(2004) Expression of matrix metalloproteinase 9 is a prognostic factor in patients with non-Hodgkin lymphoma. Cancer 100:356–365

Sehgal I, Baley PA, Thompson TC (1996) Transforming growth factor beta1 stimulates contrasting responses in metastatic versus primary mouse prostate cancer-derived cell lines in vitro. Cancer Res 56:3359–3365

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

Tanioka Y, Yoshida T, Yagawa T, Saiki Y, Takeo S, Harada T, Okazawa T, Yanai H, Okita K (2003) Matrix metalloproteinase-7 and matrix metalloproteinase-9 are associated with unfavourable prognosis in superficial oesophageal cancer. Br J Cancer 89:2116–2121

Wroblewski LE, Pritchard DM, Carter S, Varro A (2002) Gastrin-stimulated gastric epithelial cell invasion: the role and mechanism of increased matrix metalloproteinase 9 expression. Biochem J 365:873–879

Yasui W, Yokozaki H, Fujimoto J, Naka K, Kuniyasu H, Tahara E (2000) Genetic and epigenetic alterations in multistep carcinogenesis of the stomach. J Gastroenterol 35:111–115

Zhang B, Ye S, Hermann SM, Eriksson P, de Maat M, Evans A, Arveiler D, Luc G, Cambien F, Hamsten A, Watkins H, Henney AM (1999) Functional polymorphism in the regulatory region of gelatinase B gene in relation to severity of coronary atherosclerosis. Circulation 99:1788–794

Zhang S, Li L, Lin JY, Lin H (2003) Imbalance between expression of matrix metalloproteinase-9 and tissue inhibitor of metalloproteinase-1 in invasiveness and metastasis of human gastric carcinoma. World J Gastroenterol 9:899–904

Oncogene (2005) 24, 157–164

www.nature.com/onc

Loss of heterozygosity and histone hypoacetylation of the *PINX1* gene are associated with reduced expression in gastric carcinoma

Tomohiro Kondo[1], Naohide Oue[1], Yoshitsugu Mitani[1], Hiroki Kuniyasu[2], Tsuyoshi Noguchi[3], Kazuya Kuraoka[1], Hirofumi Nakayama[1] and Wataru Yasui*[,1]

[1]*Department of Molecular Pathology, Hiroshima University Graduate School of Biomedical Sciences, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan;* [2]*Department of Molecular Pathology, Nara Medical University, Kashihara, Japan;* [3]*Department of Oncological Science (Surgery II), Oita University Faculty of Medicine, Oita, Japan*

The expression of *PINX1*, a possible telomerase inhibitor and a putative tumor suppressor, has not been studied in human cancers, including gastric cancer (GC). We examined expression of *PINX1* by quantitative reverse transcription (RT)–PCR in 73 cases of GC, and 45 of these cases were further studied for loss of heterozygosity (LOH) by PCR with microsatellite marker D8S277. Reduced expression (tumor vs normal ratio <0.5) of *PINX1* was detected in 50 (68.5%) of 73 cases of GC. GC tissues with reduced expression of *PINX1* showed significantly higher telomerase activities as measured by telomeric repeat amplification protocol than those with normal expression of *PINX1* ($P=0.031$). LOH of *PINX1* locus was detected in 15 (33.3%) of 45 cases of GC and was correlated significantly with reduced expression of *PINX1* ($P=0.031$). Expression of *PINX1* in a GC cell line, MKN-74, was induced by treatment with trichostatin A (TSA) or nicotinamide (NAM). Chromatin immunoprecipitation assay of MKN-74 cells revealed that acetylation of histone H4 in the 5′ untranslated region (UTR) of *PINX1* was enhanced by treatment with TSA or NAM, whereas acetylation of histone H3 was not changed by TSA or NAM. In addition, TSA or NAM treatment led to inhibition of telomerase activity in MKN-74 cells. These results indicate that LOH of *PINX1* locus and hypoacetylation of histone H4 in the 5′ UTR of *PINX1* are associated with reduced expression of *PINX1* in GC. *Oncogene* (2005) **24**, 157–164. doi:10.1038/sj.onc.1207832

Keywords: *PINX1*; telomerase; LOH; histone H4; acetylation; hSir2; gastric carcinoma

Introduction

PinX1 is a Pin2/TRF1-binding protein that is a potent inhibitor of telomerase (Zhou and Lu, 2001). PinX1 binds human telomerase reverse transcriptase (hTERT) and inhibits its activity directly. A novel human liver-related putative tumor suppressor gene (*LPTS*) has been cloned previously (Liao *et al.*, 2000). *LPTS* gene is transcribed into two transcripts. The longer transcript, which is referred to as LPTS-L, encodes a 328-amino-acid protein (Liao *et al.*, 2002) that is highly homologous to PinX1 (Zhou and Lu, 2001). LPTS-L and PinX1 have different 3′-untranslated regions but encode the same protein, which is referred to as LPTS-L/PinX1 (Liao *et al.*, 2003). LPTS-L/PinX1 has strong telomerase inhibitory activity both *in vivo* and *in vitro* (Zhou and Lu, 2001; Liao *et al.*, 2002). A high percentage of tumor cells with characteristics of immortalization show high telomerase activity (Shay and Bacchetti, 1997; Cong *et al.*, 2002). We previously reported that telomerase activity is present in a majority of gastric cancer (GC) types (Tahara *et al.*, 1995; Yasui *et al.*, 1998, 1999). Loss of heterozygosity (LOH) of *LPTS-L/PINX1* locus, which maps to human chromosome 8p23, was identified in 34.5% of hepatocellular carcinoma cases (Park *et al.*, 2002). 8p23, but not specifically *PINX1*, is frequently deleted in various cancers, including carcinomas of the liver (Emi *et al.*, 1992), lung (Ohata *et al.*, 1993), colorectum (Fujiwara *et al.*, 1993), prostate (Macoska *et al.*, 1995), and breast (Yaremko *et al.*, 1995), in addition to GC (Yustein *et al.*, 1999; Baffa *et al.*, 2000). However, the significance of the *LPTS-L/PINX1* gene in human cancers including GC remains unclear.

A variety of genetic and epigenetic alterations are associated with GC (Yasui *et al.*, 2000; Oue *et al.*, 2004). Several lines of evidence suggest that histone acetylation plays an important role in transcriptional regulation (Grunstein, 1997). Histone hyperacetylation is thought to relax the chromatin structure and allow transcription factors to access promoters (Luger *et al.*, 1997; Luger and Richmond, 1998). We have reported that trichostatin A (TSA), a histone deacetylase (HDAC) inhibitor, induces expression of $p21^{Waf1}$ and *HLTF* in GC cell lines (Suzuki *et al.*, 2000; Hamai *et al.*, 2003). HDACs are separated into three distinct classes on the basis of their homologies to yeast transcriptional repressors (North *et al.*, 2003). Class I and II deacetylases are homologs of the yeast Rpd3p and Hda1p proteins, respectively. Class III deacetylases are classified on the basis of homology to the yeast transcriptional repressor, Sir2, which is nicotinamide adenine dinucleotide (NAD)-dependent HDAC (Imai *et al.*, 2000; Landry *et al.*, 2000; Smith

*Correspondence: W Yasui; E-mail: wyasui@hiroshima-u.ac.jp
Received 3 December 2003; revised 22 April 2004; accepted 22 April 2004

et al., 2000). TSA inhibits class I and II deacetylases, whereas nicotinamide (NAM) inhibits hSir2 (Luo *et al.*, 2001; Bitterman *et al.*, 2002).

In the present study, we examined expression and LOH of *PINX1* locus in GC. We also examined the association between expression of *PINX1* and acetylation of histones in the 5′ untranslated region (UTR) of *PINX1* in GC cell lines.

Results

Expression of PINX1 *in GC*

Levels of *PINX1* expression in 73 cases of GC are shown in Figure 1a. *PINX1* levels were significantly lower in GC tissues (0.347 ± 0.138) than in non-neoplastic mucosa (1.766 ± 0.758, $P<0.0001$, Wilcoxon signed rank test). Reduced expression of *PINX1* (tumor *vs* normal ratio <0.5) was detected in 50 (68.5%) of 73 cases of GC. Reduced expression of *PINX1* in GC tissues was not significantly associated with T grade (depth of tumor invasion), N grade (degree of lymph node metastasis), tumor stage, or histological type (Table 1).

Correlation between expression of PINX1 and telomerase activity in GC

We examined the correlation between *PINX1* expression and telomerase activity in 20 of 73 cases of GC. GC tissues with reduced expression of *PINX1* showed significantly higher telomerase activities than those with normal expression of *PINX1* (4.608 ± 2.596 *vs* 1.438 ± 1.018, $P=0.031$, Mann–Whitney *U*-test, Figure 1b).

LOH analysis of PINX1 locus in GC

LOH analysis was performed for 45 of 73 cases of GC, and representative results are shown in Figure 2a. LOH of *PINX1* locus was detected in 15 (33.3%) of 45 cases of GC. Of 15 cases with LOH of *PINX1* locus, 14 (93.3%) showed reduced expression of *PINX1* (Figure 2b and Table 2). Reduced expression of *PINX1* was significantly associated with LOH of *PINX1* locus ($P=0.031$; Fisher's exact test). LOH of *PINX1* locus was not significantly associated with T grade, N grade, tumor stage, or histological type.

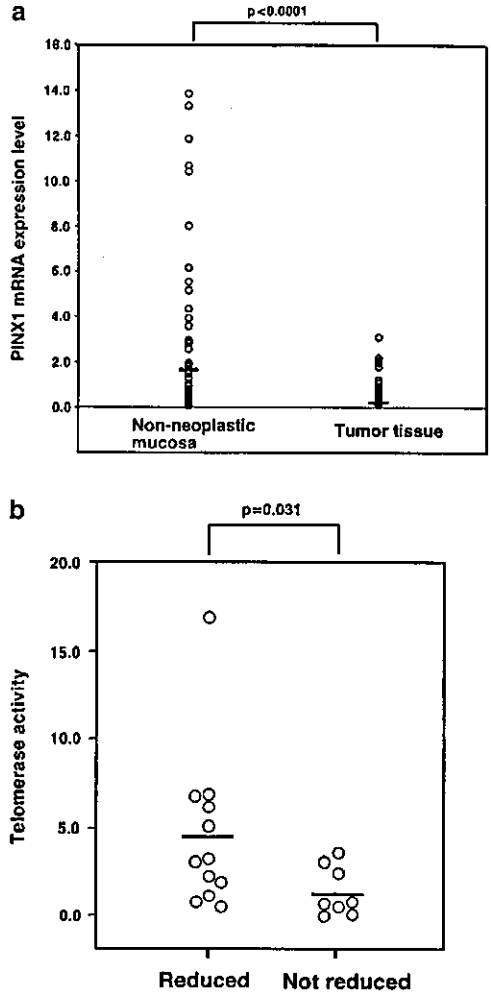


Figure 1 Expression of *PINX1* in GC. (a) Quantitative RT–PCR analysis of GC tissues and corresponding non-neoplastic mucosae. Expression of *PINX1* was significantly lower in GC tissues than in corresponding non-neoplastic mucosae ($P<0.0001$, Wilcoxon signed rank test). The expression value was calculated as the mean of three independent quantitative RT–PCR experiments. The units are arbitrary, and we calculated the *PINX1* mRNA expression level by standardization against 1 μg of total RNA from HSC-39 GC cells, which was taken as 1.0. (b) Correlation between telomerase activity and reduced expression of *PINX1*. Reduced ($T/N<0.5$), not reduced ($T/N \geq 0.5$). Telomerase activities were higher in GC tissues with reduced expression of *PINX1* than in those with normal *PINX1* levels ($P=0.031$, Mann–Whitney *U*-test). We examined the correlation between *PINX1* expression levels and telomerase activity in 20 of 73 cases for which telomerase activities were reported (Yasui *et al.*, 1998)

PINX1 expression status in GC cell lines

Next, we performed experiments with eight GC cell lines to further analyse *PINX1* expression status. Quantitative RT–PCR analysis revealed that the average level of *PINX1* expression in eight GC cell lines (0.691 ± 0.226) was less than a half of that in non-neoplastic mucosa of 73 cases (1.766 ± 0.758) (Figure 3a). In particular, *PINX1* expression level in MKN-74 cells was obviously lower than that in the seven other GC cell lines. Methylation of CpG island is an alternative way of causing the gene silencing of diverse tumor suppressor genes. To evaluate the extent of the *PINX1* methylation, bisulfite sequencing was carried out in three cell lines (MKN-28, MKN-74, and HSC-39). The *PINX1* CpG islands of these cell lines were found to be little methylated over the entire region analysed (Figure 3b and Table 3). To further examine if the reduced expression of *PINX1* is related to genomic alteration of the gene, Southern blot was performed in eight GC cell lines. No obvious alteration of *PINX1* was found in all cell lines (Figure 3c).

Table 1 Clinicopathological features of gastric cancers ($n=73$) and reduced expression of *PINX1*

	Reduced (T/N[a] *<0.5)*	*Not reduced* (T/N ≥ *0.5)*	P-*value*[b]
Histology[c]			
Intestinal	24 (63.2%)	14	
Diffuse	26 (74.3%)	9	0.221
T grade[d]			
T1, 2	22 (62.9%)	13	
T3, 4	28 (73.7%)	10	0.229
N grade[d]			
N0	14 (70.0%)	6	
N1, 2, 3	36 (67.9%)	17	0.551
Stage[d]			
I, II	20 (69.0%)	9	
III, IV	30 (68.2%)	14	0.576

[a]*T/N* ratio, *PINX1* mRNA expression levels in GC tissue relative to levels in corresponding non-neoplastic mucosa. [b]Fisher's exact test. [c]According to the Lauren criteria (Lauren, 1965). [d]According to the criteria of the TNM Stage classification system (Sobin and Wittekind, 1997)

PINX1 expression was induced by treatment with TSA and NAM

Histone acetylation plays an important role in gene expression. We presumed the possibility that histone acetylation is involved in reduced *PINX1* expression. Treatment of MKN-74 cells with TSA, an HDAC (class I and II) inhibitor, and NAM, an hSir2 (class III) inhibitor, increased *PINX1* expression in the cells (Figure 4a). The treatment with TSA and NAM together yielded an additive increase in *PINX1* expression in MKN-74 cells. In contrast, in MKN-28 cells, treatment with TSA and NAM had no effect on *PINX1* expression.

Histone acetylation status of the PINX1 gene in GC cell lines

To examine acetylation of histones H3 and H4 in the 5′ UTR of *PINX1* in MKN-74 cells, we performed chromatin immunoprecipitation (ChIP) assays. The level of acetylation of histone H3 in the 5′ UTR of *PINX1* in MKN-74 cells was similar to that in MKN-28 cells (Figure 4b). Moreover, no change in the level of acetylation of histone H3 in the 5′ UTR of *PINX1* was observed in MKN-74 cells treated with TSA and NAM. However, acetylation of histone H4 in the 5′ UTR of *PINX1* in MKN-74 cells was significantly lower than that in MKN-28 cells (Figure 4c). After treatment with both TSA and NAM, acetylation of histone H4 in the 5′ UTR of *PINX1* was increased in MKN-74 cells. No significant changes in histone H4 acetylation were observed in MKN-28 cells (data not shown).

Alteration of telomerase activity in MKN-74 treated with TSA and NAM

PinX1 binds to hTERT and inhibits its activity directly (Zhou and Lu, 2001). To observe the alteration of telomerase activity when *PINX1* expression is changed, we performed telomeric repeat amplification protocol (TRAP) assay in MKN-74 cells treated with TSA and NAM. Telomerase activity was reduced to 65.6% with TSA, and to 10.2% with NAM (Figure 5a). Levels of *hTERT* expression were not significantly changed by treatment with TSA or NAM (Figure 5b).

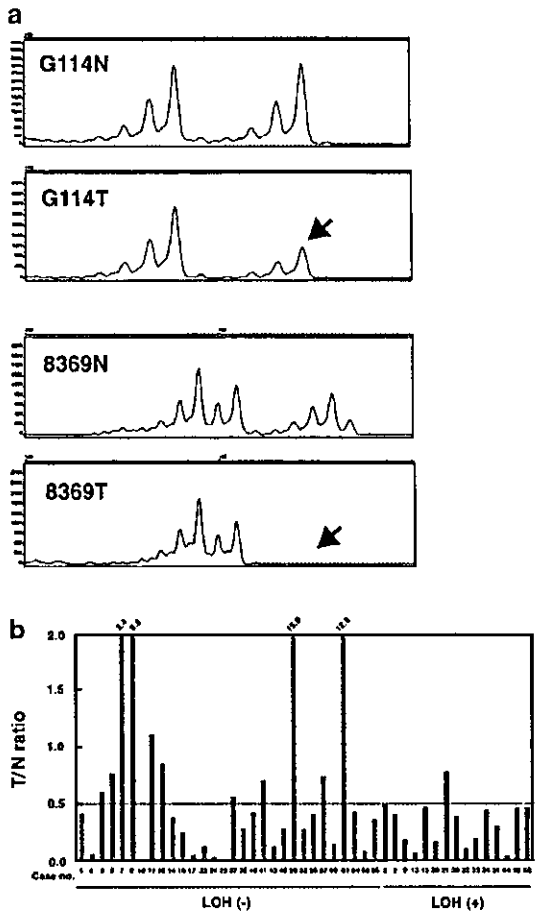


Figure 2 LOH of *PINX1* locus in GC specimens and association between LOH and *PINX1* expression. (**a**) Representative fluorescent electropherograms for LOH. Tumor tissues (T), corresponding non-neoplastic mucosae (N). Lost alleles are indicated by arrows. (**b**) Distribution of *PINX1* expression in 45 cases of GC. *T/N* ratio, *PINX1* mRNA expression levels in GC tissue relative to levels in corresponding non-neoplastic mucosa. *T/N*<0.5-fold is defined as reduced expression of *PINX1*

Discussion

The *PINX1* gene appears to function as a tumor suppressor; however, the association between *PINX1* and GC has not been studied. We report here reduced expression of *PINX1* in 50 (68.5%) of 73 GC cases. In all, 15 (33.3%) of 45 cases had LOH of *PINX1* locus, which was correlated significantly with reduced expression of *PINX1*, suggesting that LOH plays a major role in reduced expression of *PINX1*. We did not find any association between reduced expression of *PINX1* and

Table 2 Clinicopathological features of gastric cancers ($n=45$) and LOH of *PINX1* locus

	LOH		P-value*
	(+)	(−)	
PINX1 expression			
Reduced ($T/N<0.5$)	14 (42.4%)	19	
Not reduced ($T/N\geqslant 0.5$)	1 (8.3%)	11	0.031
Histology			
Intestinal	7 (33.3%)	14	
Diffuse	8 (33.3%)	16	0.625
T grade			
T1,2	8 (34.8%)	15	
T3,4	7 (31.8%)	15	0.542
N grade			
N0	4 (33.3%)	8	
N1,2,3	11 (33.3%)	22	0.645
Stage			
I,II	5 (27.8%)	13	
III,IV	10 (37.0%)	17	0.376

*Fisher's exact test

frequency of LOH and T grade, N grade, or tumor stage. This finding suggests that downregulation of *PINX1* may be involved in tumor development but not in tumor progression. We also found that telomerase activity is higher in GC tissues with reduced expression of *PINX1* than in those with normal levels of *PINX1*. Reduced expression of *PINX1* appears to be involved in activation of telomerase. Since *PINX1* is located in the subtelomeric region of human chromosome 8, LOH of the *PINX1* locus may occur easily through telomere dysfunction and chromosome instability during initiation of human cancers (Chin *et al*., 1999; Artandi *et al*., 2000; DePinho, 2000; O'Sullivan *et al*., 2002; Meeker *et al*., 2002; Van Heek *et al*., 2002). Recently, we reported that inhibition of Pot1, a single-stranded telomeric DNA-binding protein, induces telomere dysfunction and that expression of *POT1* is reduced in early-stage GC (Kondo *et al*., 2004). Although hTERT, the catalytic subunit of telomerase, is essential for activation of telomerase, it is possible that downregulation of *PINX1*, due to LOH, contributes to activation of telomerase at an early stage of stomach carcinogenesis. In fact, we previously observed that telomerase is activated in precancerous conditions such as intestinal metaplasia and adenoma of the stomach (Tahara *et al*., 1995; Kuniyasu *et al*., 1997; Yasui *et al*., 1999). Activation of telomerase is essential for cell immortality, and it may be a critical step in the development of cancers (Tahara *et al*., 1995). PinX1 may inhibit telomerase activation in normal somatic cells.

We observed reduced expression of *PINX1* in 19 (63.3%) of 30 cases of GC that did not have LOH of *PINX1* locus. We hypothesized that DNA methylation

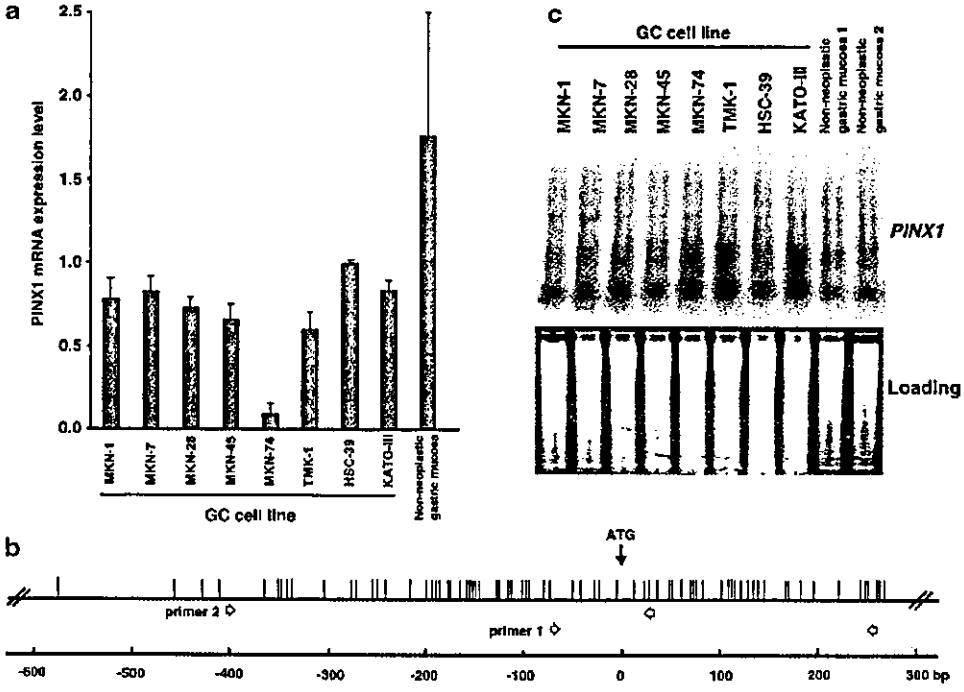


Figure 3 Expression of *PINX1* in GC cell lines. (a) Quantitative RT–PCR analysis of GC cell lines. Expression of *PINX1* in MKN-74 cells was lower than that in seven other GC cell lines. Each value is the mean of three independent quantitative RT–PCR experiments. Error bars indicate standard error (s.e.) from the mean. The units are arbitrary, and we calculated the level of *PINX1* expression by standardization against 1 μg of total RNA from HSC-39 GC cells, which was taken as 1.0. (b) A map of the CpG island of the *PINX1* gene. The CpG map of the sequence around exon 1 of the *PINX1* gene is shown. The CpG density is indicated by short vertical bars. Arrows represent PCR primers. The numbering in this scheme corresponds to position relative to known translation start sites. (c) Southern blot analysis of the *PINX1* gene in eight GC cell lines. All cell lines showed no change of *PINX1* status in comparison with normal gastric mucosa

Table 3 DNA methylation status of 38 CpG islands in the *PINX1*

Sample	*Expression*[a]	*CpG islands, percentage of clones exhibiting methylation*[b]																																			
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*	*33*	*34*	*35*	*36*	*37*	*38*
MKN-28	M	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	10	0	0	0	0	0	0
MKN-74	L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HSC-39	M	20	30	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0

[a]*PINX1* expression determined with quantitative RT–PCR. We calculated the level of *PINX1* expression by standardization against 1 μg of total RNA from HSC-39 cells, which was taken as 1.0. Units of expression levels were divided into three groups: Low, 0–0.5 U calculated (L); Medium, 0.5–2.0 U calculated (M); High, >2.0 U calculated (H). [b]DNA extracted from each sample was treated with sodium bisulfite and described in 'Materials and methods'. The percentage of 10 independent clones exhibiting DNA methylation at each CpG island is listed

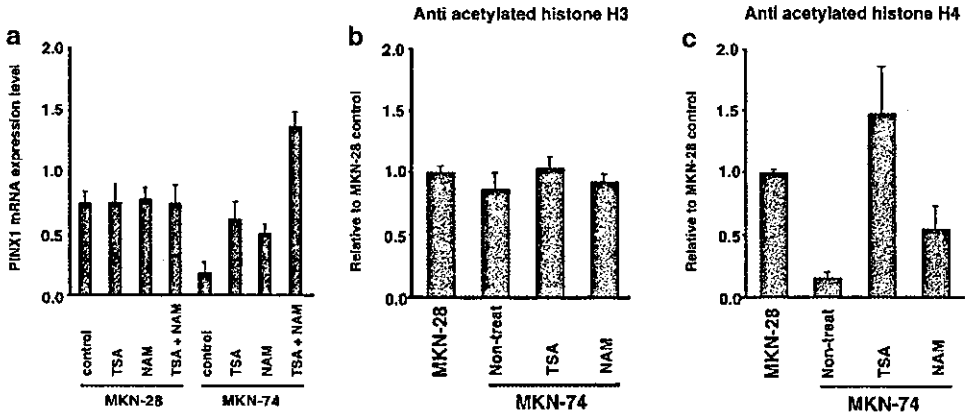


Figure 4 Effect on *PINX1* expression of TSA and NAM in GC cell lines. (a) Quantitative RT–PCR analysis of TSA- and/or NAM-treated MKN-28 and MKN-74 cells. We treated GC cell lines with TSA for 24 h and with NAM for 6 h. Cotreatment with TSA and NAM was performed for 12 h. In MKN-74 cells, treatment with TSA and/or NAM increased *PINX1* expression, whereas in MKN-28 cells, treatment with TSA and/or NAM did not significantly alter *PINX1* expression. (b, c) ChIP assay of MKN-28 and MKN-74 cells. Levels of acetylation of histone H3 in the 5′ UTR of *PINX1* in MKN-74 cells were similar to those in MKN-28 cells. Acetylation levels did not change in response to treatment with TSA or NAM. Levels of acetylation of histone H4 in the 5′ UTR of *PINX1* in MKN-74 cells were significantly lower than those in MKN-28 cells. After treatment with both TSA and NAM, acetylation of histone H4 in the 5′ UTR of *PINX1* was increased in MKN-74 cells. Bar graphs illustrate the levels of acetylation of histones H3 and H4 in MKN-74 cells relative to those of histones H3 and H4 in MKN-28 cells. Each value is the mean of three independent ChIP experiments. Error bars indicate SE from the mean. The units are arbitrary, and we calculated the acetylation level by standardization with the value from 0.1 μg of genomic DNA of HSC-39 GC cells, which was taken as 1.0

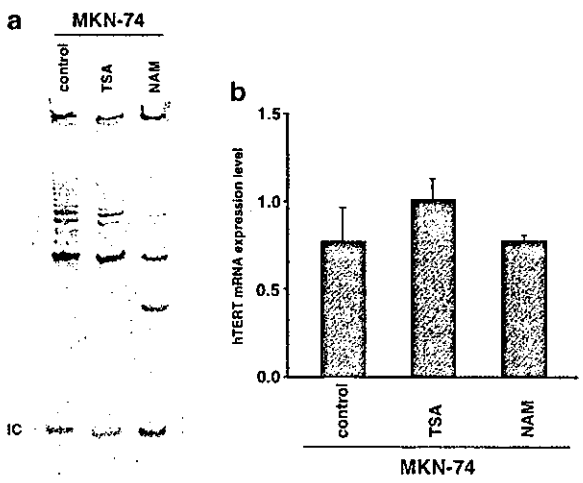


Figure 5 Alteration of telomerase activity and *hTERT* expression in MKN-74 treated with TSA and NAM. (a) Reduced telomerase activity was detected in TSA- and NAM-treated MKN-74 cells. IC = internal control (36 bp). (b) *hTERT* expression did not significantly change in TSA- and NAM-treated MKN-74 cells

might contribute to reduced expression of *PINX1* in such cases because hypermethylation of CpG islands in promoters is associated with silencing of many tumor suppressor genes in human cancers (Razin and Cedar, 1991; Kass *et al*., 1997; Jones and Baylin, 2002). We have also reported DNA methylation of tumor suppressor genes in GC (Oue *et al*., 2001, 2002, 2003; Hamai *et al*., 2003; Oshimo *et al*., 2004). Although the average level of *PINX1* expression in eight GC cell lines (0.691 ± 0.226) showed less than half that in non-neoplastic mucosa of 73 GC cases (1.766 ± 0.758), complete silencing of *PINX1* expression was not observed and bisulfite sequencing revealed that the *PINX1* CpG island methylation were little over the entire region analysed in MKN-28, MKN-74, and HSC-39 cells. Recently, it has been reported that the *PINX1* promoter is not hypermethylated in GC tissues (Akiyama *et al*., 2004).

We then investigated the possibility that histone acetylation status alters expression of *PINX1*. Both TSA and NAM treatment induce *PINX1* expression in MKN-74 cells, and acetylation of histone H4 in the 5′

UTR of *PINX1* is increased, whereas acetylation of histone H3 is not changed significantly. These data suggest that hypoacetylation of histone H4 in the 5′ UTR of *PINX1* reduces expression of *PINX1*.

Our results also indicated that *PINX1* acetylation may be controlled through two pathways of histone deacetylation. The first pathway may involve class I and/or II HDACs. The second pathway is inhibited by NAM, which suggests the involvement of the NAD-dependent TSA-resistant Sir2 deacetylase, which is a class III HDAC. Deacetylation by Sir2 occurs selectively at Lys 16 of histone H4 in yeast (Imai *et al.*, 2000). Therefore, we believe that acetylation of histone H4 in the 5′ UTR of *PINX1* was induced selectively by NAM in the present study.

Moreover, we revealed that telomerase activity was inhibited with TSA and NAM in MKN-74 cells, although *hTERT* expression was not changed. Thus, PinX1 inhibits telomerase activity without change of *hTERT* expression. Our results also provide a possibility for cancer therapy with NAM.

In conclusion, our data show that LOH of the *PINX1* locus and hypoacetylation of histone H4 in the 5′ UTR of *PINX1* are associated with reduced expression of *PINX1* in GC. However, we cannot exclude the possibility that reduced expression of *PINX1* is associated with mutation or other factors. Further studies may provide a better understanding of the physiological function of *PINX1* and its role as a tumor suppressor in stomach carcinogenesis.

Materials and methods

Tissue samples

A total of 73 GC samples from 73 patients were studied. Tumors and corresponding non-neoplastic mucosae were removed surgically at Hiroshima University Hospital, frozen immediately in liquid nitrogen, and stored at −80°C until use. We confirmed microscopically that the carcinoma tissue specimens consisted mainly (>80%) of carcinoma tissue, and that the non-neoplastic mucosae did not show any invasion by carcinoma cells or show significant inflammatory involvement. Histologic classification and tumor staging were carried out according to the Lauren classification system (Lauren, 1965) and the TNM Stage Grouping (UICC 5th Edition, 1997) (Sobin and Wittekind, 1997). Telomerase activities were determined previously in 20 of the 73 GC samples by TRAP analysis (Yasui *et al.*, 1998).

Cell lines and drug treatment

Eight cell lines derived from human GC were used. The TMK-1 cell line was established in our laboratory from a poorly differentiated adenocarcinoma (Ochiai *et al.*, 1985). Five GC 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 Dr T Suzuki. KATO-III and HSC-39 cell lines, which were established from signet ring cell carcinomas, were kindly provided by Dr M Sekiguchi and Dr K Yanagihara (Yanagihara *et al.*, 1991), respectively. All cell lines were routinely maintained in RPMI 1640 (Nissui Pharmaceutical Co., Ltd, Tokyo, Japan) containing 10% fetal bovine serum (FBS) (Bio-Whittaker, Walkersville, MD, USA) in a humidified atmosphere of 5% CO_2 and 95% air at 37°C. MKN-28 and MKN-74 cells were treated with 300 nM TSA (Wako, Tokyo, Japan) for 24 h or with 5 mM NAM (Sigma, St Louis, MO, USA) for 6 h.

Quantitative RT–PCR

Total RNA was extracted from tissues and cell lines with the RNeasy Mini Kit (Qiagen, Valencia, CA, USA). Total RNA (1 μg) was converted to cDNA with the First-Strand cDNA Synthesis Kit (Amersham Pharmacia Biotech, Uppsala, Sweden). To analyse the expression of *PINX1* gene in GC tissues specimens and GC cell lines, we performed real-time RT–PCR. PCRs were performed with the SYBR Green PCR Core Reagent kit (Applied Biosystems, Foster City, CA, USA). Real-time detection of the emission intensity of SYBR Green bound to double-stranded DNAs was performed with the ABI PRISM 7700 Sequence Detection System (Applied Biosystems). The *PINX1* cDNA and the *ACTB* cDNA (internal control) were amplified separately. Relative gene expression was determined from the threshold cycle for the *PINX1* gene and the *ACTB* gene. Reference samples (HSC-39) were included on each assay plate to verify plate-to-plate consistency. Plates were normalized to each other with these reference samples. PCR amplification was performed in 96-well optical trays with caps with a 25 μl final reaction mixture according to the manufacturer's instructions. Quantitative RT–PCRs were performed in triplicate for each sample and primer set, and the mean of the three experiments was used as the relative quantification value. We analysed *PINX1* levels in 73 cases of GC by calculating the ratio of *PINX1* mRNA expression levels between carcinoma tissue and the corresponding non-neoplastic mucosa (*T/N* ratio). We considered a $T/N<0.5$-fold to indicate reduced expression of *PINX1*. *PINX1* primer sequences were 5′-CAC TCC AGA GGA GAA CGA AAC C-3′ (sense) and 5′-CAC CGG CTT GGC AAA GTA CT-3′ (antisense). *ACTB* primer sequences were 5′-TCA CCG AGC GCG GCT-3′ (sense) and 5′-TAA TGT CAC GCA CGA TTT CCC-3′ (antisense). We used TaqMan Pre-Developed Assay Reagents Human TERT and TaqMan β-actin Control Reagents (Applied Biosystems) in *hTERT* expression analysis.

Genomic DNA extraction and LOH analysis

To examine LOH of *PINX1* locus, we extracted genomic DNAs from GC tissues with a genomic DNA purification kit (Promega, Madison, WI, USA). LOH at microsatellite marker D8S277 was evaluated by PCR of tumor and normal specimens, and microcapillary electrophoresis of PCR products was carried out in an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). GeneScan software (Applied Biosystems) was used to quantify and interpret the raw data. Allelic loss was calculated according to a previously described formula (Liloglou *et al.*, 2000, 2001). Among 73 GC cases, we evaluated 45 cases for which genomic DNA was available.

Bisulfite genomic DNA sequencing

To examine the DNA methylation patterns, we treated genomic DNA with sodium bisulfite, as described previously (Hermann *et al.*, 1996). In brief, 2 μg of genomic DNA was denatured by treatment with NaOH and modified with 3 M sodium bisulfite for 16 h. DNA samples were purified with Wizard DNA purification resin (Promega), treated with NaOH, precipitated with ethanol, and resuspended in 25 μl

of water. Treated DNAs were stored at −20°C until needed. Sodium bisulfite-treated genomic DNAs were amplified with *PINX1* gene-specific primers (primer 1 and primer 2, Figure 3b). Primer 1 sequences were 5′-TTT GAT TTT TTT GGA GTT TTT AGT-3′ (sense) and 5′-GCG ACC CAA AAT AAT TCT AAA-3′ (antisense). Primer 2 sequences were 5′-GGG TTT TTT GAT GGA GAT TTT A-3′ (sense) and 5′-ACG TTC AAC CAA CAT AAA CAT ATC-3′ (antisense). Conditions for the PCR were 1 cycle at 94°C for 10 min; 34 cycles at 94°C for 1 min, at 54°C for 1 min, and at 72°C for 1 min, and 1 cycle at 72°C for 4 min. The PCR product was cloned into the TA vector pCR2.1 (Invitrogen, Carlsbad, CA, USA). A total of 10 subclones were confirmed by restriction analysis and sequenced using the M13 reverse primer (Invitrogen) with a Prism 310 DNA Sequencer (Perkin-Elmer Applied Biosystems).

ChIP assay

ChIP assay of GC cell lines was performed as described previously with modification (Ferreira *et al*., 2001). In brief, chromatin proteins were crosslinked to DNA by addition of formaldehyde directly to the culture medium to a final concentration of 1%. After 10-min incubation at room temperature, the cells were washed and scraped off the dishes into ice-cold phosphate-buffered saline (PBS) containing protease inhibitors. Cells were pelleted and then resuspended in SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.1, protease inhibitor) for 10 min on ice. The lysate was sonicated to reduce the mean DNA fragment size to 300–1000 bp. The sample was centrifuged to remove cell debris and diluted 10-fold in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1, 167 mM NaCl, protease inhibitors). The chromatin solution was precleared with 40 μl of a mixture of salmon sperm DNA–protein A agarose slurry (Upstate Biotechnology, Lake Placid, NY, USA) to reduce nonspecific background. After precleaning, the solution was centrifuged, and the supernatant was collected. In all, 5 μl of either antiacetylated histone H3 or H4 antibody (Upstate Biotechnology) was added to the chromatin solution and incubated overnight at 4°C with agitation. A no-antibody control was also performed for each ChIP assay. After the overnight incubation, immune complexes were collected by addition of 60 μl of salmon sperm DNA–protein A agarose slurry (Upstate Biotech) and incubated at 4°C with agitation for 1 h. Beads were washed five times, and the bound immune complexes were eluted with buffer containing 1% SDS and 0.1 M $NaHCO_3$. Crosslinks were reversed by addition of 5 M NaCl and incubation at 65°C for 4 h. Samples were then treated with proteinase K for 1 h, and DNA was purified by phenol/chloroform extraction and ethanol precipitation. We performed PCR analysis of immunoprecipitated DNA using primers specific for the 5′ region of the *ACTB* gene. Each PCR product (15 μl) was loaded onto 8% nondenaturing polyacrylamide gels, separated by electrophoresis, stained with ethidium bromide, and visualized under UV light to confirm that there was no genomic DNA contamination of the no-antibody control. For quantitative PCR analysis of immunoprecipitated DNAs, we performed real-time PCR as described above.

PINX1 primer (5′ region) sequences were 5′-CCT GAG TCC AGT GCC CTA CTT T-3′ (sense) and 5′-GAA TTT TCC CAG CCA AGG C-3′ (antisense). *ACTB* primer (5′ region) sequences were 5′-CCC ACC CGG TCT TGT GTG-3′ (sense) and 5′-GGG AAG ACC CTG TCC TTG TCA-3′ (antisense).

Southern blot analysis

High molecular weight genomic DNA was extracted with a DNA Extraction Kit (Stratagene Cloning System, La Jolla, CA, USA). Tissue DNA was digested with *Msp*I, electrophoresed on 0.6% agarose gels, and blotted onto nitrocellulose filters. The filters were hybridized with the full-length cDNA PINX1 probe and then autoradiographed.

TRAP assay

TRAP assay was carried out with a TRAPEZE Telomerase Detection Kit (Intergen Company, Oxford, UK). Intensity of the TRAP product bands and of the internal control bands was determined with the use of NIH Image.

Statistical analysis

Statistical significance was assessed by Wilcoxon signed rank test, Mann–Whitney *U*-test, or Fisher's exact test. StatView 5.0 Macintosh software was used. The *P*-values of less than 0.05 were considered statistically significant.

Acknowledgements
We thank Mr M Takatani and Ms M Ueda for their excellent technical assistance and advice.

References

Akiyama Y, Maesawa C, Wada K, Fujisawa K, Itabashi T, Noda Y, Honda T, Sato N, Ishida K, Takagane A, Saito K and Masuda T. (2004). *Oncol. Rep.*, **11**, 871–874.

Artandi SE, Chang S, Lee SL, Alson S, Gottlieb GJ, Chin L and DePinho RA. (2000). *Nature*, **406**, 641–645.

Baffa R, Santoro R, Bullrich F, Mandes B, Ishii H and Croce CM. (2000). *Clin. Cancer Res.*, **6**, 1372–1377.

Bitterman KJ, Anderson RM, Cohen HY, Latorre-Esteves M and Sinclair DA. (2002). *J. Biol. Chem.*, **277**, 45099–45107.

Chin L, Artandi SE, Shen Q, Tam A, Lee SL, Gottlieb GJ, Greider CW and DePinho RA. (1999). *Cell*, **97**, 527–538.

Cong YS, Wright WE and Shay JW. (2002). *Microbiol. Mol. Biol. Rev.*, **66**, 407–425.

DePinho RA. (2000). *Nature*, **408**, 248–254.

Emi M, Fujiwara Y, Nakajima T, Tsuchiya E, Tsuda H, Hirohashi S, Maeda Y, Tsuruta K, Miyaki M and Nakamura Y. (1992). *Cancer Res.*, **52**, 5368–5372.

Ferreira R, Naguibneva I, Mathieu M, Ait-Si-Ali S, Robin P, Pritchard LL and Harel-Bellan A. (2001). *EMBO Rep.*, **2**, 794–799.

Fujiwara Y, Emi M, Ohata H, Kato Y, Nakajima T, Mori T and Nakamura Y. (1993). *Cancer Res.*, **53**, 1172–1174.

Grunstein M. (1997). *Nature*, **389**, 349–352.

Hamai Y, Oue N, Mitani Y, Nakayama H, Ito R, Matsusaki K, Yoshida K, Toge T and Yasui W. (2003). *Cancer Sci.*, **94**, 692–698.

Herman JG, Graff JR, Myohanen S, Nelkin BD and Baylin SB. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 9821–9826.

Imai S, Armstrong CM, Kaeberlein M and Guarente L. (2000). *Nature*, **403**, 795–800.

Jones PA and Baylin SB. (2002). *Nat. Rev. Genet.*, **3**, 415–428.

Kass SU, Pruss D and Wolffe AP. (1997). *Trends Genet.*, **13**, 444–449.

Kondo T, Oue N, Yoshida K, Mitani Y, Naka K, Nakayama H and Yasui W. (2004). *Cancer Res.*, **64**, 523–529.
Kuniyasu H, Domen T, Hamamoto T, Yokozaki H, Yasui W, Tahara H and Tahara E. (1997). *Jpn. J. Cancer Res.*, **88**, 103–107.
Landry J, Sutton A, Tafrov ST, Heller RC, Stebbins J, Pillus L and Sternglanz R. (2000). *Proc. Natl. Acad. Sci. USA*, **97**, 5807–5811.
Lauren P. (1965). *Acta. Pathol. Microbiol. Scand.*, **64**, 31–49.
Liao C, Zhao M, Song H, Uchida K, Yokoyama KK and Li T. (2000). *Hepatology*, **32**, 721–727.
Liao C, Zhao MJ, Zhao J, Jia D, Song H and Li ZP. (2002). *World J. Gastroentrol.*, **8**, 1050–1052.
Liao C, Zhao MJ, Zhao J, Song H, Pineau P, Marchio A, Dejean A, Tiollais P, Wang HY and Li TP. (2003). *World J. Gastroentrol.*, **9**, 89–93.
Liloglou T, Maloney P, Xinarianos G, Fear S and Field JK. (2000). *Int. J. Oncol.*, **16**, 5–14.
Liloglou T, Maloney P, Xinarianos G, Hulbert M, Walshaw MJ, Gosney JR, Turnbull L and Field JK. (2001). *Cancer Res.*, **61**, 1624–1628.
Luger K, Mader AW, Richmond RK, Sargent DF and Richmond TJ. (1997). *Nature*, **389**, 251–260.
Luger K and Richmond TJ. (1998). *Curr. Opin. Genet. Dev.*, **8**, 140–146.
Luo J, Nikolaev AY, Imai S, Chen D, Su F, Shiloh A, Guarente L and Gu W. (2001). *Cell*, **107**, 137–148.
Macoska JA, Trybus TM, Benson PD, Sakr WA, Grignon DJ, Wojno KD, Pietruk T and Powell IJ. (1995). *Cancer Res.*, **55**, 5390–5395.
Meeker AK, Hicks JL, Platz EA, March GE, Bennett CJ, Delannoy MJ and De Marzo AM. (2002). *Cancer Res.*, **62**, 6405–6409.
North BJ, Marshall BL, Borra MT, Denu JM and Verdin E. (2003). *Mol. Cell*, **11**, 437–444.
Ochiai A, Yasui W and Tahara E. (1985). *Jpn. J. Cancer Res.*, **76**, 1064–1071.
Ohata H, Emi M, Fujiwara Y, Higashino K, Nakagawa K, Futagami R, Tsuchiya E and Nakamura Y. (1993). *Genes Chromosomes Cancer*, **7**, 85–88.
Oshimo Y, Oue N, Mitani Y, Nakayama H, Kitadai Y, Yoshida K, Chayama K and Yasui W. (2004). *Int. J. Cancer*, **110**, 212–218.
O'Sullivan JN, Bronner MP, Brentnall TA, Finley JC, Shen WT, Emerson S, Emond MJ, Gollahon KA, Moskovitz AH, Crispin DA, Potter JD and Rabinovitch PS. (2002). *Nat. Genet.*, **32**, 280–284.
Oue N, Hamai Y, Mitani Y, Matsumura S, Oshimo Y, Aung PP, Kuraoka K, Nakayama H and Yasui W. (2004). *Cancer Res.*, **64**, 2397–2405.
Oue N, Matsumura S, Nakayama H, Kitadai Y, Taniyama K, Matsusaki K and Yasui W. (2003). *Oncology*, **64**, 423–429.
Oue N, Motoshita J, Yokosaki H, Hayashi K, Tahara E, Taniyama K, Matsusaki K and Yasui W. (2002). *J. Pathol.*, **198**, 55–59.
Oue N, Shigeishi H, Kuniyasu H, Yokosaki H, Kuraoka K, Ito R and Yasui W. (2001). *Int. J. Cancer*, **93**, 805–809.
Park WS, Lee JH, Park JY, Jeong SW, Shin MS, Kim HS, Lee SK, Lee SN, Lee SH, Park CG, Yoo NJ and Lee JY. (2002). *Cancer Lett.*, **178**, 199–207.
Razin A and Cedar H. (1991). *Microbil. Rev.*, **55**, 451–458.
Shay JW and Bacchetti S. (1997). *Eur. J. Cancer*, **33**, 787–791.
Smith JS, Brachmann CB, Celic I, Kenna MA, Muhammad S, Starai VJ, Avalos JL, Escalante-Semerena JC, Grubmeyer C, Wolberger C and Boeke JD. (2000). *Proc. Natl. Acad. Sci. USA*, **97**, 6658–6663.
Sobin LH and Wittekind CH (ed). (1997). *TNM Classification of Malignant Tumors* 5th edn. Wiley-Liss, Inc: New York, pp. 59–62.
Suzuki T, Yokozaki H, Kuniyasu H, Hayashi K, Naka K, Ono S, Ishikawa T, Tahara E and Yasui W. (2000). *Int. J. Cancer*, **88**, 992–997.
Tahara H, Kuniyasu H, Yokosaki H, Yasui W, Shay JW, Ide T and Tahara E. (1995). *Clin. Cancer Res.*, **1**, 1245–1251.
Van Heek NT, Meeker AK, Kern SE, Yeo CJ, Lillemoe KD, Cameron JL, Offerhaus GL, Hicks JL, Wilentz RE, Goggins MG, De Marzo AM, Hruban RH and Maitra A. (2002). *Am. J. Pathol.*, **161**, 1541–1547.
Yanagihara K, Seyama T, Tsumuraya M, Kamada N and Yokoro K. (1991). *Cancer Res.*, **51**, 381–386.
Yaremko ML, Recant WM and Westbrook CA. (1995). *Genes Chromosomes Cancer*, **13**, 186–191.
Yasui W, Tahara E, Tahara H, Fujimoto J, Naka K, Nakayama J, Ishikawa F, Ide T and Tahara E. (1999). *Jpn. J. Cancer Res.*, **90**, 589–595.
Yasui W, Tahara H, Tahara E, Fujimoto J, Nakayama J, Ishikawa F, Ide T and Tahara E. (1998). *Jpn. J. Cancer Res.*, **89**, 1099–1103.
Yasui W, Yokozaki H, Fujimoto J, Naka K, Kuniyasu H and Tahara E. (2000). *J. Gastroenterol.*, **35**, 111–115.
Yustein AS, Harper JC, Petroni GR, Cummings OW, Moskaluk CA and Powell SM. (1999). *Cancer Res.*, **59**, 1437–1441.
Zhou XZ and Lu KP. (2001). *Cell*, **107**, 347–359.

[CANCER RESEARCH 64, 2397–2405, April 1, 2004]

Gene Expression Profile of Gastric Carcinoma: Identification of Genes and Tags Potentially Involved in Invasion, Metastasis, and Carcinogenesis by Serial Analysis of Gene Expression

Naohide Oue, Yoichi Hamai, Yoshitsugu Mitani, Shunji Matsumura, Yasuhiro Oshimo, Phyu Phyu Aung, Kazuya Kuraoka, Hirofumi Nakayama, and Wataru Yasui

Department of Molecular Pathology, Hiroshima University Graduate School of Biomedical Sciences, Hiroshima, Japan

ABSTRACT

Gastric carcinoma (GC) is one of the most common malignancies worldwide. To better understand the genetic basis of this disease, we performed serial analysis of gene expression (SAGE) on four primary GC samples and one associated lymph node metastasis. We obtained a total of 137,706 expressed tags (Gene Expression Omnibus accession number GSE 545, SAGE Hiroshima gastric cancer tissue), including 38,903 that were unique. Comparing tags from our GC libraries containing different stages and different histologies, we found several genes and tags that are potentially involved in invasion, metastasis, and carcinogenesis. Among these, we selected 27 genes and measured mRNA expression levels in an additional 46 GC samples by quantitative reverse transcription-PCR. Frequently overexpressed genes (tumor/normal ratio > 2) were *COLIA1* (percentage of cases with overexpression, 78.3%), *CDH17* (73.9%), *APOC1* (67.4%), *COLIA2* (58.7%), *YF13H12* (52.2%), *CEACAM6* (50.0%), *APOE* (50.0%), *REGIV* (47.8%), *S100A11* (41.3%), and *FUS* (41.3%). Among these genes, mRNA expression levels of *CDH17* and *APOE* were associated with depth of tumor invasion ($P = 0.0060$ and $P = 0.0139$, respectively), and those of *FUS* and *APOE* were associated with degree of lymph node metastasis ($P = 0.0416$ and $P = 0.0006$, respectively). In addition, mRNA expression levels of *FUS*, *COLIA1*, *COLIA2*, and *APOE* were associated with stage ($P = 0.0414$, $P = 0.0156$, $P = 0.0395$, and $P = 0.0125$, respectively). Quantitative reverse transcription-PCR analysis also showed a high level of *REGIV* expression (>100 arbitrary units) in 14 of 46 GC samples (30.4%) but not in noncancerous tissues. We detected V5-tagged RegIV protein in the culture media of cells transfected with pcDNA-RegIV-V5 by Western blot. Our results provide a list of candidate genes that are potentially involved in invasion, metastasis, and carcinogenesis of GC. *REGIV* may serve as a specific biomarker for GC.

INTRODUCTION

Gastric carcinoma (GC) is one of the most common human cancers. Despite improvements in cancer therapy, ~650,000 patients with GC die/year (1). A variety of genetic and epigenetic alterations are associated with GC (2–4). However, the underlying mechanism of gastric carcinogenesis is still poorly understood. To identify potential molecular markers for GC and to better understand the development of GC at the molecular level, comprehensive gene expression analysis may be useful. Although several large-scale gene expression studies with cDNA or oligonucleotide arrays have been performed in GC (5–8), they have used different platforms that varied in the number and identity of genes printed on them. On the other hand, serial analysis of gene expression (SAGE) analyzes 14-bp tags derived from defined positions of cDNAs without *a priori* knowledge of the sequence of the genes expressed (9). Thus, SAGE offers an unbiased, comprehensive gene expression profiling approach. Recently, three SAGE studies of GC were reported, and several up-regulated and down-regulated genes were identified (10–12). However, only one (10) or two samples (11, 12) were examined, and the relation to invasion and metastasis was not analyzed. In the present study, we performed SAGE analysis on four samples of GC of different stages and different histologies. In addition, we performed SAGE analysis on one lymph node metastasis of GC. We report here the identification of several genes and tags potentially involved in invasion, metastasis, and carcinogenesis of GC. Among these, we focused on the *REGIV* gene because this gene is frequently overexpressed in GC, and *REGIV* expression is narrowly restricted in noncancerous tissues. In addition, the amino acid sequence of the RegIV protein suggests that it may be secreted.

MATERIALS AND METHODS

Tissue Samples. For SAGE analysis, four primary GC samples and 1 associated lymph node metastasis were used (Table 1). We confirmed microscopically that the tumor specimens consisted mainly (>80%) of carcinoma tissue with the exception of S219T. For quantitative reverse transcription-PCR (RT-PCR), 46 GC samples and corresponding nonneoplastic mucosa samples were used. Of the 46 GC samples, lymph node metastasis samples were available for 9. The samples were obtained at surgery at Hiroshima University Hospital and affiliated hospitals. Noncancerous samples of the heart, aorta, lung, tongue, esophagus, stomach, duodenum, ileum, colon, liver, gallbladder, pancreas, kidney, urinary bladder, thyroid gland, adrenal gland, spleen, skin, endometrium, and lymph node were obtained at autopsy from a 28-year-old woman diagnosed with multiple sclerosis. Samples were frozen immediately in liquid nitrogen and stored at −80°C until use. Histological classification of GC was performed according to the Lauren classification system (13). In addition, diffuse-type GC samples were additionally classified into diffuse-adherent and diffuse-scattered subtypes (14). Tumor staging was carried out according to the Tumor-Node-Metastasis Stage Grouping (15).

SAGE. SAGE was performed 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 (The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine). Clinicopathological details of the 5 samples are shown in Table 1. To identify genes involved in tumor progression, we analyzed 2 GC samples (W226T and W246T). Both samples were classified as intestinal type GC. However, W226T was early, and W246T was advanced. Early GC is limited to the mucosa or the mucosa and submucosa, regardless of nodal status (16). We confirmed microscopically that these 2 samples showed similar histological features (Fig. 1). To identify genes involved in tumor metastasis, we analyzed 1 GC sample (P208T) and its lymph node metastasis (P208L). Histologically, these samples were classified as diffuse-adherent type, and we confirmed microscopically that both the primary tumor (P208T) and the metastatic tumor (P208L) contained few stromal cells and lymphocytes (Fig. 1). Scirrhous-type GC belongs to the diffuse-scattered type, often occurs in young women, and is characterized by extensive fibrous stroma, infiltrative and rapid growth, and poor prognosis (17). Sample S219T was a scirrhous-type GC showing scattering growth in an abundant fibrous stroma (Fig. 1). To

Received 11/10/03; revised 1/20/04; accepted 1/26/04.

Grant support: Grants-in-Aid for Cancer Research from the Ministry of Education, Culture, Science, Sports, and Technology of Japan and from the Ministry of Health, Labor, and Welfare of Japan.

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

Note: Supplementary data for this article are available at *Cancer Research Online* (http://cancerres.aacrjournals.org). This work was carried out with the kind cooperation of the Research Center for Molecular Medicine, Faculty of Medicine, Hiroshima University.

Requests for reprints: Wataru Yasui, Department of Molecular Pathology, Hiroshima University Graduate School of Biomedical Sciences, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan. Fax: 81-82-257-5149; E-mail: wyasui@hiroshima-u.ac.jp.

Table 1 *Clinicopathological details of the 5 samples analyzed by serial analysis of gene expression*

Sample name	Sample type	Sex	Age (yrs)	Tumor[a]	Node[a]	Metastasis[a]	Stage[a]	Histological classification[b]	Total number of tags	Number of unique tags in each sample
W226T	Primary	Male	59	1	0	0	IA	Intestinal	43,908	16,082
W246T	Primary	Male	44	2	2	0	IIIA	Intestinal	32,174	12,792
S219T	Primary	Female	29	3	3	0	IV	Diffuse-scattered	34,660	14,576
P208T	Primary	Male	60	4	3	0	IV	Diffuse-adherent	11,582	6,135
P208L	Lymph node metastasis of P208T								15,382	7,425
									137,706[c]	38,903[d]

[a] Tumor staging of gastric carcinoma (GC) were done according to the Tumor-Node-Metastasis Stage Grouping (15).
[b] Histological classification of GC was done according to the Lauren classification system (13). In addition, diffuse-type GC were additionally classified into diffuse-adherent and diffuse-scattered subtypes (14).
[c] Total number of tags of 5 GC samples.
[d] Total number of unique tags among 5 GC samples.

permit direct comparison, each library was normalized to a total of 1,000,000 tags.

Cluster Analysis. The Cluster and TreeView computer programs were obtained from online resources.[1] We compared SAGE tags from 4 primary GC samples with those from samples of normal gastric epithelia [GSM784, SAGE normal gastric body epithelial (10)], available from SAGEmap (18).[2] We also compared SAGE tags from 2 primary GC samples, also available from SAGEmap [GSM757, SAGE gastric cancer-G234 (10) and GSM2385, SAGE gastric cancer-G189] with those from normal gastric epithelia (GSM784) and obtained the 20 most up-regulated and 20 most down-regulated tags. This produced a dataset of 128 tags. These data were imported into the Cluster program and were log-transformed, and complete linkage clustering was performed.

Quantitative RT-PCR Analysis. Total RNA was extracted with an RNeasy Mini kit (Qiagen, Hilden, Germany), and 1 μg of total RNA was converted to cDNA with a first-strand cDNA synthesis kit (Amersham Pharmacia Biotech, Uppsala, Sweden). PCR was performed with a SYBR Green PCR Core Reagents kit (Applied Biosystems, Foster City, CA). Real-time detection of the emission intensity of SYBR green bound to double-stranded DNA was performed with an ABI PRISM 7700 Sequence Detection System (Applied Biosystems) as described previously (19). The sequences primer are listed in Supplementary Table 1. We calculated the ratio of target gene mRNA expression levels between GC tissue (T) and corresponding nonneoplastic mucosa (N). T/N ratios > 2-fold were considered to represent overexpression. Genes with T/N ratios > 2 in >40% of the samples examined were defined as frequently up-regulated genes.

Cell Lines, Expression Vector, and Western Blot. Two cell lines derived from human GC were used. MKN-28 was kindly provided by Dr. Toshimitsu Suzuki. HSC-39 was kindly provided by Dr. Kazuyoshi Yanagihara (20). All cell lines were maintained in RPMI 1640 (Nissui Pharmaceutical Co., Ltd., Tokyo, Japan) containing 10% fetal bovine serum (BioWhittaker, Walkersville, MD) in a humidified atmosphere of 5% CO_2 and 95% air at 37°C. For constitutive expression of the *RegIV* gene, cDNA was PCR amplified and subcloned into pcDNA 3.1 (Invitrogen Corp., Carlsbad, CA) in-frame with a COOH-terminal V5 epitope tag. MKN-28 cells were transfected transiently with *REGIV* cDNA using FuGene6 Transfection Reagent (Roche Diagnostics Co., Indianapolis, IN) according to the manufacturer's instructions. For Western blot analysis, cells and culture media from MKN-28 cells transfected with pcDNA 3.1 or pcDNA-RegIV-V5 were lysed as described previously (21). The culture media was concentrated with a PROTEIN concentrate kit (Takara Bio, Inc., Shiga, Japan). The lysates (40 μg) were solubilized in Laemmli sample buffer by boiling and then were subjected to 15% SDS-PAGE followed by electrotransfer onto a nitrocellulose filter. Anti-V5 monoclonal antibody was purchased from Invitrogen Corp. Peroxidase-conjugated antimouse IgG was used as the secondary antibody. The immune complex was visualized with an ECL Western blot detection system (Amersham Pharmacia Biotech).

Statistical Methods. Statistical analyses were performed with the Mann-Whitney *U* test. *P* of <0.05 was regarded as statistically significant.

[1] Internet address: http://www.microarrays.org/software.html.
[2] Internet address: http://www.ncbi.nlm.nih.gov/SAGE/.

RESULTS

Generation of SAGE Data. A total of 137,706 tags was generated, including 38,903 that were unique. The numbers of tags and unique tags are shown in Table 1. Sequence data from our SAGE libraries are publicly available at SAGEmap (GEO accession number GSE 545, SAGE Hiroshima gastric cancer tissue).

Comparison of Expression Patterns in GCs and Normal Stomach. We compared SAGE tags from 4 primary GC samples with those from normal gastric epithelia (GSM784). The 20 most up-regulated tags and 20 most down-regulated tags in each GC are shown in Supplementary Table 2. Among the up-regulated tags, 12 were commonly up-regulated in both W226T (intestinal type) and W246T (intestinal type). These tags included *lysozyme* (*LYZ*), *trefoil factor 3* (*TFF3*), *aldolase A* (*ALDOA*), and *S100 calcium-binding protein*, which may participate in the genesis of intestinal type GC. P208T (diffuse-adherent type) and S219T (diffuse-scattered type) showed many different tags from those of W226T and W246T. The down-regulated tags were similar in all 4 GC samples and included *lipase* (*LIPF*), *pepsinogen* (*PGA5*), and *antrum mucosa protein* (*AMP18*), which are expressed physiologically in normal gastric glands.

The SAGE data were also analyzed by a clustering algorithm to delineate patterns in the expression of 128 tags among all four libraries (our four GC libraries, our one lymph node metastasis library, and three libraries available from SAGEmap[2]; Fig. 2). These tags were selected as described in "Materials and Methods." Clusters of coexpressed tags suggested that the two intestinal type GC libraries (W226T and W246T), despite being derived from 2 different patients at different stages, were the most similar to each other. The primary GC (P208T) and its lymph node metastasis (P208L) appeared not to be similar to each other. To identify ideal biomarkers for GC, we focused on a cluster of 14 tags, the expression of which was up-regulated in 6 GC samples (our four GC libraries plus two GC libraries available from SAGEmap) and 1 lymph node metastasis. Because some genes share the same SAGE tag, these 14 tags represented 22 genes (Fig. 2 and Table 2). To validate the SAGE data, the expression of 12 known genes was analyzed by quantitative RT-PCR of an additional 46 GC samples and corresponding nonneoplastic mucosa samples. Frequently overexpressed genes were *APOC1* (percentage of samples with overexpression: T/N ratio > 2; 67.4%), *YF13H12* (52.2%), and *CEACAM6* (50.0%; Fig. 3*A*, see also Supplementary Fig. 1). Other genes were less frequently overexpressed. The expression levels of all 12 genes were not associated with T grade (depth of tumor invasion), N grade (degree of lymph node metastasis), or tumor stage.

Comparison of Expression Patterns in Early and Advanced GC. To identify genes involved in tumor progression, we compared tags from early GC (W226T) and advanced GC (W246T). The 10

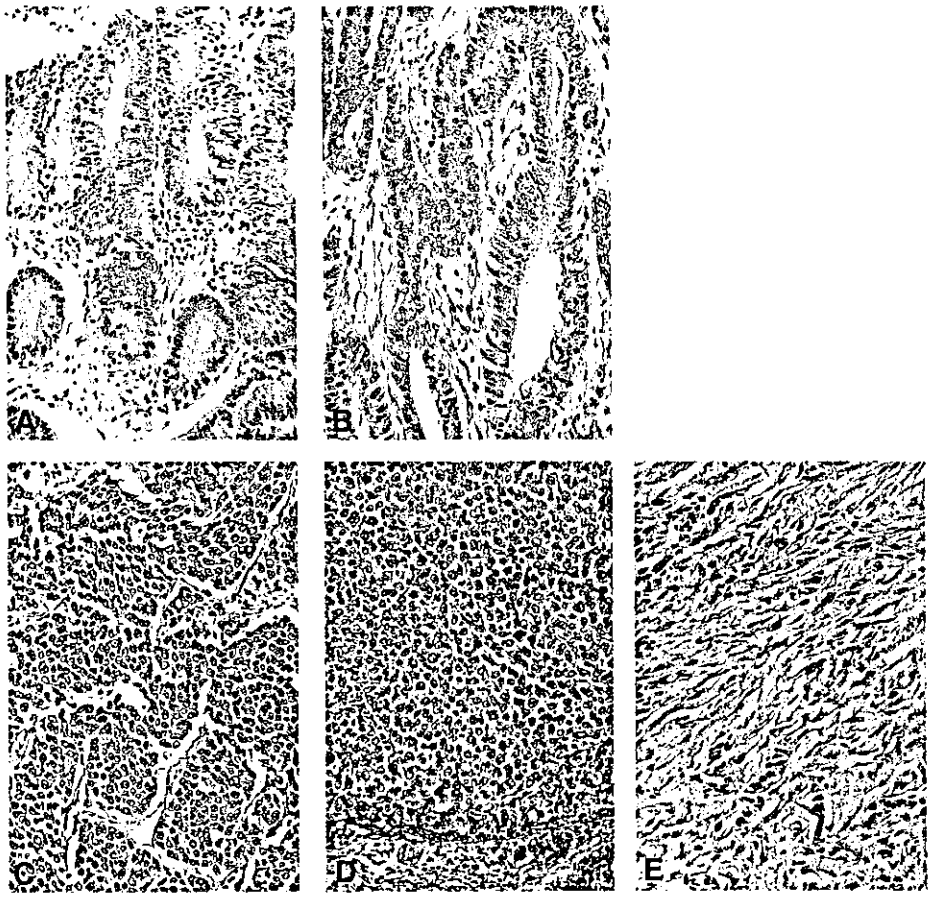


Fig. 1. Histological features of gastric carcinoma (GC) samples analyzed by serial analysis of gene expression. Formalin-fixed, paraffin-embedded sections were stained with H&E. Both W226T (*A*) and W246T (*B*) were to intestinal type GC, and histological features were similar. Both P208T (*C*) and P208L (*D*) were diffuse-adherent type GC. P208L was a lymph-node metastasis of P208T. S219T (*E*) was diffuse-scattered type GC. Obvious histological heterogeneity was not seen in all specimens (original magnification, ×100).

most up-regulated tags are shown in Table 3, and the 10 most down-regulated tags are shown in Supplementary Table 3. Because some genes share the same SAGE tag, these up-regulated 10 tags represented 12 genes (Table 3). To validate the SAGE data, the expression of 9 known genes was analyzed by quantitative RT-PCR of an additional 46 GC samples and corresponding nonneoplastic mucosa samples. Genes frequently overexpressed in GC compared with nonneoplastic mucosa were *COL1A1* (78.3%), *CDH17* (73.9%), *COL1A2* (58.7%), and *FUS* (41.3%; Fig. 3*B*, see also Supplementary Fig. 2). Other genes were less frequently overexpressed. The mRNA expression levels of *CDH17* were associated with T grade ($P = 0.0060$). The mRNA expression levels of *FUS* were associated with N grade ($P = 0.0416$). The mRNA expression levels of *FUS*, *COL1A1*, and *COL1A2* were associated with tumor stage ($P = 0.0414$, $P = 0.0156$, and $P = 0.0395$, respectively; Table 4).

Comparison of Expression Patterns in Primary GC and Associated Lymph Node Metastasis. To identify genes involved in tumor metastasis, we compared tags from primary GC (P208T) and its lymph node metastasis (P208L). The 10 most up-regulated tags are shown in Table 5, and the 10 most down-regulated tags are shown in Supplementary Table 4. The up-regulated tags represented 12 genes (Table 5). To validate the SAGE data, the expression of 5 known genes was analyzed by quantitative RT-PCR of an additional 46 GC samples and their lymph node metastases in 9 samples. A frequently overexpressed gene in lymph node metastasis compared with primary GC was not found (Fig. 3*C*, see also Supplementary Fig. 3). *APOE* mRNA expression in lymph node metastasis tended to be higher than that in primary GC. Other genes were less frequently overexpressed. Genes frequently overexpressed in GC compared with nonneoplastic

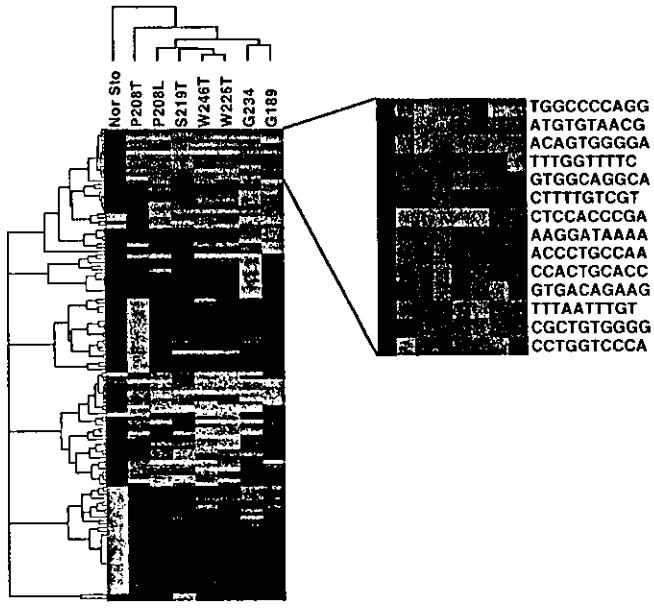


Fig. 2. Cluster analysis of 128 tags form eight serial analysis of gene expression (SAGE) libraries and dendrogram showing similarities in expression patterns among libraries. Tags were selected as described in the "Materials and Methods." *Brackets* indicate the cluster of tags commonly up-regulated in gastric carcinoma, which is expanded in size on the right for visualization. On the dendrogram, two intestinal type gastric carcinoma samples cluster together, indicating their high degree of similarity. Each row represents a tag, whereas each column corresponds to a SAGE library sample. The absolute abundance of the SAGE tag in the library (SAGE tag number) correlates with the intensity of the red color (*black*, not present; *intense red*, highly abundant).

Table 2 *Genes and tags commonly up-regulated in gastric carcinoma obtained by serial analysis of gene expression*

Tag sequence	UniGene ID	Symbol	Description
TGGCCCCAGG	Hs.268571	*APOC1*	Apolipoprotein C-I
ATGTGTAACG	Hs.81256	*S100A4*	S100 calcium binding protein A4 (calcium protein, Calvasculin, metastasin, murine placental homologue)
	Hs.173611	*NDUFS2*	NADH dehydrogenase (ubiquinone) Fe-S protein 2, 49 kDa (NADH-coenzyme Q reductase)
ACAGTGGGGA	Hs.278270	*TEBP*	Unactive progesterone receptor, 23 kDa
	Hs.288443		Homosapiens transcribed sequences
	Hs.355693		Homosapiens transcribed sequence with strong similarity to protein pir:A56211 (H. sapiens) A56211 progesterone receptor-related protein P23-human
TTTGGTTTTC	Hs.179573	*COL1A2*	Collagen, type I, α2
	Hs.21431	*SUFU*	Suppressor of fused homologue (Drosophila)
GTGGCAGGCA	Hs.47334	*SYAP1*	Synapse associated protein 1, SAP47 homologue (Drosophila)
	Hs.13255	*KIAA0930*	KIAA0930 protein
CTTTTGTCGT	Hs.19597	*KIAA1694*	KIAA1694 protein
CTCCACCCGA	Hs.82961	*TFF3*	Trefoil factor 3 (intestinal)
AAGGATAAAA	Hs.73848	*CEACAM6*	Carcinoembryonic antigen-related cell adhesion molecule 6 (nonspecific cross reacting antigen)
ACCCTGCCAA	Hs.405871	*FLJ20249*	Hypothetical protein FLJ20249
CCACTGCACC	Hs.6853	*FLJ22167*	Hypothetical protein FLJ22167
	Hs.146844		na similar to hypothetical protein FLJ10891
GTGACAGAAG	Hs.356129	*EIF4A1*	Eukaryotic translation initiation factor 4A, isoform 1
TTTAATTTGT	Hs.182793	*GOLPH2*	Golgi phosphoprotein 2
	Hs.220689	*G3BP*	Ras-GTPase-activating protein SH3-domain-binding protein
CGCTGTGGGG	Hs.7486	*YF13H12*	Protein expressed in thyroid
CCTGGTCCCA	Hs.23881	*KRT7*	Keratin 7
	Hs.167679	*SH3BP2*	SH3-domain binding protein 2
TGGAAATGAC	Hs.172928	*COL1A1*	Collagen, type I, α1
	Hs.20506	*LOC284371*	Hypothetical protein LOC284371

mucosa were *APOE* (50.0%) and *S100A11* (41.3%; Fig. 3*D*). The mRNA expression levels of *APOE* were associated with T grade ($P = 0.0139$), N grade ($P = 0.0006$), and tumor stage ($P = 0.0125$; Table 4).

REGIV Overexpression in GC. Among the 20 up-regulated tags in each GC sample (Supplementary Table 2), we focused on *REGIV* because *REGIV* expression was narrowly restricted by Virtual Northern analysis by SAGEmap (Fig. 4*A*). Besides GCs, *REGIV* was detected at low levels in only eight libraries, including one colon cancer and two normal colon libraries. Quantitative RT-PCR analysis showed overexpression of the *REGIV* gene in 22 samples of the 46 GC samples (47.8%; Fig. 4*B*). When we focused on *REGIV* gene expression in GC, high levels of *REGIV* expression (>100, arbitrary units) were found in 14 of 46 samples (30.4%; Fig. 4*C*). Among various normal tissues obtained from an autopsy, obvious *REGIV* expression was found in noncancerous stomach, duodenum, ileum, colon, and pancreas, as reported elsewhere (22). However, the levels of *REGIV* expression were low (<60 arbitrary units).

Analysis of the amino acid sequence of the RegIV protein suggests that it may be secreted. To investigate whether RegIV is a secreted protein, we performed Western blot analysis of cell extracts and culture media of MKN-28 cells transiently transfected with pcDNA 3.1 or pcDNA-RegIV-V5. With an anti-V5 antibody, we detected an approximate M_r 20,000 band corresponding to V5-tagged RegIV protein in cell extracts and culture media from RegIV-V5-expressing MKN-28 cells but not in control cells (Fig. 4*D*).

DISCUSSION

To identify potential molecular markers for GC and to better understand the development of GC at the molecular level, we performed SAGE on 5 GC samples from 4 patients that showed distinct histological types and tumor stages. We analyzed with respect to (*a*) commonly up-regulated genes in GC compared with normal stomach, (*b*) up-regulated genes in advanced compared with early GC (genes potentially involved in tumor progression), (*c*) up-regulated genes in GC lymph node metastasis compared with primary GC (genes potentially involved in tumor metastasis), and (*d*) genes specifically ex-

Table 3 *The 10 most up-regulated tags in advanced gastric carcinoma in comparison with early gastric carcinoma*

Tag sequence	Tags/million W226T	Tags/million W246T	UniGene ID	Symbol	Description
TCCCCGTAAA	22[a] (1)[b]	559 (18)			No match
TCCCGTACAT	0 (0)	279 (9)			No match
AAAAGAGTGG	0 (0)	217 (7)	Hs.89436	*CDH17*	Cadherin 17, LI cadherin (liver-intestine)
			Hs.99969	*FUS*	Fusion, derived from t(12;16) malignant liposarcoma
CCAGAGAACT	0 (0)	217 (7)	Hs.356442	*PRO1073*	PRO1073 protein
AACCTCCCCA	0 (0)	186 (6)	Hs.137396		Sapiens cDNA FLJ36926 fis, clone BRACE2005196
			Hs.232092		Sapiens cDNA FLJ30146 fis, clone BRACE2000256
AATACTTTTG	0 (0)	186 (6)	Hs.356427	*PAI-RBP1*	PAI-1 mRNA-binding protein
			Hs.179573	*COL1A2*	Collagen, type I, α2
TCCTATTAAG	22 (1)	372 (12)			No match
TGGAAATGAC	0 (0)	186 (6)	Hs.172928	*COL1A1*	Collagen, type I, α1
			Hs.193076	*GRAP2*	GRB2-related adaptor protein 2
TCCCCGTACA	227 (10)	3325 (107)	Hs.2730	*HNRPL*	Heterogeneous nuclear ribonucleoprotein L
			Hs.151734	*NUTF2*	Nuclear transport factor 2
AAGTGAAACA	0 (0)	155 (5)	Hs.93659	*ERP70*	Protein disulfide isomerase-related protein (calcium-binding protein, intestinal-related)

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

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	
COLIA1				
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	
COLIA2				
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

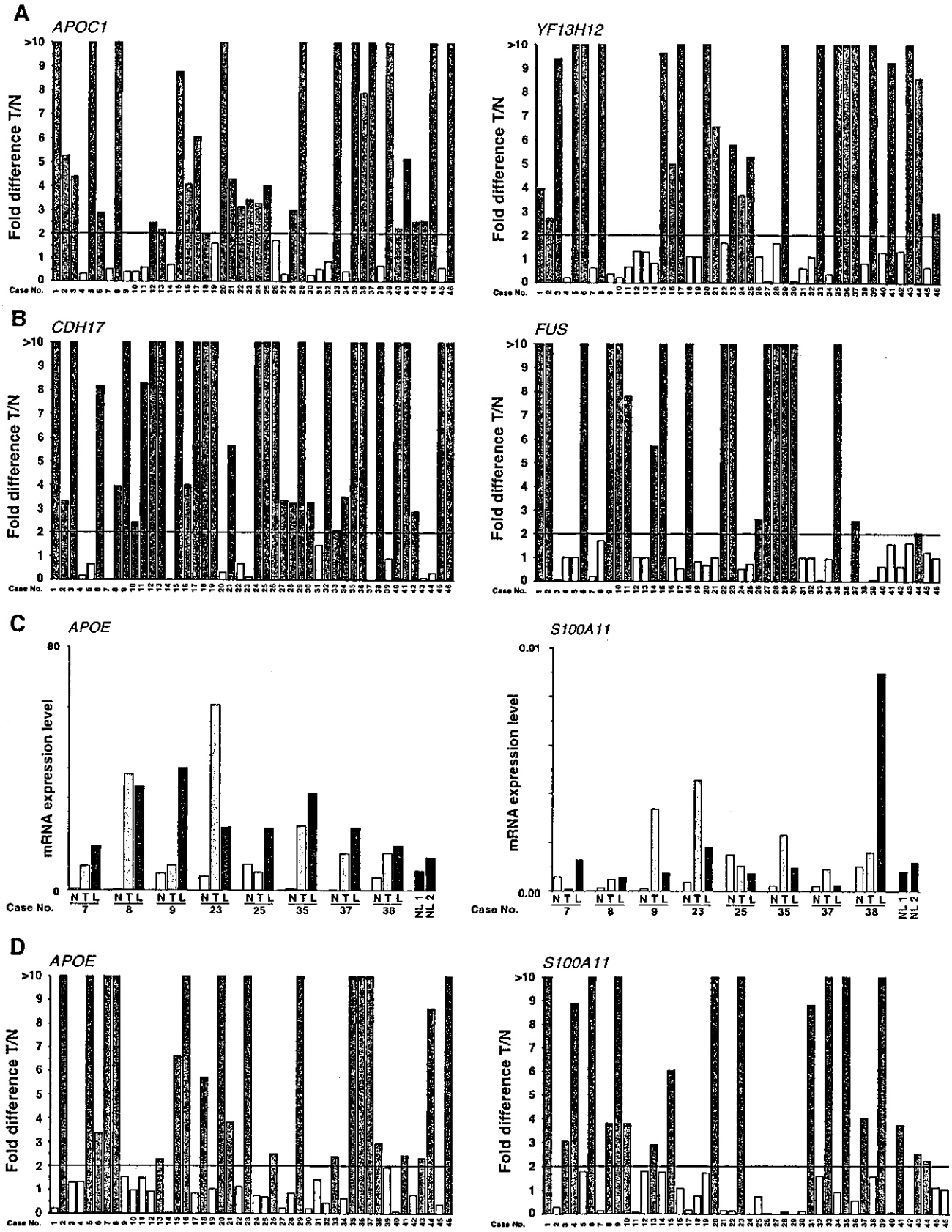
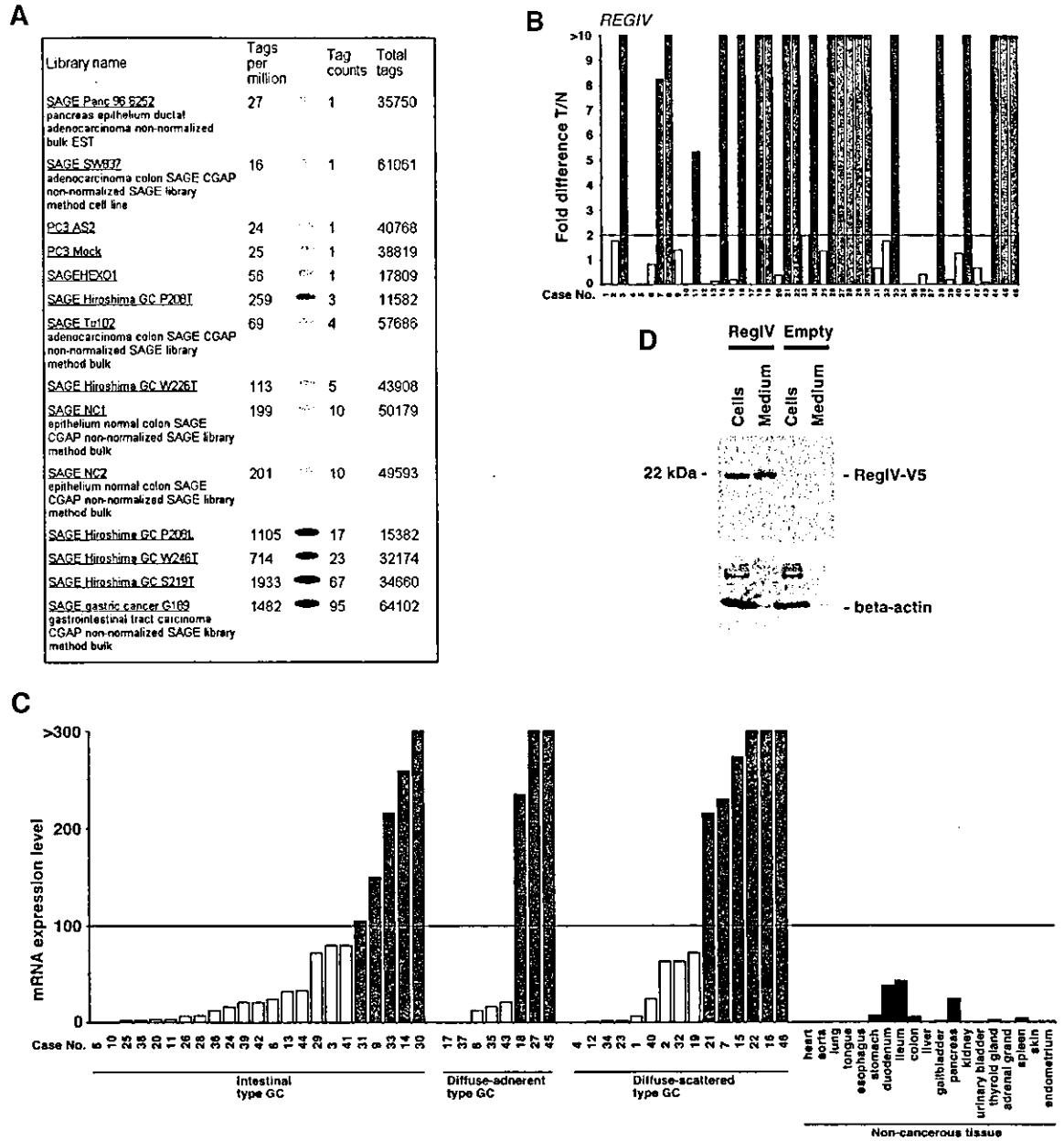


Fig. 3. Validation of serial analysis of gene expression (SAGE) data by quantitative reverse transcription-PCR (RT-PCR). Fold change indicates the ratio of target gene mRNA level in gastric carcinoma (GC) to that in corresponding nonneoplastic mucosa. *A*, quantitative RT-PCR analysis of genes commonly 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
SAGEHEXO1	56	1	17809
SAGE Hiroshima GC P208T	259	3	11582
SAGE Tu102 adenocarcinoma colon SAGE CGAP non-normalized SAGE library method bulk	69	4	57686
SAGE Hiroshima GC W226T	113	5	43908
SAGE NC1 epithelium normal colon SAGE CGAP non-normalized SAGE library method bulk	199	10	50179
SAGE NC2 epithelium normal colon SAGE CGAP non-normalized SAGE library method bulk	201	10	49593
SAGE Hiroshima GC P208L	1105	17	15382
SAGE Hiroshima GC W246T	714	23	32174
SAGE Hiroshima GC S219T	1933	67	34660
SAGE gastric cancer G189 gastrointestinal tract carcinoma CGAP non-normalized SAGE library method bulk	1482	95	64102

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

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

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

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

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

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

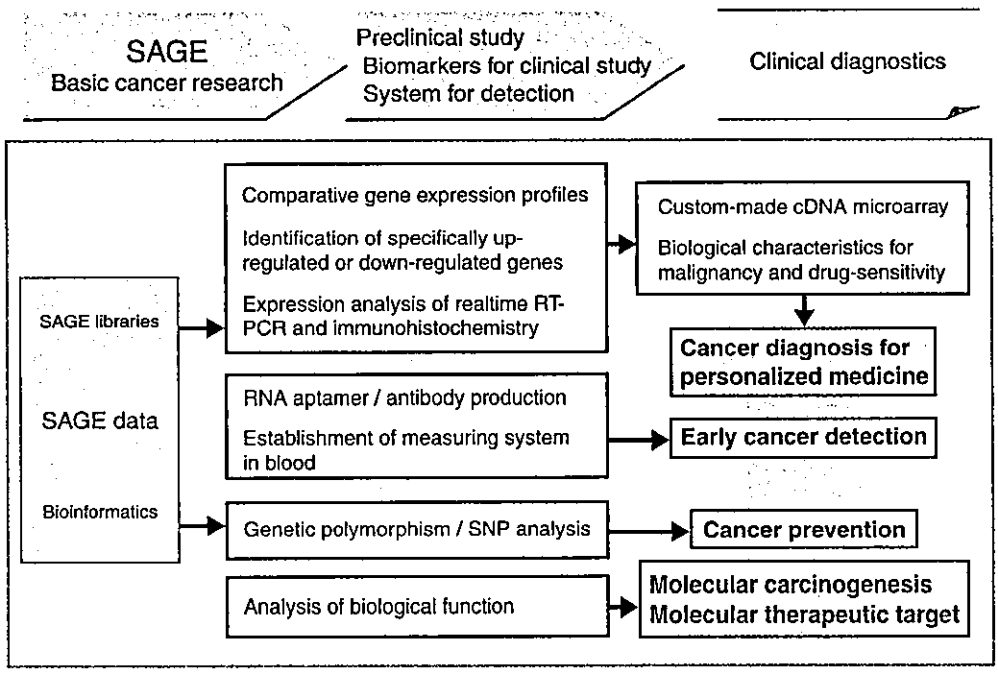


Fig. 1. Strategy to search for novel genes associated with gastric cancer through SAGE, and its clinical implication.