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

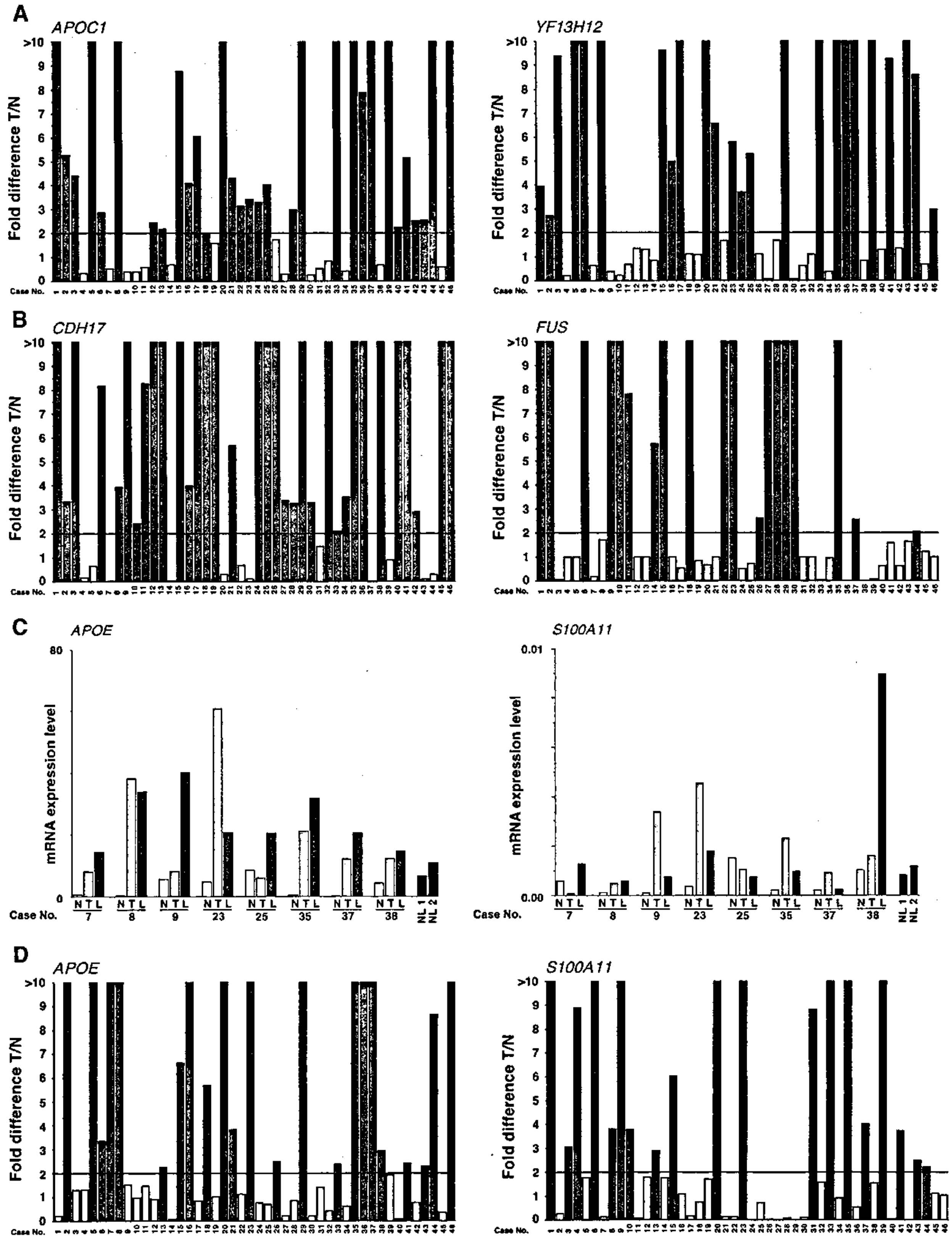
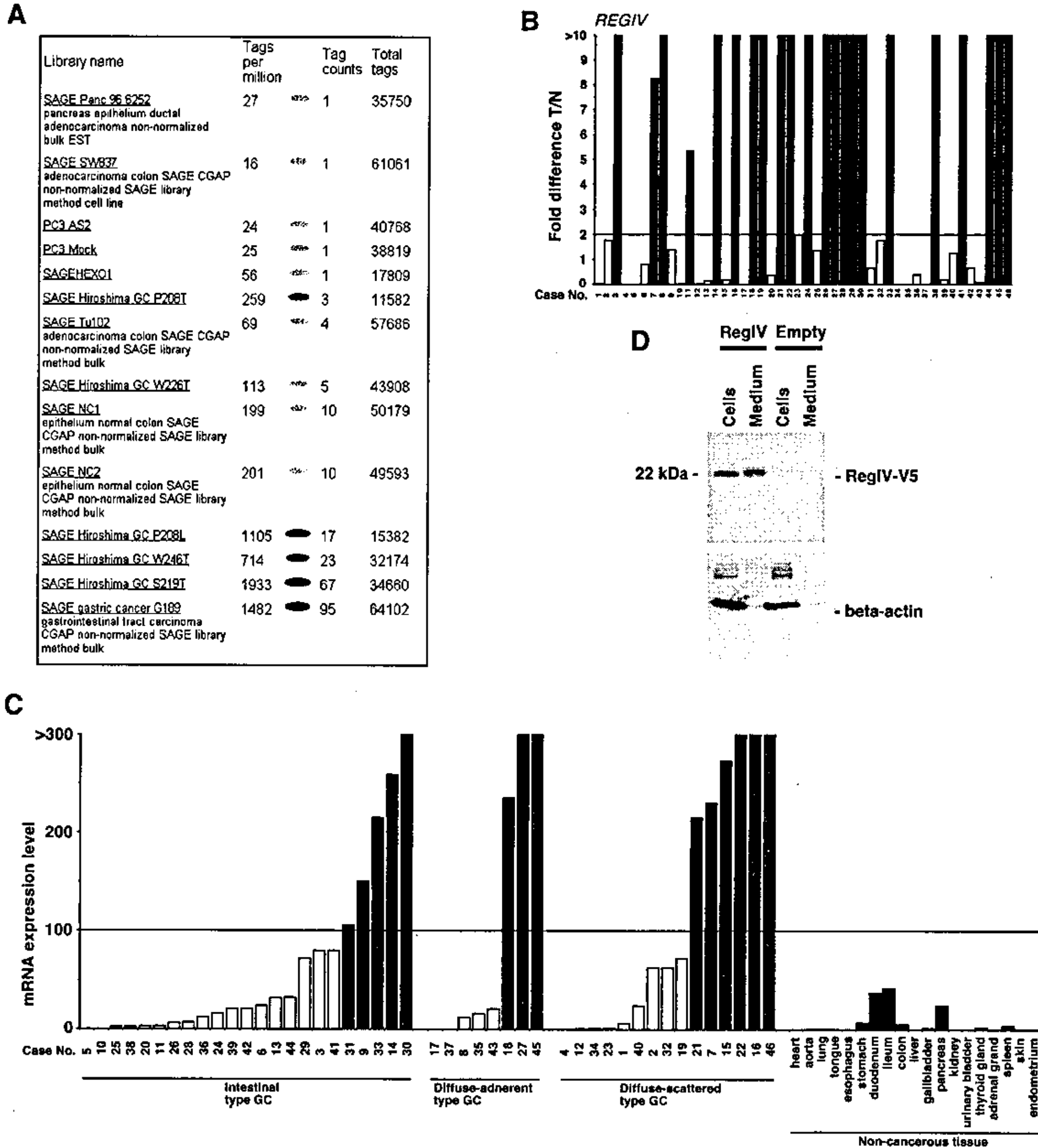


Fig. 3. Validation of serial analysis of gene expression (SAGE) data by quantitative reverse transcription-PCR (RT-PCR). Fold change indicates the ratio of target gene mRNA level in gastric carcinoma (GC) to that in corresponding nonneoplastic mucosa. *A*, quantitative RT-PCR analysis of genes commonly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



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
SAGEHEX01	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*

	Tags per million				
Tag sequence	P208T	P208L	UniGene ID	Symbol	Description
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 I 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.

J Cancer Res Clin Oncol (2004) 130: 259–265
DOI 10.1007/s00432-004-0543-1

ORIGINAL PAPER

Shunji Matsumura · Naohide Oue · Yasuhiko Kitadai
Kazuaki Chayama · Kazuhiro Yoshida
Yoshiyuki Yamaguchi · Tetsuya Toge · Kazue Imai
Kei Nakachi · Wataru Yasui

A single nucleotide polymorphism in the *MMP-1* promoter is correlated with histological differentiation of gastric cancer

Received: 1 August 2003 / Accepted: 28 December 2003 / Published online: 18 February 2004
© Springer-Verlag 2004

Abstract *Purpose*: Matrix metalloproteinase-1 (MMP-1) plays a key role in cancer invasion and metastasis by degradation of extracellular matrix (ECM) and basement membrane barriers. The 1G/2G single nucleotide polymorphism (SNP) in the *MMP-1* promoter at position −1607 bp has been reported to affect the transcriptional activity. In the light of these findings, we investigated whether this SNP in the *MMP-1* promoter is associated with the development, differentiation, and progression of gastric cancer. *Methods*: The 215 gastric cancer patients and 166 controls were used in this study. The SNP of the *MMP-1* promoter was analyzed by PCR-RFLP and sequencing. The genotype frequency was compared between cases and controls, and the association with clinicopathological parameters among cases was studied. *Results*: The frequency of 1G/2G genotypes in gastric cancer patients was similar to those in controls ($p = 0.57$). The degree of tumor invasion, the presence of lymph node metastasis, and clinical stage showed no significant association with the SNP. On the other hand, we found a significant association with histological differentiation and gender among gastric cancer patients ($p < 0.05$, respectively). *Conclusions*: The presence of 2G allele in the *MMP-1* promoter did not enhance the risk of gastric cancer; however, it may be involved in differentiation of gastric cancer.

Keywords SNP · MMP-1 · Gastric cancer

S. Matsumura · N. Oue · W. Yasui (✉)
Department of Molecular Pathology,
Hiroshima University Graduate School
of Biomedical Sciences, 1-2-3 Kasumi,
Minami-ku, 734-8551 Hiroshima, Japan
E-mail: wyasui@hiroshima-u.ac.jp
Tel.: +81-82-2575145
Fax: +81-82-2575149

Y. Kitadai · K. Chayama
Department of Medicine and Molecular Science,
Hiroshima University Graduate School
of Biomedical Sciences,
Hiroshima, Japan

K. Yoshida · Y. Yamaguchi · T. Toge
Department of Surgical Oncology,
Research Institute for Radiation
Biology and Medicine,
Hiroshima University,
Hiroshima, Japan

K. Imai · K. Nakachi
Department of Radiobiology/Molecular Epidemiology,
Radiation Effects Research Foundation,
Hiroshima, Japan

Introduction

Gastric cancer is one of the most common cancers in many Asian countries including Japan and Korea. The poor prognosis depends on the degree of stomach wall invasion and on metastatic spread to regional lymph nodes. Degradation of extracellular matrix (ECM) and basement membrane barriers by MMPs plays an important role in tumor invasion and metastasis (Forget et al. 1999; Kohn and Liotta 1995; Liotta et al. 1991). A prognostic value of MMPs expression in tumor tissue has been reported (McDonnel and Matrisian 1991). Overexpression of MMP-1 has been demonstrated in a variety of cancers (Hewitt et al. 1991; Murray et al. 1998a; Templeton et al. 1990), and the expression of MMP-1 is associated with poor prognosis of esophageal cancer patients (Murray et al. 1998b). In colorectal cancer, the expression of MMP-1 correlated with pathological factors such as Dukes' stage, differentiation, lymphatic or vascular invasion, and tumor depth (Baker and Leaper 2003).

Genetic polymorphism of insertion of a guanine (G) nucleotide at −1607 bp in the *MMP-1* gene promoter sequence, which generates the sequence 5'-GGA-3', has been identified. This sequence generates a new binding site for ETS transcription factor, influencing its transcriptional activity (Rutter et al. 1998). Moreover, the presence of 2G allele in the *MMP-1* promoter has been reported to associate with the development and

progression of carcinomas of the ovary, endometrium, and colorectum (Ghilardi et al. 2001; Kanamori et al. 1999; Nishioka et al. 2000). The frequency of ovarian cancer patients carrying 2G alleles was significantly higher than that in non-cancer individuals (Kanamori et al. 1999); hence, the presence of 2G allele is thought to be a risk factor of endometrial cancer (Nishioka et al. 2000). Similarly, the frequency of 2G allele was higher in colorectal patients than that in controls (Ghilardi et al. 2001). The levels of *MMP-1* expression in ovarian cancer tissues among the patients carrying 2G alleles were significantly elevated, compared with those homozygously carrying 1G alleles (Kanamori et al. 1999; Nishioka et al. 2000).

In gastric cancer, *MMP-1*expression has been associated with both peritoneal and lymph node metastasis (Inoue et al. 1999); however, there is no report on the association between the *MMP-1* promoter polymorphism and the development of gastric cancer. In this study we investigated whether the 1G/2G polymorphism in the *MMP-1* promoter is associated with the development of gastric cancer. Moreover, we examined the relationship between the 1G/2G polymorphism and the clinicopathological factors among gastric cancer patients.

Materials and methods

Samples

A total of 381 peripheral blood samples from 166 healthy control subjects and 215 gastric cancer patients were employed in this study. Controls were randomly selected from those visited Hiroshima University Hospital for regular healthy check or symptoms such as appetite loss or epigastralgia. They were proved to be free from malignancy by medical examination with gastrointestinal scope and biopsy. Gastric cancer patients underwent surgical operation or endoscopic mucosal resection (EMR) at Hiroshima University Hospital. All patients and controls gave informed consent prior to enrollment in the study. The human genome research ethics screening committee of Hiroshima University School of Medicine approved this study. Gastric cancer patients were 153 males and 62 females (median age 67.7±11.4 years). Histology of gastric cancer was classified according to the criteria of Lauren (1965). There were 122 patients who had an intestinal type of gastric cancer, and 93 patients who had a diffuse type. Intestinal type and diffuse type correspond to well-differentiated type and poorly differentiated type, respectively, in the histological classification of the Japanese Gastric Cancer Association (Japanese Gastric Cancer Association 1998). Alternative histological classification of the 215 gastric carcinomas was those with either intestinal or diffuse type components (pure type) or with coexistence of both types of components (mixed type; Stelzner and Emmrich 1997). Gastric cancer patients were grouped according to TNM classification (Sobin and Wittekind 2002), on the basis of the postoperative histopathological evaluation. Moreover, they were assigned to two subgroups according to the presence (N+) or absence (N–) of detectable lymph node metastasis at the time of diagnosis.

Eight human gastric carcinoma cell lines (MKN-1, MKN-7, MKN-28, MKN-45, MKN-74, TMK-1, HSC-39, and KATO-III) were used. The TMK-1 cell lines was established in our laboratory from poorly differentiated adenocarcinoma (Ochiai et al. 1985). Five gastric carcinoma cell lines of the MKN series (MKN-1, adenosquamous cell carcinoma; MKN-7, MKN-28 and MKN-74, well-differentiated adenocarcinomas; and MKN-45, poorly differentiated adenocarcinoma) were kindly provided by T. Suzuki (Fukushima Medical University, Fukushima, Japan). KATO-III and HSC-39 cell lines, which were established from signet ring cell carcinoma, were kindly provided by M Sekiguchi (University of Tokyo, Tokyo, Japan) and by K. Yanagihara (National Cancer Center, Tokyo, Japan), respectively, (Yokozaki 2000). All of these cell lines were routinely maintained in RPMI-1640 (Nissui Pharmaceutical, Tokyo, Japan) supplemented with 10% fetal bovine serum (Whittaker, Walkersville, Md.), penicillin (0.1 mg/ml), and streptomycin (0.1 mg/ml) under conditions of 5% CO_2 in air at 37°C.

DNA extraction and PCR-RFLP analysis

The genomic DNA purification kit (Promega, Madison, Wis.) and QIAamp 96 DNA Blood kit (QIAGEN, Valencia, USA) were used for DNA extraction. The PCR-restriction fragment length polymorphism (RFLP) assay was used to determine the *MMP-1* genotypes. The PCR primers used for amplifying MMP-1 polymorphism were: forward primer 5'-TGACTTTTAAAACA TAGTCTATGTTCA-3'; reverse primer 5'-TCTTGGATTGATT TGAGATAAGTCATAGC-3'. The reverse primer was specially designed to introduce a recognition site of restriction enzyme *Alu*I (AGCT) by replacing a T with a G at the second position close to the 3' end of the primer (Zhu et al. 2001). The 1G alleles have this recognition site, whereas the 2G alleles destroy the recognition site by inserting a guanine. The target sequence was amplified in a 25-μl reaction volume containing 10–20 ng of genomic DNA, 0.2 μM dNTP, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2 mM MgCl2, 0.3 μM of each primer, and 0.75 units of Ampli Taq Gold (Perkin-Elmer, Norwalk, Conn.). The PCR amplification was carried out with 40 cycles of denaturation at 94°C for 30 s, annealing at 58°C (depending on the primer) for 30 s, and extension at 72°C for 30 s after the initial activation step of 94°C for 10 min. The 269-bp fragment was then digested with *Alu*I (TaKaRa Biomedicals, Shiga, Japan) overnight at 37°C. After overnight digestion, 269-bp (2G allele), 241-bp, and 28-bp (1G allele) fragments were loaded on an ethidium bromide stained 2.5% NuSieve 3:1 agarose (FMC Bioproducts, Rockland, Md.) gel for 60 min at 100 V. Heterozygotes displayed a combination of both alleles (269, 241, and 28 bp).

Sequencing analysis of PCR products

The PCR products were purified and cloned into the pCR2.1 vector (Invitrogen, Carlsbad, Calif.). Plasmid DNA was extracted from individual clones by alkaline lysis plasmid minipreparation. The inserted PCR fragments obtained from each sample were sequenced with both M13 reverse and M13 forward primer using the PRISM AmpliTaq DNA polymerase FS Ready Reaction Dye Terminator Sequencing kit (Perkin-Elmer ABI, Foster City, Calif.). Reamplified DNA fragments were purified with Centri-Sep Columns (Princeton, Adelphia, N.J.) and were sequenced by ABI PRISM 310 genetic analyzer (Perkin-Elmer ABI, Foster City, Calif.).

Statistical analysis

Statistical analysis was performed with the use of Fisher's exact test. A value of $p<0.05$ was considered significant. Odds ratios (OR) and 95% confidence intervals (95% CI) were used for estimating the risk of association with genotypes. Odds ratios for the genotypes were calculated by the logistic regression model, adjusting for age and gender. The logistic regression analysis was performed for the association between the genotypes and clinicopathological factors (SPSS software, ver 11.0).

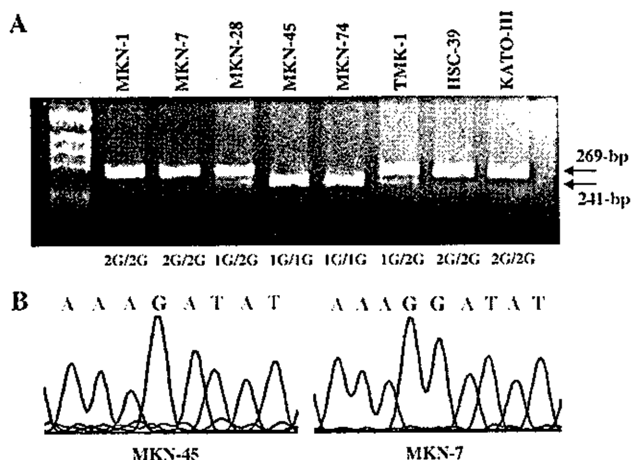


Fig. 1 **A** Representative PCR-RFLP analysis to confirm the variants in eight human gastric cancer cell lines. Ethidium bromide-stained 2.5% NuSieve 3:1 agarose gel used for genotyping. The target products (269-bp) in the *MMP-1* gene promoter was PCR amplified and digested with *AluI*, which cleaved the 1G allele at the polymorphic site, generating two fragments (241-bp and 28-bp, respectively), but did not cut the 2G allele. **B** Representative electropherogram of each type of polymorphism. The genotype of the SNP was proved to be 1G/1G in MKN-45 and 2G/2G in MKN-7

Results

We first examined the 1G/2G polymorphism in the *MMP-1* gene promoter by PCR-RFLP in eight gastric cancer cell lines, followed by sequencing: three patterns of PCR-RFLP were detected (Fig. 1A). MKN-1, MKN-7, HSC-39, and KATO-III cells had a single 269-bp DNA band; MKN-45 and MKN-74 cells showed a single 241-bp DNA band. On the other hand, MKN-28 and TMK-1 cells had heterozygous bands. Subsequent sequencing confirmed the 1G/2G genotyping of MKN-45 and MKN-7 cells (Fig. 1B). MKN-45 and MKN-7 cells were identified with 1G/1G and 2G/2G genotypes, respectively; MKN-45 and MKN-74 cells carried 1G/1G genotype; MKN-1, MKN-7, HSC-39, and KATO-III cells carried 2G/2G genotype; MKN28 and TMK-1 cells carried heterozygous 1G/2G genotype. Eight human gastric cancer cell lines used in this study had been derived from Japanese gastric cancer patients and established in Japan. There were no particularly prominent genotypes among these cell lines.

We next determined the 1G/2G genotyping among gastric cancer patients and controls; typical PCR-RFLP patterns are shown in Fig. 2 (i.e., 1G/1G or 2G/2G genotype with a single 241-bp or 269-bp band, respectively; 1G/2G genotype with 241-bp and 269-bp bands). The 1G/2G genotype distribution in gastric cancer patients and controls is shown in Table 1. The genotype distribution among controls was in good agreement with Hardy-Weinberg equilibrium. The allelic frequency in controls in our study was similar to the allele frequency reported in healthy Japanese subjects (Kanamori et al. 1999). The frequency of those carrying at least one 2G allele (1G/2G and 2G/2G) was almost equal between patients with gastric cancers (87.9%) and controls (89.7%; $p=0.57$, OR = 0.83, 95% CI = 0.43–1.59). In addition, we found that the frequency of genotypes (1G/1G vs 1G/2G + 2G/2G) did not differ by gender, age, and status of *H. pylori* infection.

Finally, we analyzed the association between the 1G/2G genotyping and clinicopathological factors among gastric cancer patients (Table 2). We found a significant difference in genotype distribution (1G/1G vs 1G/2G + 2G/2G) by histological classification (intestinal type vs diffuse type, $p=0.03$, OR = 2.84, 95% CI = 1.09–7.39). The patients carrying at least one 2G allele were more frequent in diffuse type than those in intestinal type. Furthermore, the patients carrying at least one 2G allele were more frequent in mixed type than those in pure type, with an OR of 3.81, although this was not statistically significant due to a small number of the mixed type. We also found a significant difference by gender (male vs female, $p=0.04$, OR = 3.48, 95% CI = 1.00–12.04). On the other hand, depth of tumor invasion, the presence of lymph node metastasis, and TNM classification showed no significant correlation with genotyping. In general, intestinal and diffuse types of gastric carcinomas are more frequent in males and females, respectively, implying a possible interaction between histological typing and gender. The logistic

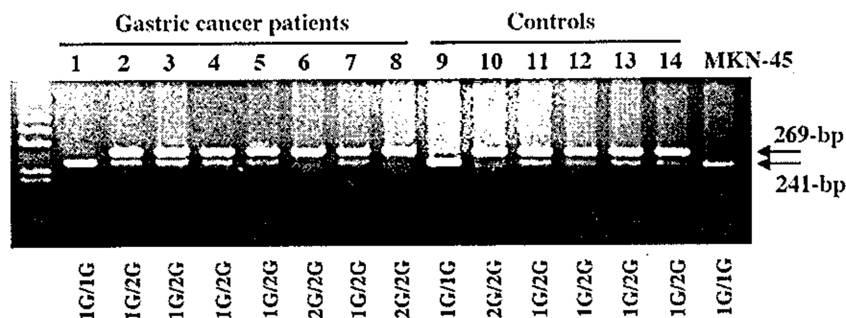


Fig. 2 Representative 8 cases of gastric cancer patients and 5 cases of controls are shown. The target products (*269-bp*) in the *MMP-1* gene promoter were PCR amplified and digested with *AluI*. MKN-45 cells were used as positive controls for digestion with *AluI*. *Numbers above the panel* are case numbers. Genotypes are shown below each case. Case numbers 1, 3, 6, and 7 are intestinal type, and the others are diffuse type

Table 1 *MMP-1*genotype distribution of the study subjects. *CI* confidence interval

Genotype	Controls (%)[a] ($n=166$)	Patients (%) ($n=215$)	*P* value[b]	Crude OR[c] (95% CI)
Overall				
1G/1G	17 (10.3)	26 (12.1)		
1G/2G	61 (36.7)	88 (40.9)		
2G/2G	88 (53.0)	101 (47.0)	0.57	0.83 (0.43–1.59)
Gender				
Male				
1G/1G	9 (9.5)	23 (15.0)		
1G/2G	41 (43.2)	60 (39.2)		
2G/2G	45 (47.3)	70 (45.8)	0.20	0.59 (0.26–1.34)
Female				
1G/1G	8 (11.3)	3 (4.8)		
1G/2G	20 (28.1)	28 (45.2)		
2G/2G	43 (60.6)	31 (50.0)	0.18	2.49 (0.63–9.86)
Age (years)				
<65				
1G/1G	14 (11.1)	11 (11.7)		
1G/2G	46 (36.5)	39 (41.5)		
2G/2G	66 (52.4)	44 (46.8)	0.89	0.94 (0.41–2.18)
≥65				
1G/1G	3 (7.5)	15 (12.4)		
1G/2G	15 (37.5)	49 (40.5)		
2G/2G	22 (55.0)	57 (47.1)	0.39	0.57 (0.16–2.09)
H. pylori infection				
Negative	$n=47$	$n=44$		
1G/1G	4 (8.5)	7 (15.9)		
1G/2G	16 (34.0)	14 (31.8)		
2G/2G	27 (57.5)	23 (52.3)	0.28	0.49 (0.13–1.81)
Positive	$n=90$	$n=49$		
1G/1G	12 (13.4)	6 (12.2)		
1G/2G	31 (34.4)	20 (40.8)		
2G/2G	47 (52.2)	23 (47.0)	0.86	1.10 (0.39–3.15)

[a]The observed genotype distribution of controls was in agreement with Hardy-Weinberg equilibrium
[b]Correlation was analyzed by Fisher's test. The *p* values <0.05 were regarded as statistically significant. The *p* values of 1G/2G + 2G/2G genotypes relative to 1G/1G genotype
[c]Odds ratio of 2G/2G + 1G/2G genotypes relative to 1G/1G genotype

Table 2 Correlation between the genotype of SNP in the *MMP-1* promoter and clinicopathological characteristics of gastric cancer patients

	Genotype (%)			OR (95%CI)[a]	
	1G/1G	1G/2G	2G/2G	Crude	Adjusted[b]
Age (years)					
<65 ($n=94$)	11 (11.7)	39 (41.5)	44 (46.8)	0.94	
≥65 ($n=121$)	15 (12.4)	49 (40.5)	57 (47.1)	(0.41–2.15)	
Gender					
Male ($n=153$)	23 (15.0)	60 (39.2)	70 (45.8)	3.48	
Female ($n=62$)	3 (4.8)	28 (45.2)	31 (50.0)	(1.00–12.04)	
Histological classification 1[c]					
Intestinal ($n=121$)	20 (16.5)	49 (40.5)	52 (43.0)	2.84	3.56
Diffuse ($n=94$)	6 (6.4)	39 (41.5)	49 (52.1)	(1.09–7.39)	(1.15–11.11)
Histological classification 2[d]					
Mixed type ($n=26$)	1 (3.8)	12 (46.4)	13 (50.0)	3.81	4.73
Pure type ($n=189$)	25 (13.2)	76 (40.2)	88 (46.6)	(0.49–29.38)	(0.58–38.53)
Depth					
m, sm ($n=85$)	10 (11.8)	29 (34.1)	46 (54.1)	0.97	1.02
mp ~ ($n=130$)	16 (12.3)	59 (45.4)	55 (42.3)	(0.42–2.25)	(0.24–4.35)
Lymph node metastasis					
N (−; $n=126$)	15 (11.9)	46 (36.5)	65 (51.6)	0.98	1.14
N (+; $n=89$)	11 (12.4)	42 (47.2)	36 (40.4)	(0.43–2.25)	(0.16–8.13)
TNM classification[e]					
Stages I, II ($n=145$)	20 (13.8)	55 (37.9)	70 (48.3)	1.71	3.04
Stages III, IV ($n=70$)	6 (8.6)	33 (47.1)	31 (44.3)	(0.65–4.46)	(0.83–11.16)

Correlation was analyzed by Fisher's test. The *P* values <0.05 were regarded as statistically significant
[a]Odds ratio of 2G/2G + 1G/2G genotypes relative to 1G/1G genotype
[b]The ORs were adjusted for age and gender
[c]Histology of gastric cancer was classified according to the criteria of the Lauren
[d]Pure-type gastric cancer with either intestinal or diffuse type components; mixed-type gastric cancer with coexistence of both types of components
[e]Tumor staging was classified according to the criterion of the UICC TNM stage grouping, 6th edition, 2002, stomach

regression analysis then revealed that histological classification remained significant ($p = 0.03$, OR = 3.56, 95% CI = 1.15–11.11) even after adjustment for age and gender.

Discussion

Studies of ovarian cancer and colorectal cancer have shown that the frequency of patients carrying at least one 2G allele in the *MMP-1* promoter was significantly higher than in control subjects (Ghilardi et al. 2001; Kanamori et al. 1999); thus, the presence of 2G allele is considered to be one of the risk factors for the development of these cancers. Furthermore, pancreatic cancers frequently showed a positive staining for *MMP-1* protein in immuno-histochemical analysis, whereas the *MMP-1* expression in fetal and normal pancreatic tissues was very faint (Ito et al. 1999). In this study, we examined whether the risk of gastric cancer is associated with the 1G/2G polymorphism in the *MMP-1* promoter region. The allelic frequency in the patients with gastric cancer was similar to that in controls. It seems that the presence of 2G allele did not enhance the susceptibility for the development of gastric cancer.

However, we found a significant association between the 1G/2G polymorphism and the histological classification. The frequency of those carrying at least one 2G allele was significantly higher in the diffuse type (poorly differentiated type) of gastric cancer than that in the intestinal type (well-differentiated type) with an OR of 3.56. The *MMP-1* promoter with 2G allele has displayed significantly increased transcriptional activity than that with 1G allele in melanoma cell lines and normal fibroblasts through the ETS binding site (Rutter et al. 1998). Diffuse type of gastric cancer is usually characterized by an abundant deposition of collagen fibers, possibly requiring higher levels of *MMP-1*expression for degradation of ECM. It has beensuggested that carcinogenesis was a multicellular and multistage process in which the destruction of the microenvironment was required for conversion of normal tissue to tumor (Park et al. 2000). Although MMPs are not oncogenic or mutagenic, they alter the microenvironment and may affect the process of carcinogenesis and its histology. This polymorphism may have a more profound impact on histology and differentiation of gastric cancer.

Recent studies showed the significance of a modified histological classification, pure type or mixed type. Patients with mixed-type gastric cancer revealed poorer prognosis than those with pure type did (Stelzner and Emmrich 1997). In this study, the patients carrying at least one 2G allele were more frequent in mixed type than those in pure type, with an OR of 3.81, although this was not statistically significant due to a small number of the mixed type. Take together; these findings suggest that the presence of 2G allele in the *MMP-1* promoter may contribute to the morphogenesis of gastric carcinomas.

Several MMPs, including MMP-1, have been reported to play an important role in cancer invasion through their overexpression, which is associated with metastasis and unfavorable prognosis in esophageal cancer, ovarian cancer, cutaneous malignant melanoma, and colorectal cancer (Ghilardi et al. 2001; Kanamori et al. 1999; Murray et al. 1998b; Ye et al. 2001).Kanamori et al. (1999) observed a high expression level of the *MMP-1* in tumors carrying 2G allele than 1G homozygotes in ovarian cancers; however, in this study, 1G/2G polymorphism in the *MMP-1* promoter was not associated with invasion, lymph node metastasis, and TNM classification in gastric cancer patients. Possible explanations may be the following:

1. The degree of tumor invasion in gastric carcinoma might be determined not only by the presence of 2G allele in the *MMP-1* promoter but also the response to growth factors and cytokines. In fact, cytokines, such as interleukin-1 (IL-1), influence the expression levels of *MMP-1*(Singer et al. 1997). Especially, IL-1A acts as a growth stimulator for gastric carcinoma (Ito et al. 1993), and it correlates with liver metastasis of gastric carcinoma (Furuya et al. 1999; Tomimatsu et al. 2001). Although the presence of 2G allele in the *MMP-1* promoter displays a significantly higher transcriptional activity than the *MMP-1* promoter with 1G allele (Rutter et al. 1998), the polymorphism may not significantly affect mean expression levels of *MMP-1* in normal stromal cells. The polymorphism may increase the *MMP-1* expression in response to growth factors and cytokines (Wyatt et al. 2002).
2. Some reports revealed that overexpression of *MMP-1* was observed in stromal cells of gastric carcinoma but not in carcinoma cells (Migita et al. 1999; Otani et al. 1999). So, the presence of 2G allele of *MMP-1* promoter may not necessarily contribute to the degree of tumor invasion in gastric carcinoma.
3. In stage-I cases in our study, a large portion of them (99 of 112 cases) have 2G allele, and 37 of 54 cases carrying 2G homozygotes were treated by means of EMR. So, they are in the status of before invading to the submucosa, they have no significant correlation with tumor invasion and metastasis. In fact, our results revealed a tendency of increasing risk of development of gastric cancer with an 8.33 exceeding risk for the patients with the 2G allele (stage II vs stage III + IV;$p = 0.05$). It has been reported that MMPs can alter the microenvironment and may influence tumor formation. So, MMPs may contribute to the initial stages of cancer development, and overexpression of MMPs may be associated with elevated risk of tumorigenesis. Moreover, the polymorphism in the *MMP-1* promoter is associated with early tumor stages in lung cancer (Zhu et al. 2001). The polymorphism in the *MMP-1* promoter may be one of the pathways of the increased neoplastic risk

in the stomach. However, we do not have any data concerning pre-cancerous lesions (i.e., patients with gastric adenoma, severe atrophic gastritis with intestinal metaplasia) at present; therefore, we could not verify a relationship between the polymorphism in the *MMP-1* promoter and a pre-cancerous lesion. However, since this is a very important point to elucidate the pathway of increasing risk of tumorigenesis and progression in gastric cancer, we should clarify this issue in the near future.

On the other hand, we found a significant association with gender among gastric cancer patients ($p = 0.037$, OR = 3.48, 95% CI = 1.00–12.04). This finding suggested a gender-specific effect of the *MMP-1* polymorphism. The presence of 2G allele in *MMP-1* promoter enhances the transcriptional activity. Moreover, it has been reported that the activity of MMP-1 might be regulated by sex hormones (Marbaix et al. 1992; Schneikert et al. 1996). The expression of *MMP-1* was negatively regulated by androgen (Schneikert et al. 1996), whereas the secretion and activation of *MMP-1* was inhibited by physiological concentrations of progesterone (Marbaix et al. 1992). Furthermore, being consistent with the above-mentioned information, IL-1A is a key inducer of *MMP-1* in the human endometrium. Ovarian steroids inhibited the release of IL-1A and repress MMP-1 production (Singer et al. 1997); therefore, the presence of 2G allele in the *MMP-1* promoter may affect the risk of gastric cancer in women. However, in our study, almost all women with gastric cancer are elderly and in menopausal status; the mean levels of ovarian steroids may be low among them. So, we did not find a significant association among gastric cancer patients and controls in women.

Conclusion

In conclusion, our studies suggest that the presence of the 2G allele in the *MMP-1* promoter might be associate with histological differentiation of gastric cancer; however, we could not consider lifestyle factors such as cigarette smoking, alcohol consumption, and dietary habits of our patients. These factors may contribute to the development of gastric cancer (Chen et al. 2000). Further investigations are necessary to clarify a role of this *MMP-1* polymorphism with increased number of study subjects and epidemiological data.

References

Baker EA, Leaper DJ (2003) The plasminogen activator and matrix metalloproteinase systems in colorectal cancer: relationship to tumour pathology. Eur J Cancer 39:981–988

Chen MJ, Chiou YY, Wu SL (2000) Lifestyle habits and gastric cancer in hospital-based case-control study in Taiwan. Am J Gastroenterol 95:3242–3249

Forget MA, Desrosiers RR, Beliveau R (1999) Physiological roles of matrix metalloproteinases: implications for tumor growth and metastasis. Can J Physiol Pharmacol 77:465–480

Furuya Y, Ichikura T, Mochizuki H (1999) Interleukin-1alpha concentration in tumors as a risk factor for liver metastasis in gastric cancer. Surg Today 29:288–289

Ghilardi G, Biondi ML, Mangoni J, Leviti S, DeMonti M, Guagnellini E, Scorza R (2001) Matrix metalloproteinase-1 promoter polymorphism 1G/2G is correlated with colorectal cancer invasiveness. Clin Cancer Res 7:2344–2346

Hewitt RE, Leach IH, Powe DG, Clark IM, Cawston TE, Turner DR (1991) Distribution of collagenase and tissue inhibitor of metalloproteinases (TIMP) in colorectal tumours. Int. J Cancer 49:666–672

Inoue T, Yashiro M, Nishimura S, Maeda K, Sawada T, Ogawa Y, Sowa M, Chung KH (1999) Matrix metalloproteinase-1 expression is a prognostic factor for patients with advanced gastric cancer. Int J Mol Med 4:73–77

Ito R, Kitadai Y, Kyo E, Yokozaki H, Yasui W, Yamashita U, Nikai H, Tahara E (1993) Interleukin 1 alpha acts as an autocrine growth stimulator for human gastric carcinoma cells. Cancer Res 53:4102–4106

Ito T, Ito M, Shiozawa J, Naito S, Kanematsu T, Sekine I (1999) Expression of the MMP-1 in human pancreatic carcinoma: relationship with prognostic factor. Mod Pathol 12:669–674

Japanese Gastric Cancer Association (1998) Japanese Classification of Gastric Carcinoma – 2nd English Edition. Gastric Cancer 1:10–24

Kanamori Y, Matsushima M, Minaguchi T, Kobayashi K, Sagae S, Kudo R, Terakawa N, Nakamura Y (1999) Correlation between expression of the matrix metalloproteinase-1 gene in ovarian cancers and an insertion/deletion polymorphism in its promoter region. Cancer Res 59:4225–4227

Kohn EC, Liotta LA (1995) Molecular insights into cancer invasion: strategies for prevention and intervention. Cancer Res 55:1856–1862

Lauren P (1965) The two histological main types of gastric carcinoma. Diffuse and so-called intestinal type carcinoma: an attempt at histological classification. Acta Pathol Microbiol Scand 64:31–49

Liotta LA, Steeg PS, Stetler Stevenson WG (1991) Cancer metastasis and angiogenesis: an imbalance of positive and negative regulation. Cell 64:327–336

Marbaix E, Donnez J, Courtoy PJ, Eeckhout Y (1992) Progesterone regulates the activity of collagenase and related gelatinases A and B in human endometrial explants. Proc Natl Acad Sci USA 89:11789–11793

McDonnel S, Matrisian LM (1991) Stromelysin in tumor progression and invasion. Cancer Metastasis Rev 9:305–319

Migita T, Sato E, Saito K, Mizoi T, Shiiba K, Matsuno S, Nagura H, Ohtani H (1999) Differing expression of MMPs-1 and -9 and urokinase receptor between diffuse- and intestinal-type gastric carcinoma. Int J Cancer 84:74–79

Murray GI, Duncan ME, Arbuckle E, Melvin WT, Fothergill JE (1998a) Matrix metalloproteinases and their inhibitors in gastric cancer. Gut 43:791–797

Murray GI, Duncan ME, O'Neil P, Mckay JA, Melvin WT, Fothergill JE (1998b) Matrix metalloprotease-1 is associated with poor prognosis in oesophageal cancer. J Pathol 185:256–261

Nishioka Y, Kobayashi K, Sagae S, Ishioka S, Nishikawa A, Matsushima M, Kanamori Y, Minaguchi T, Nakamura Y, Tokino T, Kudo R (2000) A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter in endometrial carcinomas. Jpn J Cancer Res 91:612–615

Ochiai A, Yasui W, Tahara E (1985) Growth-promoting effect of gastrin on human gastric carcinoma cell line TMK-1. Jpn J Cancer Res 76:1064–1071

Otani Y, Kubota T, Sakurai Y, Igarashi N, Yokoyama T, Kimata M, Wada N, Kameyama K, Kumai K, Okada Y, Kitajima M (1999) Expression of matrix metalloproteinases in

gastric carcinoma and possibility of clinical application of matrix metalloproteinase inhibitor in vivo. Ann N Y Acad Sci 30:541–543

Park CC, Bissell MJ, Barcellos-Hoff MH (2000) The influence of the microenvironment on the malignant phenotype. Mol Med Today 6:324–329

Rutter JL, Mitchell TI, Buttice G, Meyers J, Gusella JF, Ozelius LJ, Brinckerhoff CE (1998) A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter creates an Ets binding site and augments transcription. Cancer Res 58:321–5325

Schneikert J, Peterziel H, Defossez PA, Klocker H, Launoit Y, Cato ACB (1996) Androgen receptor-Ets protein interaction is a novel mechanism for steroid hormone-mediated down-modulation of matrix metalloproteinase expression. J Biol Chem 271:23907–23913

Singer CF, Marbaix E, Kokorine I, Lemoine P, Donnez J, Eeckhout Y, Courtoy PJ (1997) Paracrine stimulation of interstitial collagenase (MMP-1) in the human endometrium by interleukin 1alpha and its dual block by ovarian steroids. Proc Natl Acad Sci USA 94:10341–10345

Sobin LH, Wittekind CH (eds) (2002) TNM classification of malignant tumors, 6th edn. Wiley-Liss, New York, pp 65–68

Stelzner S, Emmrich P (1997) The mixed type in Lauren's classification of gastric carcinoma. Histologic description and biologic behavior. Gen Diagn Pathol143:39–48

Templeton NS, Brown PD, Levy AT, Margulies IM, Liotta LA, Stetler-Stevenson WG (1990) Cloning and characterization of human tumor cell interstitial collagenase. Cancer Res 50:5431–5437

Tomimatsu S, Ichikura T, Mochizuki H (2001) Significant correlation between expression of interleukin-1alpha and liver metastasis in gastric carcinoma. Cancer 91:1272–1276

Wyatt CA, Coon CI, Gibson JJ, Brinckerhoff CE (2002) Potential for the 2G single nucleotide polymorphism in the promoter of matrix metalloproteinase to enhance gene expression in normal stromal cells. Cancer Res 62:7200–7202

Ye S, Dhillon S, Turner SJ, Bateman AC, Theaker JM, Pickering RM, Day I, Howell WM (2001) Invasiveness of cutaneous malignant melanoma is influenced by matrix metalloproteinase 1 gene polymorphism. Cancer Res 61:1296–1298

Yokozaki H (2000) Molecular characteristics of eight gastric cancer cell lines established in Japan. Pathol Int 50:767–777

Zhu Y, Spitz MR, Lei L, Mills GB, Wu X (2001) A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter enhances lung cancer susceptibility. Cancer Res 61:7825–7829

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

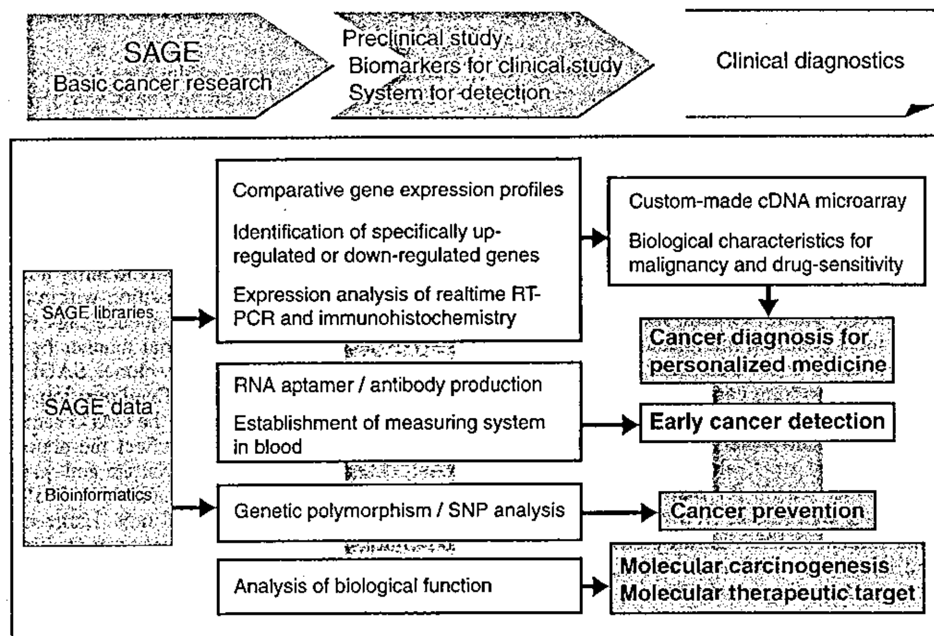


Fig. 1. Strategy to search for novel genes associated with gastric cancer through SAGE, and its clinical implication.

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.

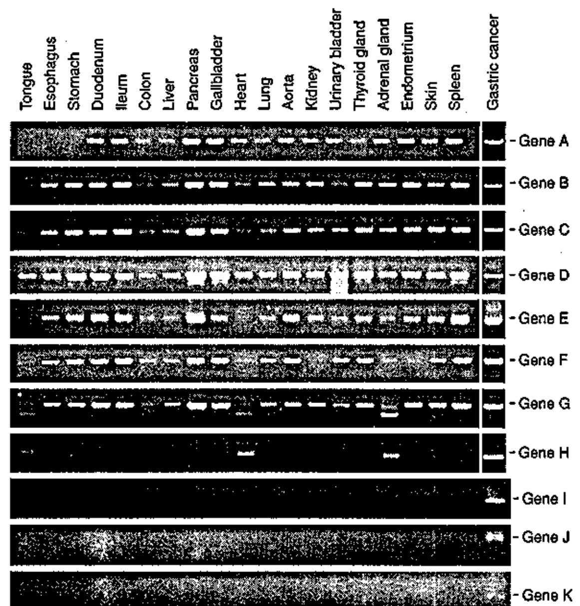


Fig. 2. Expression of genes detected in SAGE libraries of gastric cancer, but not in the libraries of 15 normal tissues. Semi-quantitative RT-PCR on various normal human tissues revealed that genes H, I, J and K are not expressed significantly in normal tissues, whereas they are detected in gastric cancer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

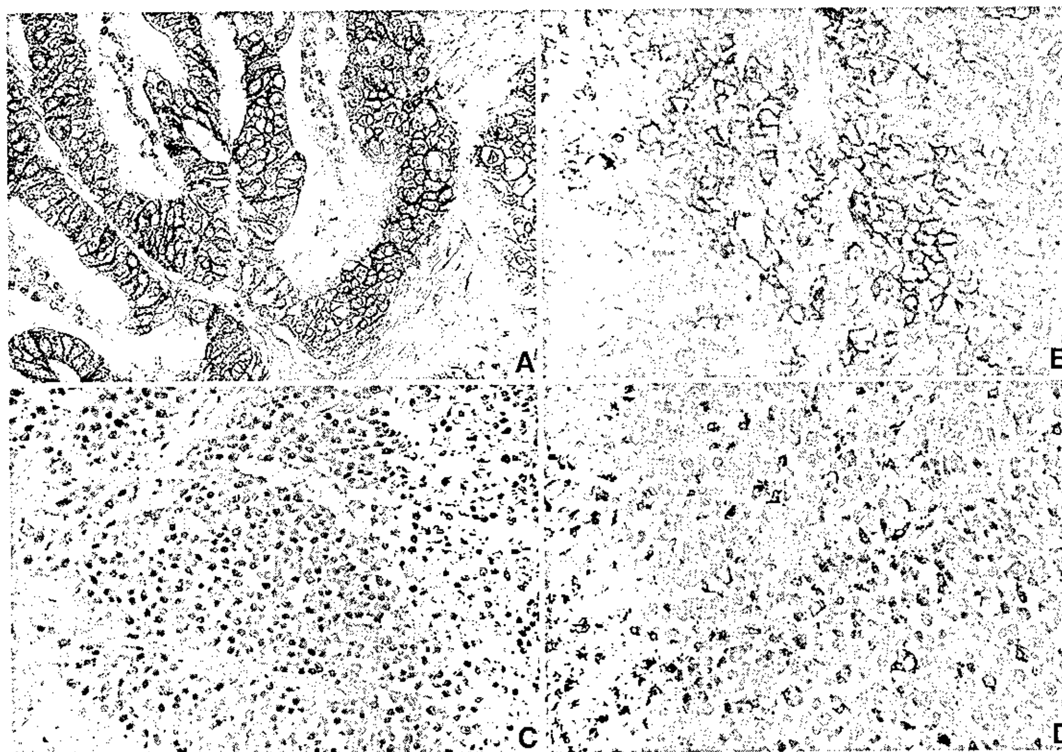


Fig. 3. Expression of cadherin 17 and apolipoprotein E in gastric carcinomas (immunohistochemical analysis). Cadherin 17 is expressed at the cell surface and cell-cell border of well-differentiated-type gastric adenocarcinoma with intestinal phenotype (A) and poorly differentiated solid-type gastric 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

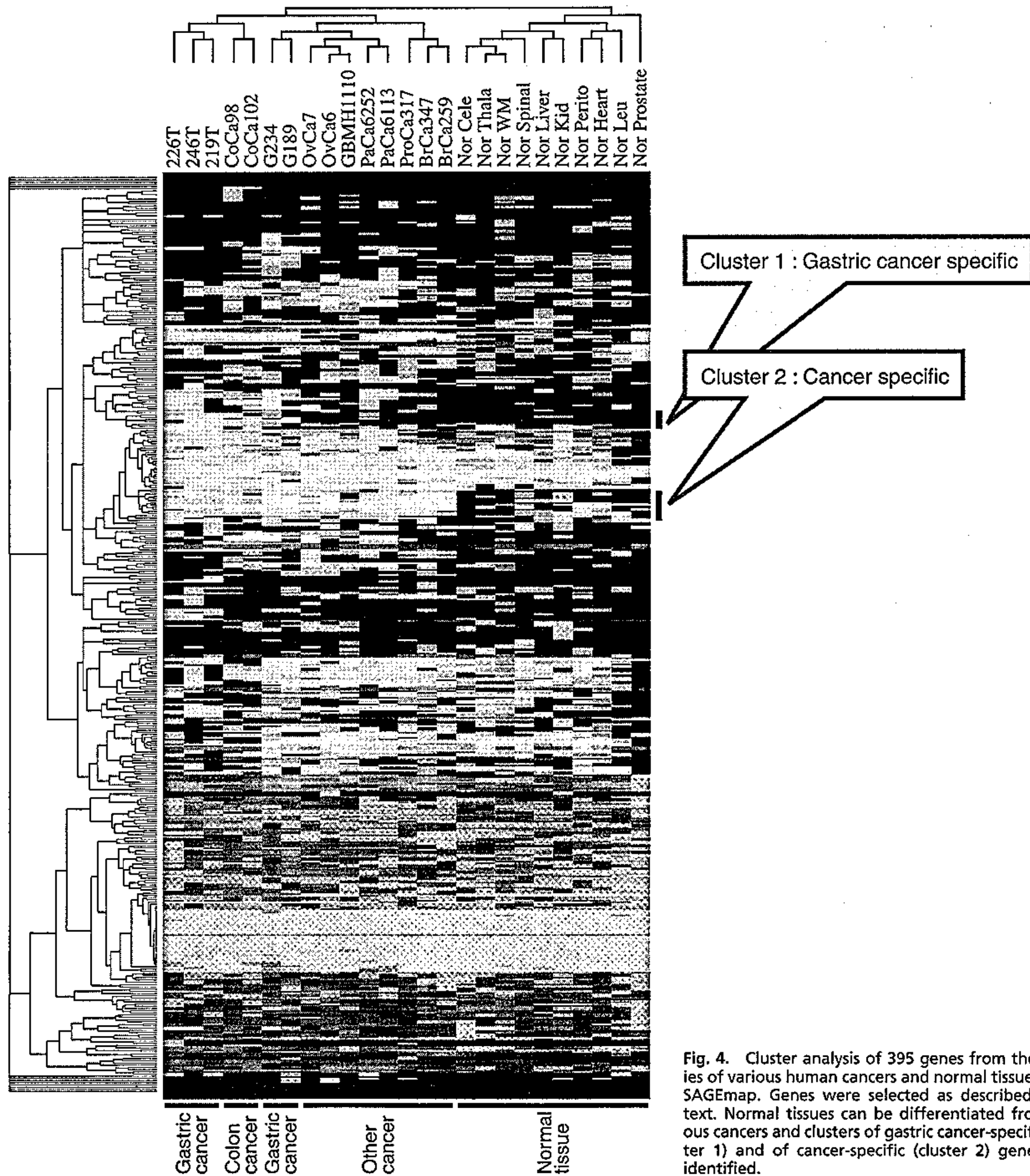


Fig. 4. Cluster analysis of 395 genes from the libraries of various human cancers and normal tissues in the SAGEmap. Genes were selected as described in the text. Normal tissues can be differentiated from various cancers and clusters of gastric cancer-specific (cluster 1) and of cancer-specific (cluster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

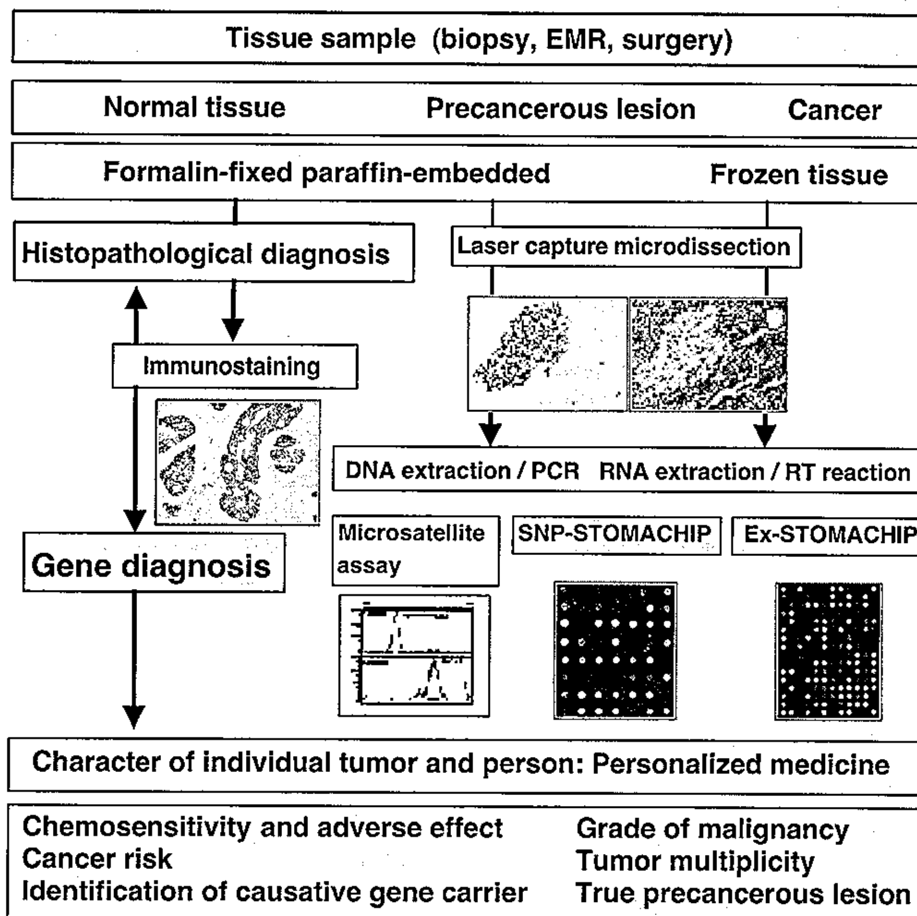


Fig. 5. New strategy of gene diagnosis of gastric cancer using pathology samples.

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.