Hs.424551	*P24B*	Integral type 1 protein
TAGAAAGGCA	0 (0)	650 (10)	Hs.457718		na LOC151103
CTCGCGCTGG	0 (0)	585 (9)	Hs.25640	*CLDN3*	Claudin 3
GCTGCTCCCT	0 (0)	585 (9)	Hs.343579	*MRPL14*	Mitochondrial ribosomal protein L14

[a] The absolute tag counts are normalized to 1,000,000 total tags per sample.
[b] Number in parentheses indicates the absolute tag counts.

remains unclear. A possible involvement of *REGIV* in drug (5-fluorouracil or methotrexate) resistance was reported recently (38). Thus, *REGIV* may inhibit apoptosis and may participate in tumor cell growth.

We found that *FUS* and *YF13H12* were overexpressed in GC. *FUS* was first identified as the 5′-part of a fusion gene with *CHOP* in myxoid liposarcomas with the translocation t(12;16)(q13;p11), and FUS protein was found to bind to RNA (39). No studies have analyzed *FUS* expression in human cancers, including GC. However, it has been shown that expression of the *FUS* domain restores liposarcoma development in *CHOP*-transgenic mice (40), suggesting that gain-of-function mutation of both *FUS* and *CHOP* is important. In the present study, *FUS* was a candidate marker for tumor progression, and we showed that a high level of *FUS* expression was associated with advanced N grade and stage. We also found *YF13H12* gene overexpression in GC. However, YF13H12 function remains unclear, and there are no reports on *YF13H12* gene expression. Additional studies will elucidate the biological role of *FUS* and *YF13H12* protein in GC.

Although we found several genes to be overexpressed in GC by SAGE, there were some exceptions of genes overexpressed by SAGE but not by quantitative RT-PCR. It is possible that inconsistent results between SAGE and quantitative RT-PCR represent more than one gene. For example, TTTAATTTGT, represented in both *GOLPH2* and *G3BP*, is commonly up-regulated in GC; however, the expression levels of both *GOLPH2* and *G3BP* were not frequently up-regulated by quantitative RT-PCR. Whether discrepancies between SAGE and quantitative RT-PCR are attributable to differences in methodology remains to be determined. Some GC samples that we analyzed showed overexpression of both *GOLPH2* and *G3BP* by quantitative RT-PCR. Recent evidence indicates that G3BP may serve as an important downstream effector of Ras signaling, and G3BP has been shown to be overexpressed in cancers of the colon, thyroid, breast, and head and neck (41). Thus, genes not frequently overexpressed may play an important role in restricted cases of GC.

Interestingly, among the 20 up-regulated tags in each GC sample, the 2 intestinal-type samples showed distinct tumor stages but showed many of the same tags. Cluster analysis showed that the two intestinal-type GC libraries were the most similar to each other. These results lead us to speculate that morphological phenotype reflects the gene expression profile. Our present results may be due to the selection of samples that represented similar histological features among many variations of intestinal type GC. Additional studies should investigate gene expression profile with respect to morphology. Comparison of expression patterns of W226T and W246T will provide a list of genes involved in tumor progression without the potential bias of histology. Our cluster analysis also showed that the gene expression pattern of SAGE gastric cancer-G234, which is a moderately differentiated tumor and is categorized as an intestinal type GC, was not similar to that of our 2 intestinal type GC samples but is similar to that of SAGE gastric cancer-G189, which is a poorly differentiated tumor and is categorized as a diffuse type GC. The gene expression patterns of GC in Japan may differ from those in the United States. Because we analyzed a limited number of GC samples, additional experiments are needed.

In conclusion, our present SAGE data provide a list of genes potentially involved in invasion, metastasis, and carcinogenesis of GC. We identified several genes by quantitative RT-PCR that have not previously been implicated in GC. Among these, a high level of *REGIV* expression was detected in GC, and expression of *REGIV* was narrowly restricted. Because the RegIV protein is secreted, it may serve as a biomarker for diagnosis of GC.

ACKNOWLEDGMENTS

We thank Dr. Kenneth W. Kinzler for SAGE software; Drs. Yukio Takeshima and Mayumi Kaneko for lung and breast cancer tissues; and Kenshi Tominaga and Yoshie Kaneko for excellent technical assistance and advice.

REFERENCES

1. Pisani P, Parkin DM, Bray F, Ferlay J. Estimates of the worldwide mortality from 25 cancers in 1990. Int J Cancer 1999;83:18–29.
2. Yasui W, Yokozaki H, Fujimoto J, Naka K, Kuniyasu H, Tahara E. Genetic and epigenetic alterations in multistep carcinogenesis of the stomach. J Gastroenterol 2000;35:111–5.
3. Oue N, Motoshita J, Yokozaki H, et al. Distinct promoter hypermethylation of p16INK4a, CDH1, and RAR-β in intestinal, diffuse-adherent, and diffuse-scattered type gastric carcinomas. J Pathol 2002;198:55–9.
4. Oue N, Oshimo Y, Nakayama H, et al. DNA methylation of multiple genes in gastric carcinoma: association with histological type and CpG island methylator phenotype. Cancer Sci 2003;94:901–5.
5. El-Rifai W, Frierson HF Jr, Harper JC, Powell SM, Knuutila S. Expression profiling of gastric adenocarcinoma using cDNA array. Int J Cancer 2001;92:832–8.
6. Hasegawa S, Furukawa Y, Li M, et al. Genome-wide analysis of gene expression in intestinal-type gastric cancers using a complementary DNA microarray representing 23,040 genes. Cancer Res 2002;62:7012–7.
7. Hippo Y, Taniguchi H, Tsutsumi S, et al. Global gene expression analysis of gastric cancer by oligonucleotide microarrays. Cancer Res 2002;62:233–40.
8. Lee S, Baek M, Yang H, et al. Identification of genes differentially expressed between gastric cancers and normal gastric mucosa with cDNA microarrays. Cancer Lett 2002;184:197–206.
9. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science (Wash. DC) 1995;270:484–7.
10. El-Rifai W, Moskaluk CA, Abdrabbo MK, et al. Gastric cancers overexpress S100A calcium-binding proteins Cancer Res 2002;62:6823–6.
11. Oien KA, Vass JK, Downie I, Fullarton G, Keith WN. Profiling, comparison and validation of gene expression in gastric carcinoma and normal stomach. Oncogene 2003;22:4287–300.
12. Lee JY, Eom EM, Kim DS, Ha-Lee YM, Lee DH. Analysis of gene expression profiles of gastric normal and cancer tissues by SAGE. Genomics 2003;82:78–85.
13. Lauren P. The two histological main types of gastric carcinoma. Diffuse and so-called intestinal type carcinoma: an attempt at histological classification. Acta Pathol Microbiol Scand 1965;64:31–49.

14. Shimoyama Y, Hirohashi S. Expression of E- and P-cadherin in gastric carcinomas. Cancer Res 1991;51:2185–92.
15. Sobin LH, Wittekind CH, editors. TNM classification of malignant tumors, 5th ed. New York: Wiley-Liss, Inc.; 1997, p. 59–62.
16. Hohenberger P, Gretschel S. Gastric cancer Lancet 2003;362:305–15.
17. Yokota T, Teshima S, Saito T, Kikuchi S, Kunii Y, Yamauchi H. Borrmann's type IV gastric cancer: clinicopathologic analysis. Can J Surg 1999;42:371–6.
18. Lal A, Lash AE, Altschul SF, et al. A public database for gene expression in human cancers. Cancer Res 1999;59:5403–7.
19. Kondo T, Oue N, Yoshida K, et al. Expression of POT1 is associated with tumor stage and telomere length in gastric carcinoma. Cancer Res 2004;64(2):523–9.
20. Yanagihara K, Seyama T, Tsumuraya M, Kamada N, Yokoro K. Establishment and characterization of human signet ring cell gastric carcinoma cell lines with amplification of the c-myc oncogene. Cancer Res 1991;51:381–6.
21. Yasui W, Ayhan A, Kitadai Y, et al. Increased expression of p34cdc2 and its kinase activity in human gastric and colonic carcinomas. Int J Cancer 1993;53:36–41.
22. Hartupee JC, Zhang H, Bonaldo MF, Soares MB, Dieckgraefe BK. Isolation and characterization of a cDNA encoding a novel member of the human regenerating protein family: reg IV. Biochim Biophys Acta 2001;1518:287–93.
23. Ryu B, Jones J, Hollingsworth MA, Hruban RH, Kern SE. Invasion-specific genes in malignancy: serial analysis of gene expression comparisons of primary and passaged cancers. Cancer Res 2001;61:1833–8.
24. Iacobuzio-Donahue CA, Argani P, Hempen PM, Jones J, Kern SE. The desmoplastic response to infiltrating breast carcinoma: gene expression at the site of primary invasion and implications for comparisons between tumor types. Cancer Res 2002; 62:5351–7.
25. St Croix B, Rago C, Velculescu V, et al. Genes expressed in human tumor endothelium. Science (Wash. DC) 2000;289:1197–202.
26. Berndorff D, Gessner R, Kreft B, et al. Liver-intestine cadherin: molecular cloning and characterization of a novel Ca(2+)-dependent cell adhesion molecule expressed in liver and intestine. J Cell Biol 1994;125:1353–69.
27. Grotzinger C, Kneifel J, Patschan D, et al. LI-cadherin: a marker of gastric metaplasia and neoplasia. Gut 2001;49:73–81.
28. Takamura M, Sakamoto M, Ino Y, et al. Expression of liver-intestine cadherin and its possible interaction with galectin-3 in ductal adenocarcinoma of the pancreas. Cancer Sci 2003;94:425–30.
29. Niemi M, Hakkinen T, Karttunen TJ, et al. Apolipoprotein E and colon cancer. Expression in normal and malignant human intestine and effect on cultured human colonic adenocarcinoma cells. Eur J Intern Med 2002;13:37–43.
30. Ohta M, Kitadai Y, Tanaka S, et al. Monocyte chemoattractant protein-1 expression correlates with macrophage infiltration and tumor vascularity in human gastric carcinomas. Int J Oncol 2003;22:773–8.
31. Venanzoni MC, Giunta S, Muraro GB, et al. Apolipoprotein E expression in localized prostate cancers. Int J Oncol 2003;22:779–86.
32. Yonemura Y, Endou Y, Kimura K, et al. Inverse expression of S100A4 and E-cadherin is associated with metastatic potential in gastric cancer. Clin Cancer Res 2000;6:4234–42.
33. Paxton RJ, Mooser G, Pande H, Lee TD, Shively JE. Sequence analysis of carcinoembryonic antigen: identification of glycosylation sites and homology with the immunoglobulin supergene family. Proc Natl Acad Sci USA 1987;84:920–4.
34. Oikawa S, Inuzuka C, Kuroki M, Matsuoka Y, Kosaki G, Nakazato H. Cell adhesion activity of non-specific cross-reacting antigen (NCA) and carcinoembryonic antigen (CEA) expressed on CHO cell surface: homophilic and heterophilic adhesion. Biochem Biophys Res Commun 1989;164:39–45.
35. Jantscheff P, Terracciano L, Lowy A, et al. Expression of CEACAM6 in resectable colorectal cancer: a factor of independent prognostic significance. J Clin Oncol 2003;21:3638–46.
36. Mashimo H, Wu DC, Podolsky DK, Fishman MC. Impaired defense of intestinal mucosa in mice lacking intestinal trefoil factor. Science (Wash. DC) 1996;274: 262–5.
37. Leung WK, Yu J, Chan FK, et al. Expression of trefoil peptides (TFF1, TFF2, and TFF3) in gastric carcinomas, intestinal metaplasia, and non-neoplastic gastric tissues. J Pathol 2002;197:582–8.
38. Violette S, Festor E, Pandrea-Vasile I, et al. Reg IV, a new member of the regenerating gene family, is overexpressed in colorectal carcinomas. Int J Cancer 2003;103: 185–93.
39. Crozat A, Aman P, Mandahl N, Ron D. Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. Nature (Lond.) 1993;363:640–4.
40. Perez-Mancera PA, Perez-Losada J, Sanchez-Martin M, et al. Expression of the FUS domain restores liposarcoma development in CHOP transgenic mice Oncogene 2002; 21:1679–84.
41. Guitard E, Parker F, Millon R, Abecassis J, Tocque B. G3BP is overexpressed in human tumors and promotes S phase entry. Cancer Lett 2001;162:213–21.

Review Article

Search for new biomarkers of gastric cancer through serial analysis of gene expression and its clinical implications

Wataru Yasui, Naohide Oue, Reiko Ito, Kazuya Kuraoka and Hirofumi Nakayama

Department of Molecular Pathology, Hiroshima University Graduate School of Biomedical Sciences, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551

(Received February 4, 2004/Revised March 12, 2004/Accepted April 7, 2004)

Gastric cancer is one of the most common human cancers and is the second most frequent cause of cancer-related death in the world. Serial analysis of gene expression (SAGE) is a powerful technique to allow genome-wide analysis of gene expression in a quantitative manner without prior knowledge of the gene sequences. SAGE on 5 samples of gastric cancer with different histology and clinical stages have created large SAGE libraries of gastric cancer that enable us to identify new cancer biomarkers. Commonly up-regulated genes in gastric cancer in comparison with normal gastric epithelia included *CEACAM6*, *APOC1* and *YF13H12*. By comparing gene expression profiles of gastric cancers at early and advanced stages, several genes differentially expressed by tumor stage were also identified, including *FUS*, *CDH17*, *COL1A1* and *COL1A2*, which should be novel genetic markers for high-grade malignancy. Regenerating gene type IV (*REGIV*) is one of the most up-regulated genes in a SAGE library of a scirrhous-type gastric cancer. *In vitro* studies using RegIV-transfected cells revealed that RegIV is secreted by cancer cells and inhibits apoptosis, suggesting that RegIV may serve as a novel biomarker and therapeutic target for gastric cancer. Production of RNA aptamers could be a useful approach to establish a detection system in blood. A custom-made array, named Ex-STOMACHIP, consisting of 395 genes, including highly differentially expressed genes identified by our SAGE and other known genes related to carcinogenesis and chemosensitivity, is useful to study the molecular pathogenesis of gastric cancer and to obtain information about biological behavior and sensitivity to therapy in the clinical setting. Combined analyses of gene expression profile, genetic polymorphism and genetic instability will aid not only cancer detection, but also characterization of individual cancers and patients, leading to personalized medicine and cancer prevention. (Cancer Sci 2004; 95: 385–392)

According to the World Health Organization, gastric cancer is the fourth most common malignancy in the world, with some 870,000 new cases every year, and mortality from gastric cancer is second only to lung cancer.[1] The incidence of gastric cancer is declining worldwide. This trend is mainly due to decreased consumption of salt-preserved food, avoidance of high-salt diet and availability of fresh fruit and vegetables throughout the year. Another reason for the high incidence of gastric cancer in Japan is a high rate of *Helicobacter pylori* infection among Japanese. *Helicobacter* causes chronic active or atrophic gastritis and intestinal metaplasia, which are believed to be precancerous or predisposing conditions for gastric cancer. Advances in diagnosis and treatment have resulted in excellent long-term survival for patients with early cancer, but the prognosis of advanced cancer remains poor.

Cancer is a chronic proliferative disease with multiple genetic and epigenetic alterations; that is, it is a disease with altered gene expression. Integrated research in molecular pathology over the past 15 years has uncovered many of the molecular mechanisms of the development and progression of gastric cancer.[2–6] Genetic polymorphism is an important endogenous cause and a fundamental factor influencing cancer risk. Genetic instability, DNA hypermethylation and histone hypoacetylation are early events. Multiple alterations accumulate, including inactivation of tumor suppressor genes, activation of oncogenes and abnormalities of cell cycle regulators and growth factors. Some of these changes occur commonly in both differentiated and undifferentiated types and some differ depending on the histological type. A better knowledge of changes in gene expression during stomach carcinogenesis may lead to new paradigms and possible improvements in cancer diagnosis, treatment and prevention. Although several large-scale gene expression studies using microarray techniques have been performed on gastric cancer,[7–9] they have utilized different platforms that varied in the number and identity of the genes printed on them. Besides the microarray technique, serial analysis of gene expression (SAGE) is a powerful technique for global analysis of gene expression in a quantitative manner without prior knowledge of the gene sequences.[10]

This review presents an outline of our approach to search for new genes of gastric cancer through SAGE and discusses its implications for diagnosis, treatment and prevention.

Advantage of SAGE in global analysis of gene expression

SAGE is based on the following two principles.[10] First, a short nucleotide sequence tag (about 10 base pairs) contains sufficient information to uniquely identify a transcript, provided it is isolated from a defined position within the transcript. Second, concentration of short sequence tags allows the efficient analysis of transcripts in a serial manner by the sequencing of multiple tags within a single clone. SAGE analyzes tags of about 10 bp derived from a defined position, near the polyA tail of the cDNAs, downstream of the CATG sequence. Because the SAGE tag numbers directly reflect the abundance of the mRNAs, SAGE data are highly accurate and quantitative. Completion of the human genome sequence has facilitated the mapping of specific genes to individual tags. Large numbers of normal and tumor tissues and cells have been analyzed by SAGE, creating large databases. Now, database including about 250 SAGE libraries are online and available to the public. Since the database contains accurate distribution and frequency data of

E-mail: wyasui@hiroshima-u.ac.jp

the genes and tags, we can study the expression of genes of interest in other tissues in the database by Virtual northern analysis.[11] Molecular characterization has been performed by SAGE in cancers of the lung, breast, colon, ovary and prostate, tumor endothelium, and other organs with the aim of developing diagnostic and therapeutic strategies.[12–16] Although SAGE is thus useful for global analysis of gene expression and to identify genes up-regulated or down-regulated in tissues and cells of interest, it is not suitable to study gene expression in large numbers of cases because at least 10,000 tags should be sequenced to obtain a SAGE library. To verify whether gene expression data obtained by SAGE reflect general phenomena, real-time RT-PCR or other methods must be combined.

Strategy to search for novel genes associated with gastric cancer by SAGE and its clinical implication

Our strategy to search for novel biomarkers using SAGE and to apply the results to clinical diagnostics, treatment and prevention is shown in Fig. 1. First, we perform SAGE on typical gastric cancer tissues, compare gene expression profiles among them or with those in normal gastric tissue and identify specifically up-regulated or down-regulated genes. The expression of these genes is confirmed in large numbers of cases by real-time RT-PCR and immunohistochemistry if antibodies are available. With the specific genes identified by SAGE, known genes participating in the development and progression of gastric cancer and known genetic markers for chemosensitivity, we prepare a custom-made cDNA microarray. If a specific gene encodes a secretory protein, this may be detected in the blood and should be a novel biomarker of gastric cancer. For such a molecule, we produce an RNA aptamer or antibody and establish a measuring system such as ELISA for blood. Genetic polymorphism is an important determinant among endogenous causes of cancer. The majority of genetic variation between individual humans is believed to be due to single nucleotide polymorphisms (SNPs), and 1% of all SNPs results in functional variation in proteins and alters cancer predisposition.[17, 18] Polymorphisms of genes whose expression is highly altered in cancer may be candidates for novel risk factors, and this information will be useful for cancer prevention. By functional analysis, we can understand the molecular mechanisms of stomach carcinogenesis in more detail and determine whether the genes can be novel therapeutic targets.

Generation of gene expression profiles of gastric cancer by SAGE

So far, three SAGE studies of gastric cancer have been reported that identified several up-regulated and down-regulated genes.[19–21] Up-regulated genes include *S100A* calcium-binding protein family and *TFF3*. However, the reported studies examined only one[19] or two samples[20, 21] of gastric cancer. We have performed SAGE analysis on 5 samples of gastric cancer of different stages and histologies from 4 patients.[22] They included early and advanced cancers of well-differentiated type (tubular adenocarcinoma of well-differentiated type; tub1), primary and metastatic tumors of poorly differentiated adenocarcinoma of solid type (por1) and one poorly differentiated adenocarcinoma of non-solid type (por2; scirrhous) gastric cancer. Histological classification was made according to the Japanese Classification of Gastric Cancer.[23] SAGE was carried out according to SAGE protocol version 1.0e, June 23, 2000. Tags were extracted from the raw sequence data with SAGE2000 analysis software version 4.12, kindly provided by Dr. Kenneth W. Kinzler (Johns Hopkins University School of Medicine). Our SAGE generated a total of 137,706 tags including 38,903 unique tags. Our SAGE libraries are the largest gastric cancer libraries in the world and the sequence data are publicly available at SAGEmap (GEO accession number GSE 545, SAGE Hiroshima gastric cancer tissue) (http://www.ncbi.nlm.nih.gov/SAGE/). Besides our libraries, five other SAGE libraries of gastric tissue are available from SAGEmap. Those include two gastric cancers (GSM757, SAGE_gastric_cancer-G234 and GSM2385, SAGE_gastric_cancer-G189) and normal gastric epithelia (GSM874, SAGE_normal_gastric_body_epithelia).[19] As already mentioned, SAGE is useful for comprehensive gene expression analysis and an SAGE database contains accurate data on the distribution and frequency of tags, so we can directly

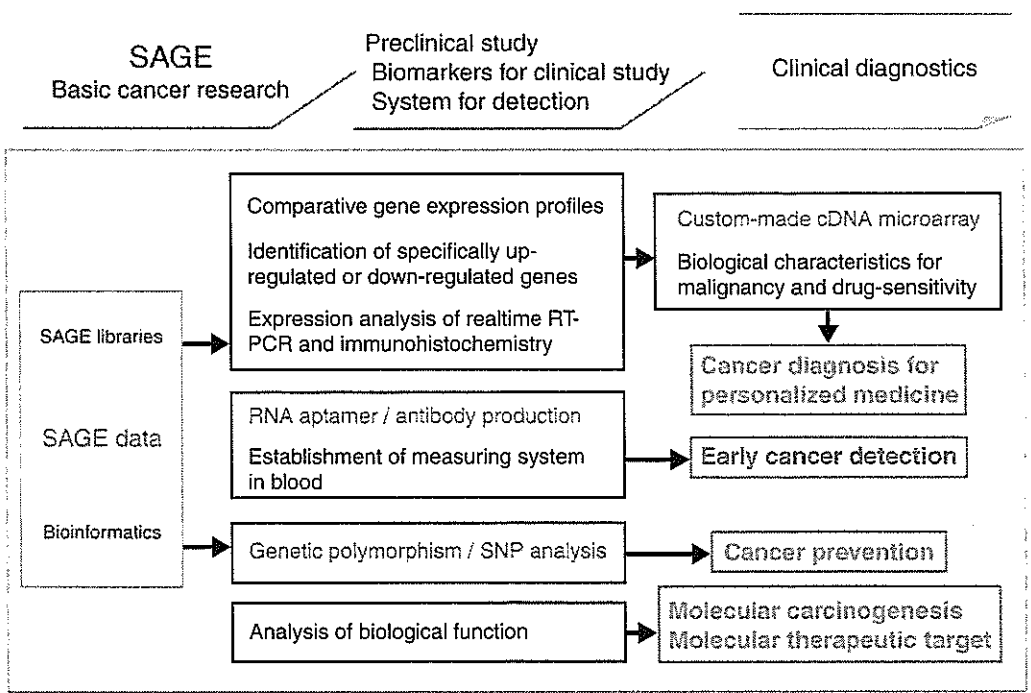

Fig. 1. Strategy to search for novel genes associated with gastric cancer through SAGE, and its clinical implication.

Table 1. Up-regulated and down-regulated tags and genes in gastric cancer obtained by SAGE

Commonly up-regulated and down-regulated tags and genes in gastric cancer in comparison with normal gastric epithelia	
Up-regulated	APOC1, S100A4, NDUF2, TEBP, COL1A2, SUFU, SYAP1, KIAA0930, KIAA1694, TFF3, CEACAM6, FLJ20249, FLJ22167, EIF4A1, COLPH2, G3BP, YF13H12, KRT7, SH3BP2, COL1A1, LOC284371
Down-regulated	CAGCGCTTCT (no match), CACCTCCCCA (no match), AGCCTCCCCA (no match), ACCCTCCCCA (no match), LIPF, AACCTCCCCC (no match), CHIA, TAGTGCTTCT (no match), TACAAGGTCC (no match), GTGGTCAGCT (no match), ATP4B, FLJ20410, MBD3, CAGTGCTTTT (no match), Hs.199360, Hs.353061
The 20[1] most up-regulated and down-regulated tags and genes in advanced gastric cancer in comparison with early gastric cancer	
Up-regulated	TCCCCGTAAA (no match), TCCCGTACAT (no match), CDH17, FUS, PRO1073, FLJ36926, FLJ30146, PAI-RBP1, COL1A2, TCCTATTAAG (no match), COL1A1, GRAP2, HNRPL, NUTF2, ERP70, PES1, CYP2J2, DAG1, IQGAP1, IL16, FXYD3, COQ4, LOC91966, CTBP1, TTCGGTTGGT (no match), alpha4GnT, Hs.290723, AKT3, CCT3, HMG20A
Down-regulated	Hs.216636, LOC116228, SH3MD2, NAB1, TTCCCCCAAA (no match), DDX5, VMP1, LOC51123, LZK1, CGCAGATCAG (no match), IFRD2, Hs.284464, RPS4Y, RPS4Y2, UAP1, Hs.180804, CATTAAATTA (no match), IKBKAP, ARPC3, NAGA, UBE3A, TRAG3, PNN, CTAATTCTTT (no match), TCCATCGTCC (no match)
The 20[1] most up-regulated and down-regulated tags and genes in lymph-node metastasis in comparison with primary tumor of gastric cancer	
Up-regulated	SCAND1, RGS5, S100A11, RNPC2, APOE, FLJ10815, RNASE1, H3F3B, P24B, LOC151103, CLDN3, MRPL14, PRex1, TCCCCTATTA (no match), Hs.105379, ATP5G1, NPD007, MGC3180, WDR11, ARPC1B, ABTB2, DNAJB1, HMGN2, KIAA1393, RAP1B, FLJ12150, STUB1
Down-regulated	ERdj5, RPL27A, DHRS3, E2IG5, USP7, CTSL, KRTHB1, KRTHB3, TGCACTACCC (no match), ALG12, S100A9, CTAGCTTTTA (no match), ELOVL5, LOC375463, GGGGGAGTTT (no match), ACTGCCCTCA (no match), SPC18, CTNND1, CYP20A1, FLJ11151, RPS17, ZYX, RPS16, GCTTTCTCAC (no match), BCL2L2

The gene symbol is shown, while the UniGene ID is given if the symbol is not present. No match, tag sequence is not matched to any known gene.
1) Because some genes share the same SAGE tag, gene numbers are more than 20.

compare our tags with those of other SAGE libraries using SAGEmap.

Identification of gastric cancer-specific genes by SAGE

To identify gastric cancer-specific genes, SAGE tags from six primary gastric cancers were compared with those from normal gastric epithelia, and the 20 most up-regulated and 20 most down-regulated tags were selected.[22] This produced a dataset of 128 tags because of overlapping. By complete linkage clustering among eight gastric libraries in the SAGEmap database, clusters with commonly up-regulated genes and down-regulated genes in all the gastric cancers were identified (Table 1). The up-regulated gene cluster included *APOC1, NDUF2, TEBP, COL1A1* and so on, in addition to *TFF3* and *S100A4*, which are known to be up-regulated in gastric cancer. Quantitative real-time RT-PCR on 46 cases of gastric cancer revealed that *APOC1, CEACAM6* and *YF13H12* were frequently overexpressed (more than 50% of cases showed a tumor/normal ratio >2) in gastric cancer tissues, and these should be novel genetic markers for gastric cancer. On the other hand, the down-regulated gene cluster included *LIPF* (gastric lipase), *CHIA, ATP4B, MBD3* and many unknown genes (tags whose corresponding genes have not been identified). Because they were commonly down-regulated in gastric cancer, the unknown genes are novel candidates for gastric-specific tumor suppressors. The reverse SAGE technique will help to identify these genes.[24]

Another approach to identify gastric cancer-specific genes is the selection of candidate genes by comparing SAGE libraries of gastric cancer with those of various normal tissues in the SAGEmap database. We picked up about 60 genes which were detected in our gastric cancer libraries, but not in the libraries from 15 kinds of normal tissues, especially important or crucial organs, including brain, lung, heart, liver, kidney, etc. We then examined the expression of these genes in normal human tissues by RT-PCR, and representative results are shown in Fig. 2. Many genes were expressed at various levels in normal tissues, although the comparison of SAGE data suggested them to be gastric cancer-specific. Some genes (genes H, I, J, K) were not expressed significantly in normal tissues, but were expressed in gastric cancer cells. Therefore, these may be gastric cancer-specific, or at least cancer-specific, and could be candidates for novel genetic markers.

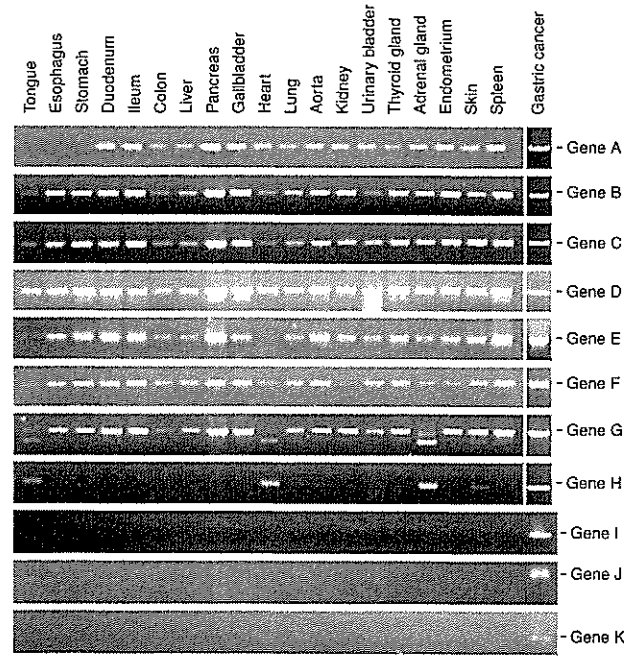

Fig. 2. Expression of genes detected in SAGE libraries of gastric cancer, but not in the libraries of 15 normal tissues. Semi-quantitative RT-PCR on various normal human tissues revealed that genes H, I, J and K are not expressed significantly in normal tissues, whereas they are detected in gastric cancer cells.

Identification of genes involved in invasion and metastasis by SAGE

The genes participating in invasion and metastasis can be identified by comparing SAGE libraries between cancers with and without invasion and metastasis. For this purpose, we first compared SAGE libraries between an early cancer (T1, N0,

M0, stage IA)[25] and an advanced cancer (T2, N2, M0, stage IIIA).[22] Both were well-differentiated-type gastric cancer (tub1) with intestinal phenotype expressing MUC2. The 20 most up-regulated and down-regulated tags and corresponding genes in the advanced cancer are shown in Table 1. Up-regulated genes include some unknown genes, *CDH17, FUS* and so on. Quantitative real-time RT-PCR demonstrated the frequent overexpression of *FUS, CDH17, COL1A1* and *COL1A2* and its positive correlation with degree of invasion, metastasis and advanced stage. *FUS* is a tumor-associated fusion gene, especially in myxoid liposarcoma, and it may have a role in regulating transcription and maintaining chromosomal stability.[26] *CDH17* (cadherin 17, liver-intestine cadherin) is known to be up-regulated in intestinal metaplasia and well-differentiated-type gastric cancer of intestinal phenotype.[27] Our immunohistochemical study confirmed that overexpression of cadherin 17 is associated with advanced stage and intestinal histology (Fig. 3, A and B), and cadherin 17-positive patients showed a poorer prognosis than the negative patients.

Using the same strategy, we tried to identify genes involved in metastasis. We compared SAGE libraries between the primary gastric cancer (por1) and its lymph-node metastatis from the same patient.[22] The 20 most up-regulated tags and corresponding genes in the metastatic tumor included *SCAND1, RGS5, S100A11, RNPC2, APOE* and so on (Table 1). Among them, *APOE* (apolipoprotein E) expression was confirmed to be associated with T grade, N grade and advanced stage. Immunohistochemically, apolipoprotein E was expressed mainly in stromal cells which are also positive for CD68, suggesting tumor-associated macrophages. Apolipoprotein E-positive cells were more prominent in the metastatic tumor than in the primary tumor (Fig. 3, C and D). Therefore, APOE may be a novel marker for metastasis, although the mechanism involved remains to be elucidated.

Candidate novel biomarkers of gastric cancer

Ideal biomarkers for cancer should have the following four characteristics.[14] First, they should be expressed at high levels in tumors and at greatly reduced levels in normal tissues. Second, the elevated expression should occur early and remain elevated during the neoplastic process. Third, the markers should be elevated in a majority of clinical samples. Fourth, the markers should be expressed on the cell surface or secreted to facilitate detection. Genes and molecules that meet all these criteria are especially useful for the development of diagnostic tools for early cancer detection. Moreover, if the function of the gene product is involved in the neoplastic process, such a gene is not just a biomarker, but can be a therapeutic target. In the course of the SAGE study, we found that *REGIV* (regenerating gene type IV) meets these criteria. A comparison of the expressed tags of scirrhous-type gastric cancer (por2) with those of normal gastric epithelia showed that *REGIV* was the second most up-regulated gene after *TFF3*.[22] Virtual northern analysis revealed that *REGIV* expression is narrowly restricted in comparison with other tissues and cancers in the SAGEmap database. Furthermore, quantitative RT-PCR showed that about a half of gastric cancers overexpressed *REGIV* mRNA, regardless of tumor stage and histological differentiation, whereas low levels of *REGIV* expression were limited to non-neoplastic gastrointestinal and pancreatic tissues. RegIV belongs to the calcium-dependent lectin superfamily, and is known to increase in inflammatory bowel diseases and a portion of colon

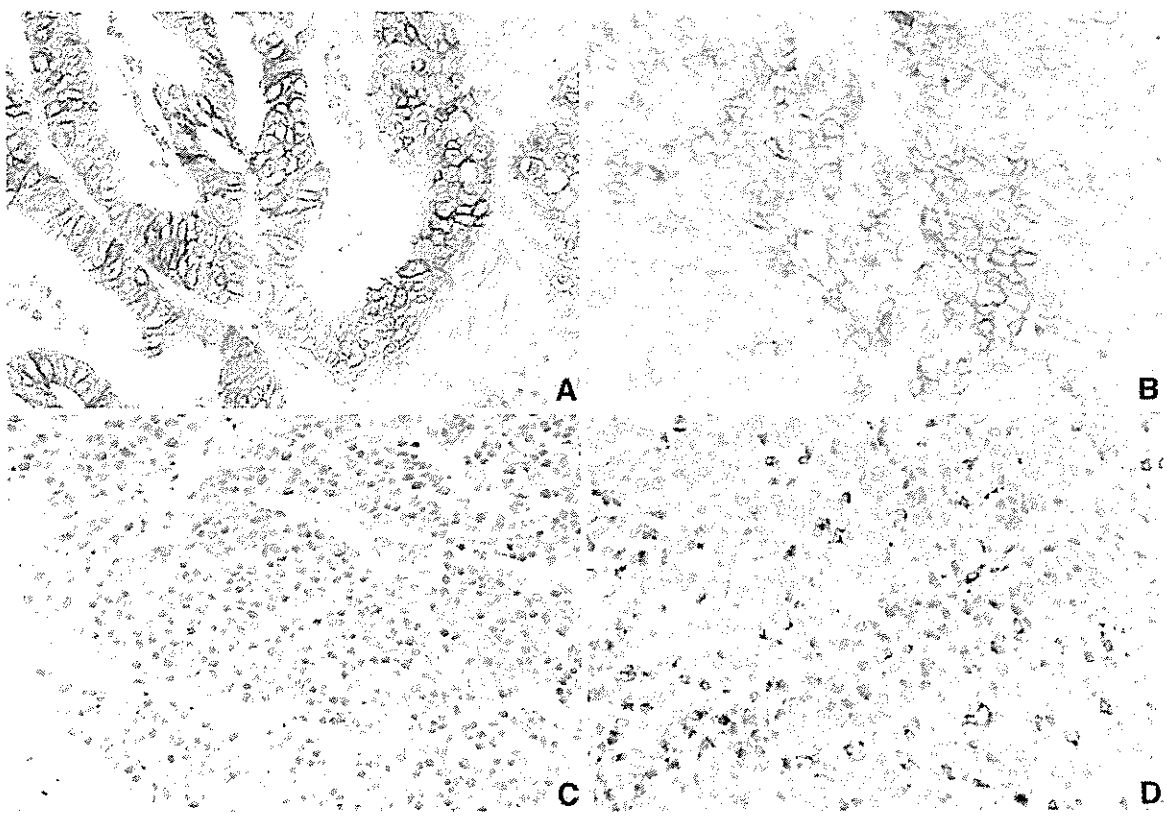

Fig. 3. Expression of cadherin 17 and apolipoprotein E in gastric carcinomas (immunohistochemical analysis). Cadherin 17 is expressed at the cell surface and cell-cell border of well-differentiated-type gastric adenocarcinoma with intestinal phenotype (A) and poorly differentiated solid-type gastric adenocarcinoma (B). Apolipoprotein E-positive cells are more prominent in metastatic tumor (D) than in primary tumor (C) of poorly differentiated solid-type adenocarcinoma.

cancers.[28,29] Transfection of a RegIV expression vector (pcDNA-RegIV-V5) into gastric cancer cell lines enhanced invasion and inhibited apoptosis. RegIV-V5 was detected in culture media of the transfected cells, indicating that RegIV is secreted by cancer cells. These findings strongly suggest that RegIV may be involved in the neoplastic process, and therefore, RegIV is not just a biomarker, but a novel therapeutic target for gastric cancer.

To obtain an anti-RegIV agent for diagnostic and therapeutic purpose, one way is the production of antibody, but another is the production of oligonucleotides that specifically bind to target molecules. Small non-coding RNA has been focused on recently as having a range of potential functions including regulation of gene expression. Oligonucleotide sequences which recognize target molecules with high affinity and specificity, called aptamers, can be isolated by systemic evolution of ligands by exponential enrichment (SELEX) process.[30,31] Recombinant RegIV protein from full-length RegIV cDNA and a random sequence oligonucleotide library are mixed and incubated. Bound oligonucleotides are selected and amplified by PCR. This cycle is repeated at least 10 times to obtain a specific aptamer. Aptamers rival antibodies because they are

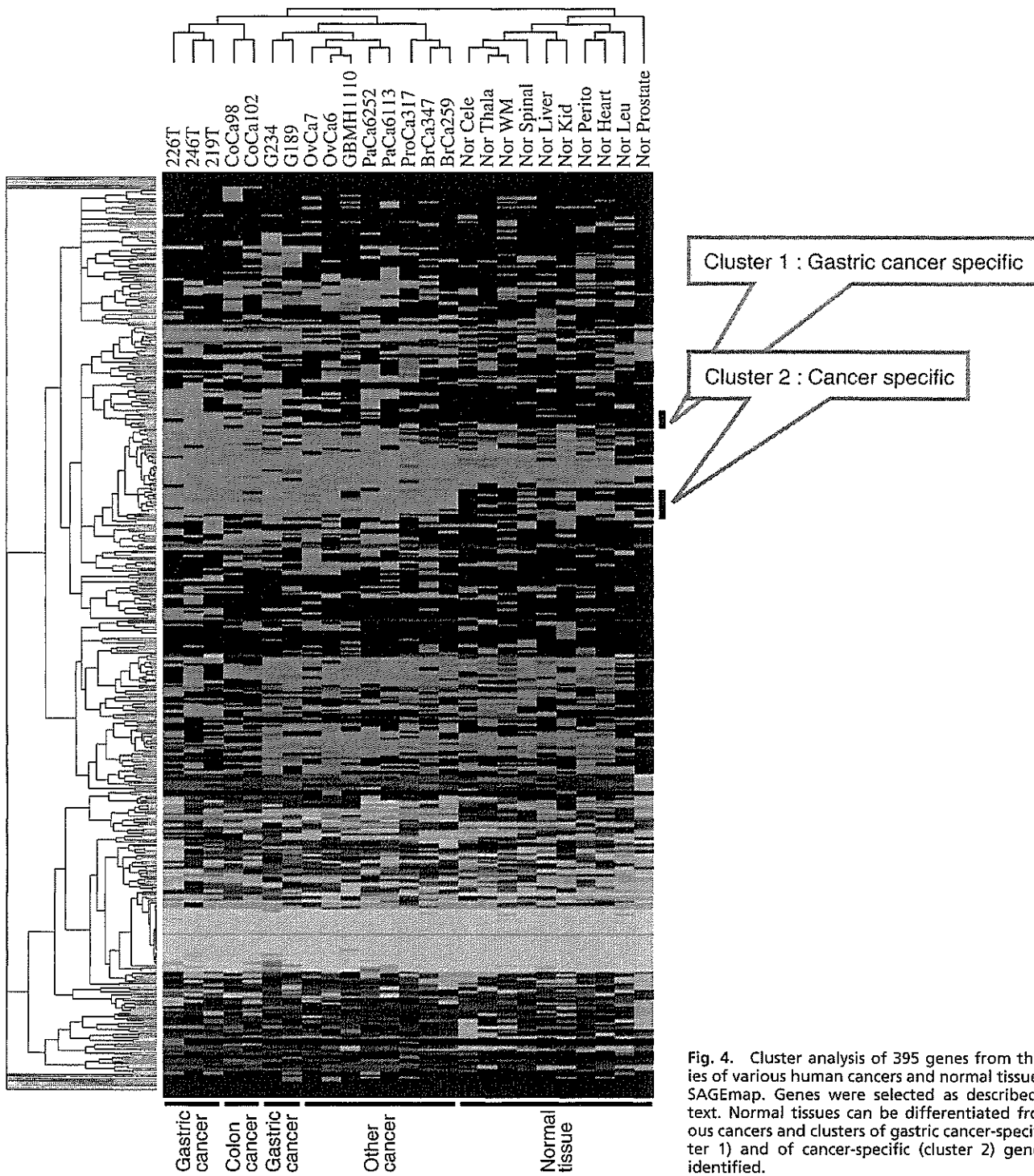

Fig. 4. Cluster analysis of 395 genes from the libraries of various human cancers and normal tissues in the SAGEmap. Genes were selected as described in the text. Normal tissues can be differentiated from various cancers and clusters of gastric cancer-specific (cluster 1) and of cancer-specific (cluster 2) genes were identified.

highly specific, able to distinguish family members, exhibit reversible equilibrium binding, and can be synthesized automatically in an *in vitro* system.

Development of specialized cDNA microarray for study and diagnosis of gastric cancer

Microarray studies have uncovered many genes related to biological behavior, such as metastasis and prognosis and sensitivity to chemotherapy.[8, 9, 32–34] On the other hand, SAGE has identified many differentially expressed genes and candidate novel biomarkers. However, it is hard to study large numbers of clinical samples and to apply the technique in a clinical setting. Because differential expression must be confirmed by other methods, many genes still remained unconfirmed. Thus, we decided to prepare a custom-made microarray for the study of stomach carcinogenesis and possible clinical application. A similar approach has been introduced in ovarian cancers.[35] The microarray, named Ex-STOMACHIP, consists of 395 genes selected based on the following three criteria. 1) The 164 genes which were selected as the 20 most up-regulated and down-regulated tags in the six SAGE libraries of gastric cancer studied by us, as already mentioned. 2) Known genes participating in stomach carcinogenesis, including genes of growth factors/receptors, cell cycle regulators, metalloproteinases, adhesion molecules, and so on. 3) Genes related to metastasis and chemosensitivity identified by other cDNA microarray studies. Using these sets of genes, hierarchical clustering was performed among various SAGE libraries of cancerous and non-cancerous tissues in the SAGEmap database. As shown in Fig. 4, gastric cancer tissues or other cancers including cancers of the ovary, pancreas and breast were differentiated from various normal tissues, and several gene clusters were identified. For instance, cluster 1 is gastric cancer-specific, and cluster 2 is specific for cancer in general. When results among gastric cancers with various differences in morphology and biological behavior were compared, clear gene clusters could be identified. Thus, Ex-STOMACHIP should be a useful tool not only to study mechanisms of stomach carcinogenesis, but also to obtain information about biological behavior and sensitivity to therapy in the clinical setting.

New strategy of gene diagnosis of gastric cancer

In 1993, we established a gene diagnosis system for gastrointestinal pathology specimens and performed this as a routine service until 2000, using so-called classical molecular and genetic markers, including *p53, APC, p27, EGFR*, microsatellite assay and so on.[3, 36] We analyzed more than 10,000 cases, and obtained much useful information concerning differential diagnosis, grade of malignancy and susceptibility of multiple primary cancers. Now, the molecular diagnosis of pathology specimens must move into the era of genomic medicine. Here, we have outlined our search for new genetic markers of gastric

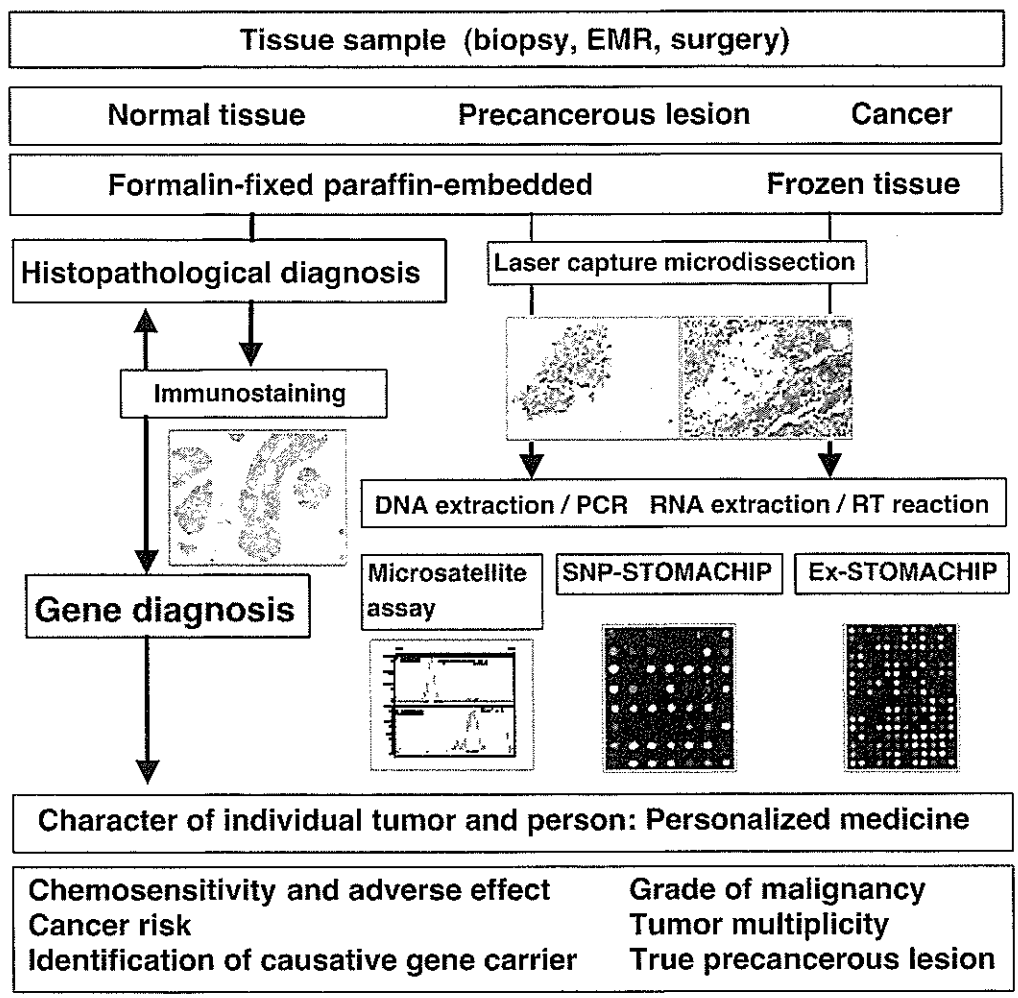

Fig. 5. New strategy of gene diagnosis of gastric cancer using pathology samples.

cancer and their clinical implications. Many new candidate biomarkers of gastric cancer were identified by SAGE, and can be introduced into clinical diagnosis through two approaches. One is the use of custom-made arrays for analysis of tissue samples, and the other is detection systems for blood samples using RNA aptamer or antibody. Information about SNPs in highly differentially expressed genes will be directly connected with cancer prevention. Fig. 5 illustrates the new strategy of gene diagnosis of gastric cancer on pathology samples obtained by biopsy, endoscopic mucosal resection or surgery. Tissues are collected from samples freshly frozen or fixed with formalin and embedded in paraffin, and DNA and RNA are recovered using laser capture microdissection, if necessary. The gene expression profile is examined by the use of Ex-STOMACHIP to obtain information about grade of malignancy and chemosensitivity, as well as possible side effects. Microsatellite analysis predicts tumor multiplicity. If these analyses are extended to mucosa without overt morphological aberration, super-early diagnosis of gastric cancer should be possible. Analysis of genetic polymorphism will give information about cancer risk and sensitivity to chemotherapy. Combinations of these approaches can not only achieve cancer detection, but also clarify the character of the individual tumor and patient, thereby leading to personalized medicine and cancer prevention.

Conclusion

SAGE is a powerful technique to identify novel genes associated with gastric cancer, and to search for new biomarkers of gastric cancer. Our SAGE libraries are the largest gastric cancer libraries in the world. By comparing the tags expressed in gastric cancers with those in normal gastric epithelia or each other, many differentially expressed genes were identified. *CEACAM6, APOC1* and *YF13H12* are commonly up-regulated in gastric cancer, while *FUS, CDH17, COL1A1, COL1A2* and *APOE* are associated with invasion and metastasis. RegIV, secreted by cancer cells, inhibits apoptosis, suggesting that RegIV may serve as a novel biomarker and therapeutic target. Production of RNA aptamer should be effective to establish a detection system for blood samples. A custom-made array, named Ex-STOMACHIP, consisting of 395 genes, including highly differentially expressed genes identified by SAGE, is useful to study molecular stomach carcinogenesis and to obtain information about biological behavior and sensitivity to therapy in the clinical setting. The combination of gene expression profiling and determination of genetic polymorphism will allow characterization of individual cancers and patients, leading directly to personalized medicine and cancer prevention.

We would like to thank Noriko Sagawa for help in preparing the manuscript. This work was supported, in part, by Grants-in-Aid for Cancer Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan and from the Ministry of Health, Labour and Welfare of Japan.

1. Ohgaki H, Matsukura N. Stomach cancer. In: Stewart BW, Kleihues P, editors. World cancer report. Lyon: IARC Press; 2003. p. 197–7.
2. Tahara E. Molecular mechanism of stomach carcinogenesis. *J Cancer Res Clin Oncol* 1993; **119**: 265–72.
3. Yasui W, Oue N, Kuniyasu H, Ito R, Tahara E, Yokozaki H. Molecular diagnosis of gastric cancer: present and future. *Gastric Cancer* 2001; **4**: 113–21.
4. Ohgaki H, Yasui W, Yokota J. In: Vainio H, Hietanen E, editors. Handbook of experimental pharmacology. Mechanisms in carcinogenesis and cancer research. Heidelberg: Springer-Verlag; 2003. p. 25–39.
5. Yokozaki H, Yasui W, Tahara E. Genetic and epigenetic changes in stomach cancer. *Int Rev Cytol* 2001; **204**: 49–95.
6. Yasui W, Oue N, Ono S, Mitani Y, Ito R, Nakayama H. Histone acetylation and gastrointestinal carcinogenesis. *Ann N Y Acad Sci* 2003; **983**: 220–31.
7. El-Rifai W, Frierson HF Jr, Harper JC, Powell SM, Knuutila S. Expression profiling of gastric adenocarcinoma using cDNA array. *Int J Cancer* 2001; **92**: 832–8.
8. Hasegawa S, Furukawa Y, Li M, Satoh S, Kato T, Watanabe T, Katagiri T, Tsunoda T, Yamaoka Y, Nakamura Y. Genome-wide analysis of gene expression in intestinal-type gastric cancers using a complementary DNA microarray representing 23,040 genes. *Cancer Res* 2002; **62**: 7–12.
9. Hippo Y, Taniguchi H, Tsutsumi S, Machida N, Chong JM, Fukayama M, Kodama T, Aburatani H. Global gene expression analysis of gastric cancer by oligonucleotide microarrays. *Cancer Res* 2002; **62**: 233–40.
10. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. *Science* 1995; **270**: 484–7.
11. Argani P, Rosty C, Reiter RE, Wilentz RE, Murugesan SR, Leach SD, Ryu B, Skinner HG, Goggins M, Jaffee EM, Yeo CJ, Cameron JL, Kern SE, Hruban RH. Discovery of new markers of cancer through serial analysis of gene expression: prostate stem cell antigen is overexpressed in pancreatic adenocarcinoma. *Cancer Res* 2001; **61**: 4320–4.
12. Nacht M, Dracheva T, Gao Y, Fujii T, Chen Y, Player A, Akmaev V, Cook B, Dufault M, Zhang M, Zhang W, Guo M-Z, Curran J, Han S, Sidranbsky D, Buetow K, Madden SL, Jen J. Molecular characteristics of non-small cell lung cancer. *Proc Natl Acad Sci USA* 2001; **98**: 15203–8.
13. Porter DA, Krop IE, Nasser S, Sgroi D, Kaelin CM, Marks JR, Riggins G, Polyak K. A SAGE (serial analysis of gene expression) view of breast tumor progression. *Cancer Res* 2001; **61**: 5697–702.
14. Buckhaults P, Rago C, St Croix B, Romans KE, Saha S, Zhang L, Vogelstein B, Kinzler KW. Secreted and cell surface genes expressed in benign and malignant colorectal tumors. *Cancer Res* 2001; **61**: 6996–7001.
15. Hough CD, Sherman-Baust CA, Pizer ES, Montz FJ, Im DD, Rosenshein NB, Cho KR, Riggins GJ, Morin PJ. Large-scale serial analysis of gene expression revealed gene differentially expressed in ovarian cancer. *Cancer Res* 2001; **61**: 3869–76.
16. St Croix B, Rago C, Velculescu V, Traverso G, Romans KE, Montgomery E, Lal A, Riggins GJ, Lengauer C, Vogelstein B, Kinzler KW. Gene expressed in human tumor endothelium. *Science* 2000; **289**: 1197–202.
17. The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001; **409**: 92841.
18. Gonzalez CA, Sala N, Capella G. Genetic susceptibility and gastric cancer risk. *Int J Cancer* 2002; **100**: 249–60.
19. El-Rafai W, Moskaluk CA, Abdrabbo MK, Harper J, Yoshida C, Riggins GJ, Frieson HF Jr, Powell SM. Gastric cancers overexpress S100A calcium-binding proteins. *Cancer Res* 2002; **62**: 6823–6.
20. Oien KA, Vass JK, Downie I, Fullarton G, Keith WN. Profiling, comparison and validation of gene expression in gastric carcinoma and normal stomach. *Oncogene* 2003; **22**: 4287–300.
21. Lee J-Y, Eom E-M, Kim D-S, Ha-Lee YM, Lee D-H. Analysis of gene expression profiles of gastric normal and cancer tissues by SAGE. *Genomics* 2003; **82**: 78–85.
22. Oue N, Hamai Y, Mitani Y, Matsumura S, Oshimo Y, Aung PP, Kuraoka K, Nakayama H, Yasui W. Gene expression profile of gastric carcinoma: identification of genes and tags potentially involved in invasion, metastasis, and carcinogenesis by serial analysis of gene expression. *Cancer Res* 2004; **64**: 2397–405.
23. Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma. 13th ed. Tokyo: Kanehara Publishers; 1999.
24. Yu J, Hwang PM, Rago C, Kinzler KW, Vogelstein B. Identification and classification of p53-regulated genes. *Proc Natl Acad Sci USA* 1999; **96**: 14517–22.
25. International Union Against Cancer. TNM classification of malignant tumors. 6th ed. New York: Wiley-Liss; 2002.
26. Hicks GG, Singh N, Nashabi A, Mai S, Bozek G, Klewes L, Arapovic D, White EK, Koury MJ, Oltz EM, Van Kaer L, Ruley HE. Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nat Genet* 2000; **24**: 175–9.
27. Grotzinger C, Kneifel J, Patschan D, Schnoy N, Anagnostopoulos I, Faiss S, Tauber R, Wiedenmann B, Gessner R. LI-cadherin: a marker of gastric metaplasia and neoplasia. *Gut* 2001; **49**: 73–81.
28. Hartupee JC, Zhang H, Bonaldo MF, Soares MB, Dieckgraefe BK. Isolation and characterization of a cDNA encoding a novel member of the human generating protein family. *Biochim Biophys Acta* 2001; **1518**: 287–93.
29. Violette S, Festor E, Pandrea-Vasile I, Mitchell V, Adida C, Dussaulx E, Lacorte JM, Chambaz J, Lacasa M, Lesuffleur T. Reg IV, a new member of the regenerating gene family, is overexpressed in colorectal carcinomas. *Int J Cancer* 2003; **103**: 185–93.
30. Jayasena SD. Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clin Chem* 1999; **45**: 1628–50.
31. Cerchia L, Hamm J, Libri D, Tavitian B, de Franciscis V. Nucleic acid aptamers in cancer medicine. *FEBS Lett* 2002; **528**: 12–6.

32. Inoue H, Matsuyama A, Mimori K, Ueo H, Mori M. Prognostic score of gastric cancer determined by cDNA microarray. *Clin Cancer Res* 2002; **8**: 3475–9.
33. Zembutsu H, Ohnishi Y, Tsunoda T, Furukawa Y, Katagiri T, Ueyama Y, Tamaoki N, Nomura T, Kitahara O, Yanagawa R, Hirata K, Nakamura Y. Genome-wide cDNA microarray screening to correlate gene expression profiles with sensitivity of 85 human cancer xenografts to anticancer drugs. *Cancer Res* 2002; **62**: 518–27.
34. Watters JW, McLeod HL. Cancer pharmacogenomics: current and future applications. *Biochim Biophys Acta* 2003; **1603**: 99–111.
35. Sawiris GP, Sherman-Baust CA, Becker KG, Cheadle C, Teichberg D, Morin PJ. Development of a highly specialized cDNA array for the study and diagnosis of epithelial ovarian cancer. *Cancer Res* 2002; **62**: 2923–8.
36. Yasui W, Yokozaki H, Shimamoto F, Tahara H, Tahara E. Molecular-pathological diagnosis of gastrointestinal tissues and its contribution to cancer histopathology. *Pathol Int* 1999; **49**: 763–74.

ADAM8 as a Novel Serological and Histochemical Marker for Lung Cancer

Nobuhisa Ishikawa,[1,2] Yataro Daigo,[1] Wataru Yasui,[3] Kouki Inai,[4] Hitoshi Nishimura,[5] Eiju Tsuchiya,[6] Nobuoki Kohno,[2] and Yusuke Nakamura[1]

[1]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo; Departments of [2]Molecular and Internal Medicine, [3]Molecular Pathology, and [4]Pathology, Graduate School of Biomedical Sciences, Hiroshima University, Hiroshima; [5]Department of Thoracic Surgery, Saitama Cancer Center, Saitama; and [6]Kanagawa Cancer Center Research Institute, Kanagawa, Japan

ABSTRACT

Purpose and Experimental Design: **We have been investigating genes involved in pulmonary carcinogenesis by examining gene expression profiles of non–small-cell lung cancers to identify molecules that might serve as diagnostic markers or targets for development of new molecular therapies. A gene encoding ADAM8, a disintegrin and metalloproteinase domain-8, was selected as a candidate for such molecule. Tumor tissue microarray was applied to examine expression of ADAM8 protein in archival lung cancer samples from 363 patients. Serum ADAM8 levels of 105 lung cancer patients and 72 controls were also measured by ELISA. A role of ADAM8 in cellular motility was examined by Matrigel assays.**

Results: **ADAM8 was abundantly expressed in the great majority of lung cancers examined. A high level of ADAM8 expression was significantly more common in advanced-stage IIIB/IV adenocarcinomas than in adenocarcinomas at stages I–IIIA. Serum levels of ADAM8 were significantly higher in lung cancer patients than in healthy controls. The proportion of the serum ADAM8-positive cases defined by our criteria was 63% and that for carcinoembryonic antigen was 57%, indicating equivalent diagnostic power of these two markers. A combined assay using both ADAM8 and carcinoembryonic antigen increased sensitivity because 80% of the lung cancer patients were then diagnosed as positive, whereas only 11% of 72 healthy volunteers were falsely diagnosed as positive. In addition, exogenous expression of ADAM8 increased the migratory activity of mammalian cells, an indication that ADAM8 may play a significant role in progression of lung cancer.**

Conclusions: **Our data suggest that ADAM8 should be useful as a diagnostic marker and probably as a therapeutic target.**

INTRODUCTION

Lung cancer is one of the most common cancers in the world, and non–small-cell lung cancer (NSCLC) accounts for ~80% of those cases (1). Because the prognosis of advanced lung cancer remains poor, development of novel therapeutic and diagnostic strategies is an urgent goal (2). Tumor markers that are currently available for lung cancer, such as carcinoembryonic antigen (CEA), serum cytokeratin 19 fragment (CYFRA 21-1), and progastrin-releasing peptide (pro-GRP), are not satisfactory for diagnosis at an early stage or for monitoring the disease because of their relatively low sensitivity and specificity in detecting the presence of cancer cells (3–5). Although the precise pathways involved in lung tumorigenesis remain unclear, some evidence indicates that tumor cells express cell surface markers unique to each histologic type at particular stages of differentiation. Because cell surface proteins are considered more accessible to immune mechanisms and drug delivery systems, identification of cancer-specific cell surface and secretory proteins is likely to be an effective approach to development of novel diagnostic markers and therapeutic strategies.

We have been screening genes encoding transmembrane/secretory proteins that are up-regulated in lung cancers, with cDNA microarrays and tumor cells purified by laser-capture microdissection (6, 7). To verify the biological and clinicopathological significance of the respective gene products, we have been performing tumor tissue microarray analysis of clinical lung cancer materials. This systematic approach revealed that a disintegrin and metalloproteinase domain-8 (ADAM8), a cell surface disintegrin and metalloproteinase domain-8 molecule, was frequently transactivated in primary lung cancers.

ADAM family members are implicated to be involved in the proteolytic processing of membrane-bound precursors, and they modulate cell-cell and cell-matrix interactions. ADAM8 encodes a protein of 824 amino acids with a COOH-terminal transmembrane domain and potential extracellular adhesion and protease domains (8, 9). This molecule, localized to the plasma membrane, is processed by autocatalysis into two forms; one is derived by removal of a prodomain and the other is a remnant protein composed of the extracellular region with a disintegrin domain at the NH_2 terminus (10). ADAM8 behaves as an active metalloprotease *in vitro*, hydrolyzing myelin basic protein and a variety of peptide substrates based on the cleavage sites of membrane-bound cytokines, growth factors, and receptors (11–

Received 7/21/04; revised 9/6/04; accepted 9/15/04.
Grant support: "Research for the Future" Program Grant 00L01402 of The Japan Society for the Promotion of Science (Y. Nakamura).
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.
Requests for reprints: Yusuke Nakamura, Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. Phone: 81-3-5449-5372; Fax: 81-3-5449-5433; E-mail: yusuke@ims.u-tokyo.ac.jp.

14). Other studies have demonstrated overexpression of some ADAM family proteins in a variety of human tumors (15, 16), but involvement of ADAM8 in human cancer was not indicated previously.

We report here the identification of ADAM8 as a novel diagnostic marker and a potential target for therapeutic agents/antibodies and also provide evidence for its possible role in human pulmonary carcinogenesis.

MATERIALS AND METHODS

Cell Lines and Clinical Samples. The 23 human lung cancer cell lines used in this study included nine adenocarcinomas (A427, A549, LC319, NCI-H1373, PC-3, PC-9, PC-14, NCI-H1666, and NCI-H1781), nine squamous cell carcinomas (EBC-1, LU61, NCI-H520, NCI-H1703, NCI-H2170, RERF-LC-AI, SK-MES-1, NCI-H226, and NCI-H647), one large-cell carcinoma (LX1), and four small-cell lung cancers (SCLCs; DMS114, DMS273, SBC-3, and SBC-5). All cells were grown in monolayers in appropriate media supplemented with 10% FCS and were maintained at 37°C in an atmosphere of humidified air with 5% CO_2.

Surgically resected primary NSCLC samples had been obtained earlier with informed consent (6). A total of 363 formalin-fixed primary tumors (206 adenocarcinomas, 118 squamous cell carcinomas, 25 large-cell carcinomas, and 14 SCLCs) and adjacent normal lung tissue samples from 336 patients undergoing surgery at Saitama Cancer Center (Saitama, Japan) and from postmortem materials (27 individuals) obtained from Hiroshima University (Hiroshima, Japan) were also used in this study.

Serum samples were obtained with informed consent from 72 healthy individuals as controls (40 males and 32 females; median age, 61.5 ± 9.1 SD; range, 42–89 years). The healthy individuals showed no abnormalities in complete blood cell counts, C-reactive proteins, erythrocyte sedimentation rates, liver function tests, renal function tests, urinalyses, fecal examinations, chest X-rays, or electrocardiograms. Serum samples were also obtained with informed consent from 105 lung cancer patients admitted to Hiroshima University Hospital (78 males and 27 females; median age, 68.0 ± 10.8 SD; range, 30–84 years). Samples were selected for the study on the basis of the following criteria: (*a*) patients were newly diagnosed and previously untreated and (*b*) their tumors were pathologically diagnosed as lung cancers (stages I–IV). These 105 cases included 62 adenocarcinomas, 25 squamous cell carcinomas, and 18 SCLCs. Clinicopathological records were fully documented. Serum was obtained at the time of diagnosis and stored at −80°C. Disease staging in all 105 cases was supported by a computed tomography scan of the chest and abdomen, bone scintigraphy, and magnetic resonance imaging of the head.

Semiquantitative Reverse Transcription-PCR Analysis. Total RNA was extracted from cultured cells and clinical tissues with Trizol reagent (Life Technologies, Inc., Gaithersburg, MD), according to the manufacturer's protocol. Extracted RNAs and normal human tissue polyA RNAs were treated with DNase I (Roche Diagnostics, Basel, Switzerland) and then reverse transcribed with oligo$(dT)_{12-18}$ primer and SuperScript II reverse transcriptase (Life Technologies, Inc.). Semiquantitative reverse transcription-PCR experiments were carried out with synthesized *ADAM8* gene-specific primers (5′-GTGTGTGTACGTGTCTCCAGGT-3′ and 5′-CAGACAAGATAGCTGACTCTCCC-3′), or with β-actin (*ACTB*)-specific primers (5′-ATCAAGATCATTGCTCCTCCT-3′ and 5′-CTGCGCAAGTTAGGTTTTGT-3′) as an internal control. All PCR reactions involved initial denaturation at 94°C for 2 minutes followed by 22 (for *ACTB*) or 28 cycles (for *ADAM8*) of 94°C 30 s, 54°C to 60°C for 30 seconds, and 72°C for 60 seconds on a GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA).

Northern Blot Analysis. Human multiple tissue blots (BD Biosciences, Palo Alto, CA) were hybridized with ^{32}P-labeled PCR products. Prehybridization, hybridization, and washing were done according to the supplier's recommendations. The blots were autoradiographed with intensifying screens at −80°C for 1 week.

Immunohistochemistry and Tissue Microarray. Tumor tissue microarrays were constructed with 363 formalin-fixed primary lung cancers, as published previously (17–19). The tissue area for sampling was selected based on visual alignment with the corresponding H&E-stained section on a slide. Three, four, or five tissue cores (diameter, 0.6 mm; height, 3–4 mm) taken from a donor tumor block were placed into a recipient paraffin block with a tissue microarrayer (Beecher Instruments, Sun Prairie, WI). A core of normal tissue was punched from each case, and 5-μm sections of the resulting microarray block were used for immunohistochemical analysis.

Three independent investigators assessed ADAM8 positivity semiquantitatively, recording staining intensity as absent (scored as 0), weak (scored as 1+), or strongly positive (scored as 2+), without prior knowledge of clinicopathological data. Cases were accepted as strongly positive only if reviewers independently defined them as such. Contingency tables were used to analyze the relationship of ADAM8 expression in NSCLCs to clinicopathological data.

To investigate the presence of ADAM8 protein in clinical samples that had been embedded in paraffin blocks, we stained the sections in the following manner. Briefly, 50 μg/mL goat polyclonal antihuman ADAM8 antibody (R&D Systems, Inc., Minneapolis, MN) were added after blocking of endogenous peroxidase and proteins, and the sections were incubated with horseradish peroxidase-labeled antigoat IgG (Histofine Simple Stain MAX PO (G), Nichirei, Tokyo, Japan) as the secondary antibody. Substrate-chromogen was added, and the specimens were counterstained with hematoxylin.

Flow Cytometric Analysis. Lung cancer cells (1×10^6 cells) were incubated with a mouse monoclonal antihuman ADAM8 antibody for detecting the ectodomain of the protein (0.34 mg/mL; R&D Systems, Inc.) or control mouse IgG (0.34 mg/mL; Santa Cruz Biotechnology, Santa Cruz, CA) at 4°C for 1 hour. The cells were washed in PBS and then incubated with AlexaFluor 488-conjugated antimouse IgG (Molecular Probes, Eugene, OR) at 4°C for 30 minutes. The cells were washed in PBS and analyzed on a FACScan flow cytometer (Becton Dickinson Labware, Bedford, MA) and analyzed by ModFit software (Verity Software House, Inc., Topsham, ME). Mean fluorescence intensity was calculated as a relative signal-intensity value, *i.e.,* of cells treated with antihuman ADAM8 antibody/cells treated with control mouse IgG.

ELISA. Serum levels of ADAM8 were measured by ELISA using a commercially available enzyme test kit (R&D Systems, Inc.). In brief, 3-fold diluted sera were added to a 96-well microplate precoated with monoclonal antibody specific for ADAM8 and incubated for 2 hours at room temperature. After washing away any unbound substances, an enzyme-linked polyclonal antibody specific for ADAM8 was added to the wells and incubated for 2 hours at room temperature. After a wash to remove any unbound antibody-enzyme reagent, a substrate solution (R&D Systems, Inc.) was added to the wells and allowed to react for 30 minutes. The reaction was stopped by adding 100 μL of 2 N sulfuric acid. Color intensity was determined by a photometer at a wavelength of 492 nm, with a reference wavelength of 630 nm. Levels of CEA in serum were measured by ELISA with a commercially available enzyme test kit (HOPE Laboratories, Belmont, CA), according to the same protocol as above. Differences in the levels of ADAM8 and CEA between tumor groups and a healthy control group were analyzed by Mann-Whitney *U* tests. The levels of ADAM8 and CEA were additionally evaluated by receiver-operating characteristic curve analysis to determine cutoff levels with optimal diagnostic accuracy and likelihood ratios. The correlation coefficients for these two markers were calculated with Pearson's correlation coefficient. Significance was defined as $P < 0.05$.

Matrigel Invasion Assay. NIH3T3 and COS-7 cells transfected either with plasmids expressing ADAM8 or with mock plasmids were grown to near confluence in DMEM containing 10% FCS. The cells were harvested by trypsinization, washed in DMEM without addition of serum or proteinase inhibitor, and suspended in DMEM at 1×10^5 cells/mL. Before preparing the cell suspension, the dried layer of Matrigel matrix (Becton Dickinson Labware) was rehydrated with DMEM for 2 hours at room temperature. DMEM (0.75 mL) containing 10% FCS was added to each lower chamber in 24-well Matrigel invasion chambers, and 0.5 mL (5×10^4 cells) of cell suspension were added to each insert of the upper chamber. The plates of inserts were incubated for 22 hours at 37°C. After incubation, the chambers were processed; cells invading through the Matrigel were fixed and stained by Giemsa as directed by the supplier (Becton Dickinson Labware).

RESULTS

ADAM8 Expression in Lung Tumors, Cell Lines, and Normal Tissues. To search for novel molecules to serve as diagnostic markers and/or targets for development of therapeutic agents for lung cancer, we had applied cDNA microarray analysis to identify genes that were transactivated in the majority of NSCLCs examined. Among 23,040 genes screened, we identified the *ADAM8* transcript, indicating 2-fold or higher expression (mean fold expression, 5.6 ± 4.5 SD; range, 2.2–16.4) in cancer cells than in normal lung cells (control) in 63% of the NSCLC samples examined, and confirmed its transactivation by semiquantitative reverse transcription-PCR experiments in 8 of 10 additional NSCLC tissues and in 11 of 23 lung cancer cell lines (Fig. 1*A*). Northern blot analysis with human cDNA as a probe detected a 3.5-kb transcript but at a very low level and only in leukocytes, lymph node, and bone marrow among 16 normal tissues examined (data not shown).

We also examined expression of ADAM8 protein in clin-

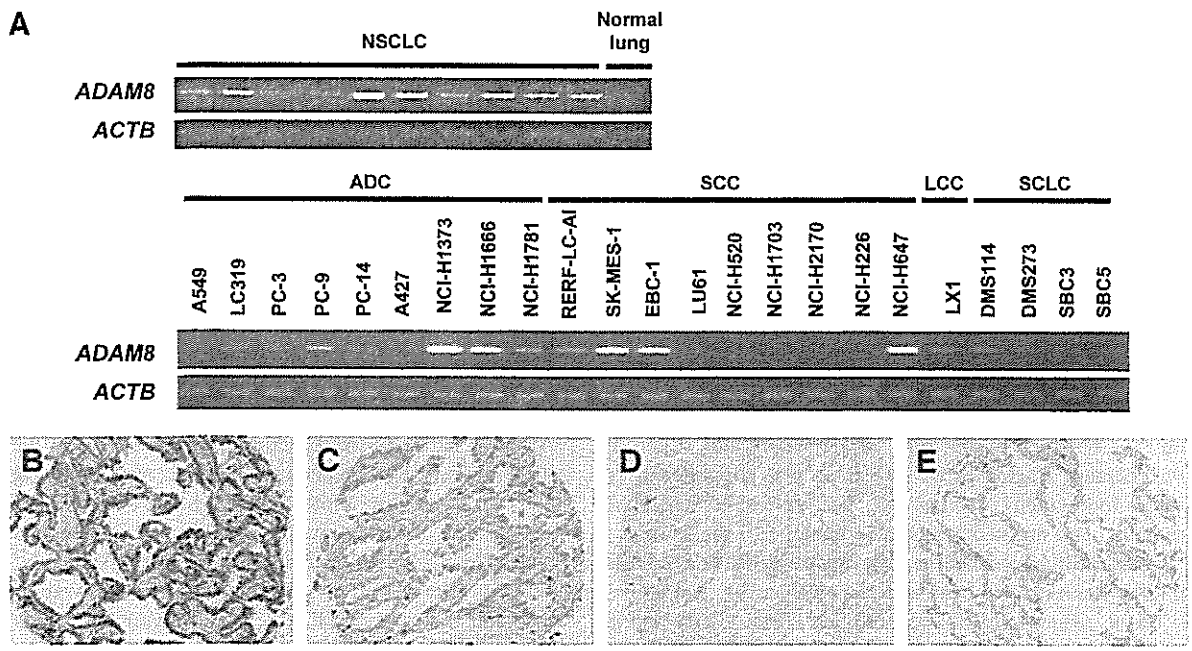

Fig. 1 Validation of ADAM8 expression in lung tumors and cell lines. *A*, expression of *ADAM8* in a normal lung tissue and 10 clinical NSCLC samples (*top panel*) and 23 lung cancer cell lines (*bottom panel*) detected by semiquantitative reverse transcription-PCR analysis. ADC, adenocarcinoma; SCC, squamous cell carcinoma; LCC, large-cell carcinoma; SCLC, small-cell lung cancer. *B–E*, representative images after immunohistochemical staining of lung tissues, with anti-ADAM8 antibody on tumor tissue microarrays (×100). Examples are shown of strong (*B*), weak (*C*), and absent (*D*) ADAM8 expression in ADCs and of no expression in normal lung (*E*).

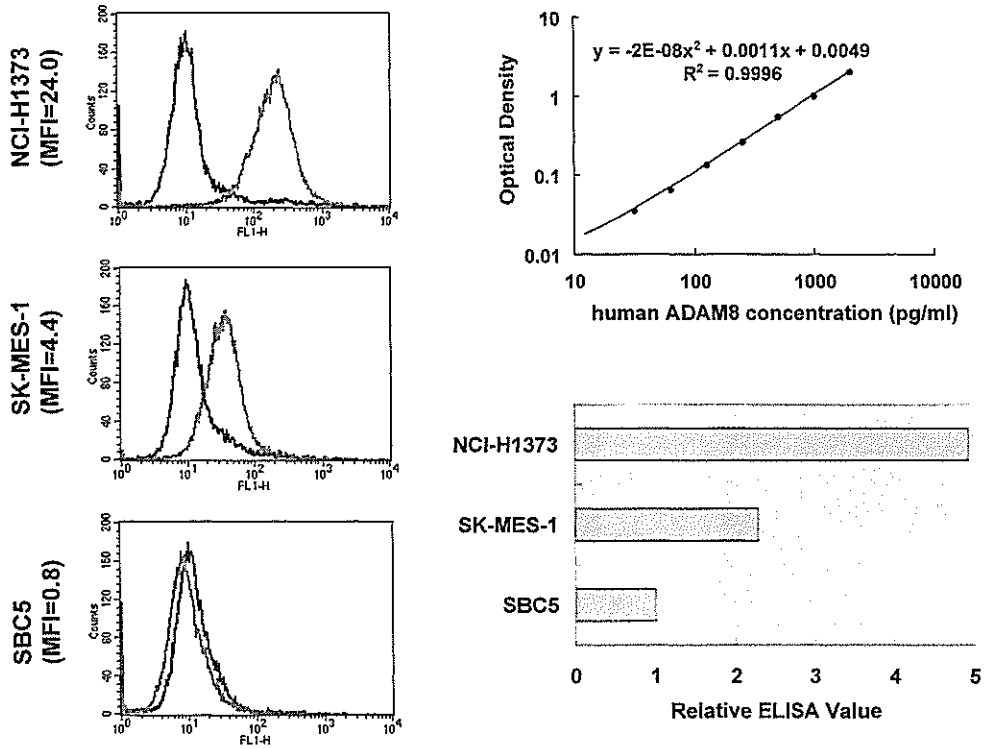

Fig. 2 Cellular localization and secretion of ADAM8 in lung cancer cells. *A*, expression of ADAM8 on cell surfaces in lung cancer lines NCI-H1373, SK-MES-1, and SBC-5 evaluated by flow cytometric analysis. Mean fluorescence intensity (MFI) was calculated as relative signal intensity values of cells treated with antihuman ADAM8 antibody (*green*)/cells treated with mouse IgG (control; *black*). *B*, standard curve of sandwich ELISA with anti-ADAM8 antibody (*X* axis, ADAM8 concentration; *Y* axis, absorbance). *C*, specific detection of ADAM8 with ELISA in conditioned medium from ADAM8-expressing NCI-H1373 and SK-MES-1 cells and nonexpressing SBC-5 cells. The *X* axis indicates the relative ELISA values (concentration of secreted ADAM8 from NCI-H1373 or SK-MES-1/concentration of secreted ADAM8 from SBC-5 cells).

ical lung cancers by tissue arrays with anti-ADAM8 antibody. ADAM8 localized at the plasma membrane, as well as in the cytoplasm of tumor cells, but was hardly detectable in surrounding normal tissues (Fig. 1*B–E*). Positive staining was observed in 158 (77%) of 206 adenocarcinomas cases examined, 55 (47%) of 118 squamous cell carcinomas, 16 (64%) of 25 large-cell carcinomas, and 11 (79%) of 14 SCLCs, whereas no staining was observed in any of the normal portions of the same tissues. We classified a pattern of ADAM8 expression on the tissue array ranging from absent/weak (scored as 0 ~ 1+) to strong (scored as 2+). Expression levels of ADAM8 were not associated with any of the clinicopathological factors in squamous cell carcinomas. However in adenocarcinomas, strong ADAM8 staining was significantly more common in stages IIIB–IV tumors (22 of 60; 37%) than in stages I–IIIA tumors (34 of 146, 23%; $P = 0.049$; χ^2 test). The sample sizes of large-cell carcinomas and SCLCs were too small to be evaluated additionally.

Secretion of ADAM8 in Lung Cancer Cells. We then validated ADAM8 expression on the surfaces of lung cancer cells with flow cytometry and anti-ADAM8 monoclonal antibody. This analysis indicated that the antibody bound to NCI-H1373 and SK-MES-1 cells in which ADAM8 transcript had been detected at a high level but not to SBC-5 cells, which did not express ADAM8 (Fig. 2*A*).

Because the extracellular domain of ADAM8 protein is thought to be secreted (10), we applied ELISA method to examine its presence in the culture media of these cell lines. High levels of ADAM8 protein were detected in media of NCI-H1373 and SK-MES-1 cultures but not in the medium of SBC-5 cells (Fig. 2, *B* and *C*). The amounts of detectable ADAM8 in the culture media accorded well with the expression levels of ADAM8 on the cell surfaces detected with flow cytometric analysis.

Serum Levels of ADAM8 in Lung Cancer Patients. Because the *in vitro* findings had suggested a possibility for development of a novel tumor maker for lung cancer, we investigated whether the extracellular domain of ADAM8 is secreted into sera of patients with lung cancer. ELISA experiments detected ADAM8 in serologic samples from lung cancer patients and also from normal individuals; serum levels of ADAM8 in lung cancer patients were 431 ± 249 pg/mL (mean ± SD) and those in healthy individuals were 267 ± 56 pg/mL. The difference was significant with P of <0.001 (Mann-Whitney U test). When classified according to histologic type, the serum levels of ADAM8 were 427 ± 286 pg/mL in adenocarcinomas patients, 467 ± 210 pg/mL in squamous cell carcinomas patients, and 400 ± 112 pg/mL in SCLC patients (Fig. 3*A*); the differences among the three histologic types were not significant. High levels of serum

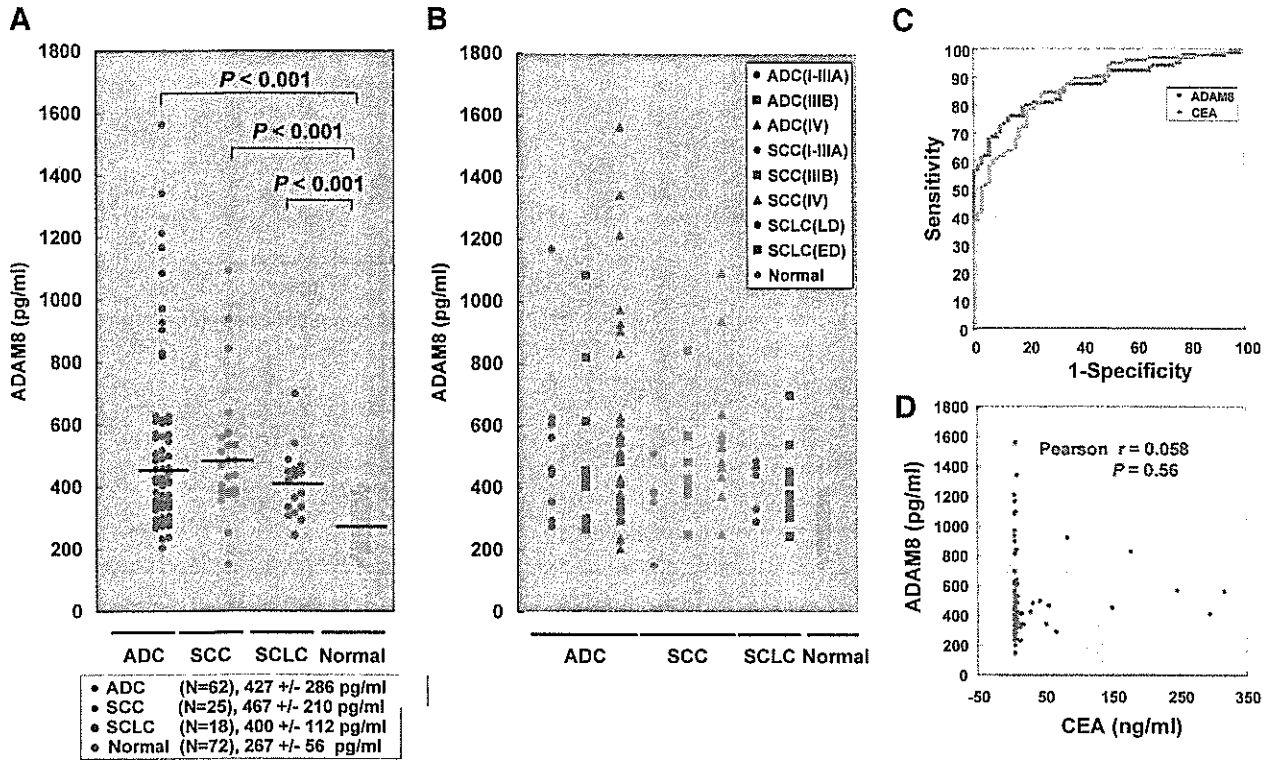

Fig. 3 Serologic concentration of ADAM8 determined by ELISA in patients with lung cancers and in healthy controls. *A*, distribution of ADAM8 in sera from patients with lung adenocarcinoma (ADC), lung squamous cell carcinoma (SCC), or SCLC. Averaged serum levels are shown as *black lines*. Differences were significant between ADC patients and healthy individuals ($P < 0.001$, Mann-Whitney U test), between SCC patients and healthy individuals ($P < 0.001$) and between SCLC patients and healthy individuals ($P < 0.001$). *B*, distribution of ADAM8 in sera from patients at various clinical stages of lung cancers. LD indicates limited disease; ED, extensive disease. *C*, receiver-operating characteristic curve analysis of ADAM8 (*blue*) and CEA (*red*) as serum markers for lung cancer (X axis, 1-specificity; Y axis, sensitivity). *D*, relationship between serum levels of ADAM8 and CEA (X axis, CEA concentration; Y axis, ADAM8 concentration).

ADAM8 were detected even in patients with earlier-stage tumors (stages I–IIIA; Fig. 3*B*). We also found no significant association between the serum ADAM8 level and age or gender (Table 1).

Comparison of ADAM8 and CEA as Tumor Markers. To evaluate the feasibility of using serum ADAM8 level as a tumor detection marker, we also measured by ELISA serum levels of CEA, a conventional tumor marker, in the same patients and controls. ADAM8 and CEA were additionally analyzed by drawing receiver-operating characteristic curves to determine their cutoff levels (Fig. 3*C*). The sum of the area under the receiver-operating characteristic curve for serum ADAM8 value was slightly larger than that for serum CEA, suggesting slightly better specificity and likelihood for ADAM8 as diagnostic marker for lung cancer. Cutoff levels in this assay were set to result in optimal diagnostic accuracy and likelihood ratios for ADAM8 and CEA, *i.e.*, 379 pg/mL for ADAM8 and 5.3 ng/mL for CEA. As shown in Fig. 3*D*, the correlation coefficient between serum ADAM8 and CEA values was not significant (Pearson's correlation: r = 0.058, $P = 0.56$), indicating that measuring both markers in serum can improve overall sensitivity for detection of NSCLC to 80% (for diagnosing NSCLC, the sensitivity of CEA alone is 57% and that of ADAM8 is 63%.). False-positive results for either of the two tumor markers among 72 normal volunteers (control group) amounted to 11% (8 of 72), whereas the false-positive rates for CEA and ADAM8 in the same control group were 5.5% (4 of 72) each.

Table 1 Correlation of serum ADAM8 level with age or gender

Factors	Serum ADAM8 Percentage (%)	No. of patients	Mean ± SD	*P*
Lung cancer patients				
Gender				
Male	74	78	442 ± 232	0.54*
Female	26	27	418 ± 299	
Age (y)				
≤65	43	45	413 ± 288	0.26*
>65	57	60	445 ± 218	
Normal individuals				
Gender				
Male	56	40	267 ± 52	0.65*
Female	44	32	273 ± 62	
Age (y)				
≤65	71	51	266 ± 53	0.08*
>65	29	21	308 ± 60	

* Mann-Whitney U test.

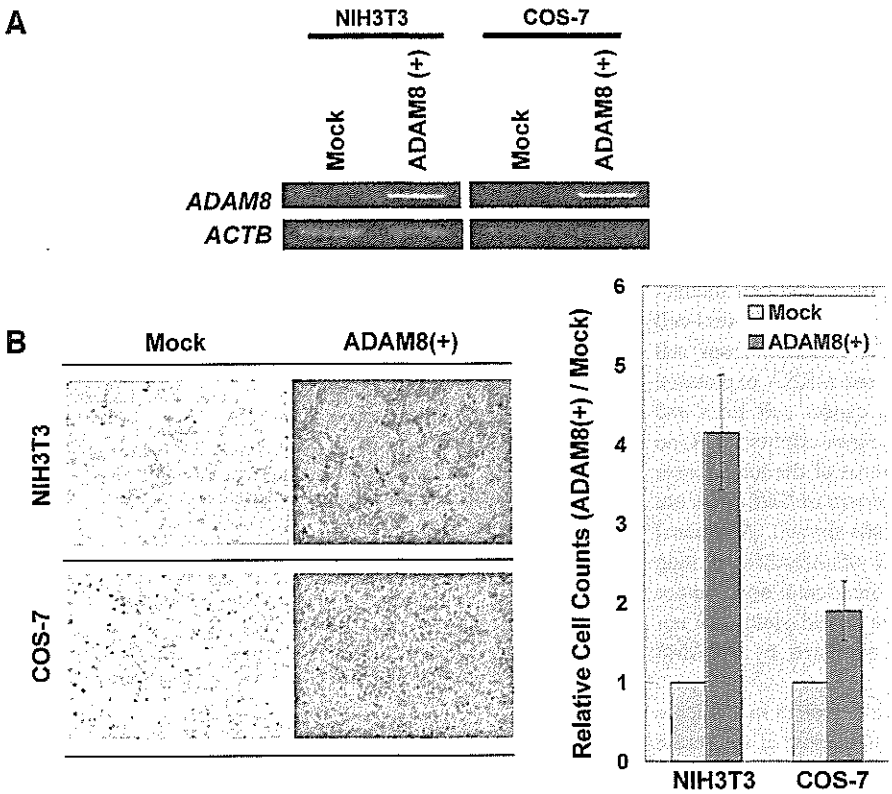

Fig. 4 Promotion of invasiveness of mammalian cells transfected with ADAM8-expressing plasmids. *A*, transient expression of *ADAM8* in NIH3T3 and COS-7 cells, detected by semiquantitative reverse transcription-PCR analysis. *B*, assay demonstrating the invasive nature of NIH3T3 and COS-7 cells in Matrigel matrix after transfection with expression plasmids for human *ADAM8*. *Left panel* shows Giemsa staining (×200), and *right panel* represents the relative number of cells migrating through the Matrigel-coated filters (cells with ADAM8-expressing plasmids/cells with mock plasmids). Assays were performed three times and in triplicate wells.

Activation of Cellular Migration/Invasive Activity by ADAM8. ADAM8 protein contains conserved ectodomains that are considered important for cell adhesion and protease activities; other ADAM family members are known to be associated with cellular motility and invasion (9). Hence, we examined a possible role of ADAM8 in cellular motility and invasion in Matrigel assays, with NIH3T3 and COS-7 cells. As shown in Fig. 4, transfection of ADAM8 cDNA into either cell line significantly enhanced its invasive activity through Matrigel, compared with cells transfected with mock vector.

DISCUSSION

Despite many advances in diagnostic imaging of tumors, combination chemotherapy, and radiation therapy, little improvement has been achieved within the last decade in terms of prognosis and quality of life for patients with lung cancer. Given the frequent failure of conventional treatment strategies, many cancer-related molecules have been characterized toward the goal of developing novel anticancer therapies such as molecular-targeted drugs and antibodies or cancer vaccines (20, 21). Molecular-targeted therapies are expected to be highly specific to malignant cells, with minimal adverse effects due to their well-defined mechanisms of action. Equally desirable in prospect are minimally invasive, highly sensitive, and specific new diagnostic methods that would adapt readily to clinical settings. From these points of view, tumor-specific transmembrane/secretory proteins should have significant advantages because they are presented either on the cell surface or within the extracellular space and/or in serum, making them easily accessible as molecular markers and therapeutic targets. Some tumor-specific markers already available, such as CYFRA or Pro-GRP, are transmembrane/secretory proteins (4, 5); the example of rituximab (Rituxan), a humanized monoclonal antibody against CD20-positive lymphomas, provides proof that targeting specific cell surface proteins can result in significant clinical benefits (22). As an approach to identifying novel cancer-specific cell surface or secretory proteins, we have been exploiting the power of genome-wide expression analysis to select genes that are overexpressed in cancer cells. Analysis of candidate molecules revealed ADAM8 as a potential target for development of novel tools for diagnosis and treatment of lung cancer.

ADAM8 protein is homologous to a snake disintegrin, Reprolysin (M12B), a zinc metalloprotease (9). Members of the ADAM family are cell surface proteins with a unique structure combining potential adhesion and protease domains. A published report has suggested that the ADAM8 ectodomain is cleaved by ADAM8 itself (10). Because various matrix metalloproteinases and ADAM family proteins had been described as being overexpressed in human cancers (23), ADAM8 seemed likely to have a potential role in tumor development or progression. In this study, we demonstrated that induction of exogenous expression of ADAM8 enhanced the cellular migration/invasive activity of mammalian cells. Concordantly, the strong ADAM8 staining in primary NSCLC tissues detected by tissue microarray analyses correlated with disease progression; overexpression of ADAM8 protein was significantly more common in tumors from patients with locally advanced lung cancer (stage IIIB) or distant organ metastases (stage IV) than in those with earlier

stage diseases (stages I–IIIA). The former patients are generally not candidates for curative resection (24). Although the precise mechanism of ADAM8 in lung carcinogenesis is unknown and the processes of cancer-cell invasion to adjacent tissues and distant metastasis consist of a complex series of sequential step, these results indicate that ADAM8 expression and its cleavage could promote dissemination of tumors by stimulating cell migration. ADAM8 has been described as a shedding enzyme of the low-affinity IgE receptor CD23 (13); it also appears to play an important role in physiologic and pathological cell interactions by specifically releasing a functional form of a neural-cell adhesion molecule, a homologue of L1 (CHL1), from cell surfaces (14). However, when we analyzed mRNA expression of *ADAM8*, *CD23*, and *CHL1* in lung cancer cell lines and cancer tissues by semiquantitative reverse transcription-PCR, no expression of *CD23* was detectable in most of the lung cancer samples examined, and the expression pattern of *CHL1* was not concordant with that of *ADAM8* (data not shown). Additional studies to identify unknown substrates of ADAM8 in lung cancers may contribute not only to identification of novel tumor markers and therapeutic targets but also should yield new understanding of the signaling pathway mediated by ADAM8 expression.

We also found high levels of ADAM8 protein in serologic samples from lung cancer patients. Because most of the serum samples used for were derived from patients with advanced cancers (stages IIIB–IV), we evaluated ADAM8 as a tumor marker for early diagnosis of this disease. Tissue microarray analyses of NSCLCs at stages I–IV documented ADAM8 positivity even in early-stage tumors; ADAM8 staining was observed in 176 (64%) of 274 cases at stages I–IIIA, including 111 (76%) of 146 adenocarcinomas cases, whereas as mentioned above, strong ADAM8 staining was observed more frequently in adenocarcinomas cases at advanced stages IIIB–IV than in cases at stages I–IIIA. These results indicated that ADAM8 should be useful for diagnosis of even early-stage lung cancers and that the high level of ADAM8 expression could indicate increased malignant potential of ADAM8-positive tumors.

To validate the feasibility of applying ADMA8 as the diagnostic tool, we compared serum levels of ADAM8 with those of CEA, a conventional diagnostic marker for NSCLCs, in terms of sensitivity and specificity for diagnosis. The proportions of positive cases among the same serum samples were 57% for CEA and 63% for ADAM8, whereas the false-positive rate for ADAM8 (5.5%) was the same as that of CEA, indicating equivalent or better diagnostic power of ADAM8. Furthermore, an assay combining both markers increased the sensitivity such that 80% of the patients with lung cancer were diagnosed as positive, whereas 11% of 72 healthy volunteers were falsely diagnosed as positive. Although additional validation with a larger set of serum samples covering various clinical stages will be necessary, the data presented here sufficiently demonstrate a potential clinical application of ADAM8 itself as a serologic/histochemical marker for lung cancers. We should mention also that we observed activation of *ADAM8* in nearly half of a group of pancreatic cancers, which have a significantly invasive phenotype (data not shown). This suggests that overexpression of *ADAM8* might play a significant role in progression of pancreatic cancer and could warrant investigation of serum levels of ADAM8 in patients with other types of invasive cancer as well.

In conclusion, we have identified ADAM8 as a potential marker for diagnosis of lung cancers. This molecule is also a likely candidate for development of therapeutic approaches such as antibody therapy. ADAM8 was specifically overexpressed in most lung cancer tissues we examined and was elevated in the sera of a large proportion of patients with lung cancer. ADAM8, combined with other tumor markers, could significantly improve the sensitivity of lung cancer diagnosis.

REFERENCES

1. Greenlee RT, Murray T, Bolden S, Wingo PA. Cancer statistics, 2000. CA - Cancer J Clin 2000;50(1):7–33.
2. Naruke T, Tsuchiya R, Kondo H, Asamura H. Prognosis and survival after resection for bronchogenic carcinoma based on the 1997 TNM-staging classification: the Japanese experience. Ann Thorac Surg 2001; 71(6):1759–64.
3. Shinkai T, Saijo N, Tominaga K, et al. Serial plasma carcinoembryonic antigen measurement for monitoring patients with advanced lung cancer during chemotherapy. Cancer (Phila.) 1986;57(7):1318–23.
4. Pujol JL, Grenier J, Daures JP, Daver A, Pujol H, Michel FB. Serum fragment of cytokeratin subunit 19 measured by CYFRA 21–1 immunoradiometric assay as a marker of lung cancer. Cancer Res 1993;53(1): 61–6.
5. Miyake Y, Kodama T, Yamaguchi K. Pro-gastrin-releasing peptide(31–98) is a specific tumor marker in patients with small cell lung carcinoma. Cancer Res 1994;54(8):2136–40.
6. Kikuchi T, Daigo Y, Katagiri T, et al. Expression profiles of non-small cell lung cancers on cDNA microarrays: identification of genes for prediction of lymph-node metastasis and sensitivity to anti-cancer drugs. Oncogene 2003;22(14):2192–205.
7. Suzuki C, Daigo Y, Kikuchi T, Katagiri T, Nakamura Y. Identification of COX17 as a therapeutic target for non-small cell lung cancer. Cancer Res 2003;63(21):7038–41.
8. Yoshiyama K, Higuchi Y, Kataoka M, Matsuura K, Yamamoto S. CD156 (human ADAM8): expression, primary amino acid sequence, and gene location. Genomics 1997;41(1):56–62.
9. Yamamoto S, Higuchi Y, Yoshiyama K, et al. ADAM family proteins in the immune system. Immunol Today 1999;20(6):278–84.
10. Schlomann U, Wildeboer D, Webster A, et al. The metalloprotease disintegrin ADAM8. Processing by autocatalysis is required for proteolytic activity and cell adhesion. Biol Chem 2002;277(50): 48210–9.
11. Primakoff P, Myles DG. The ADAM gene family: surface proteins with adhesion and protease activity. Trends Genet 2000;16(2):83–7.
12. Seals DF, Courtneidge SA. The ADAMs family of metalloproteases: multidomain proteins with multiple functions. Genes Dev 2003; 17(1):7–30.
13. Fourie AM, Coles F, Moreno V, Karlsson L. Catalytic activity of ADAM8, ADAM15, and MDC-L (ADAM28) on synthetic peptide substrates and in ectodomain cleavage of CD23. J Biol Chem 2003; 278(33):30469–77.
14. Naus S, Richter M, Wildeboer D, Moss M, Schachner M, Bartsch JW. Ectodomain shedding of the neural recognition molecule CHL1 by the metalloprotease-disintegrin ADAM8 promotes neurite outgrowth and suppresses neuronal cell death. J Biol Chem 2004;279(16): 16083–90.
15. Karan D, Lin FC, Bryan M, et al. Expression of ADAMs (a disintegrin and metalloproteases) and TIMP-3 (tissue inhibitor of metalloproteinase-3) in human prostatic adenocarcinomas. Int J Oncol 2003;23(5):1365–71.
16. O'Shea C, McKie N, Buggy Y, et al. Expression of ADAM-9 mRNA and protein in human breast cancer. Int J Cancer 2003;105(6): 754–61.