An 8-gene signature, including methylated and down-regulated *glutathione peroxidase 3*, of gastric cancer

XIANGLAN ZHANG[1*], JUNG-JIN YANG[2*], YOON SUN KIM[3], KI-YEOL KIM[4], WOONG SHICK AHN[3] and SANGHWA YANG[1]

[1]Cancer Metastasis Research Center, Yonsei University College of Medicine, Seoul 120-752;
[2]School of Computer Science and Information Engineering, The Catholic University of Korea, Bucheon 420-743;
[3]Department of Obstetrics and Gynecology, College of Medicine, The Catholic University of Korea, Seoul 137-701;
[4]Oral Cancer Research Institute, College of Dentistry, Yonsei University, Seoul 120-752, Korea

Received September 14, 2009; Accepted October 26, 2009

DOI: 10.3892/ijo_00000513

Abstract. We have identified an 8-gene signature with significant expression differences between gastric cancer and normal gastric tissues. This 8-gene set can predict the normal and cancer status of gastric tissues with more than 96% accuracy in a totally independent microarray dataset. The 8 genes are composed of down-regulated KLF4, GPX3, SST and LIPF, together with up-regulated SERPINH1, THY1 and INHBA in gastric cancer. To corroborate the differential gene expression pattern, we chose GPX3 and examined its expression pattern in detail. A comparison of GPX3 expression pattern shows a broader down-regulated pattern in multiple types of cancers, including cervical, thyroid, head and neck, lung cancers and melanoma than in healthy controls. An immuno-histostaining analysis in tissue microarrays confirms GPX3 down-regulation in gastric cancer. Mechanism-wise GPX3 down-regulation in gastric cancer is due to promoter hypermethylation. Collectively, these results show a correct identification of 8 genes as gastric cancer biomarkers.

Introduction

Gastric cancer, with its fourth-highest cancer incidence rate, is the second most important cause of cancer-related deaths and a major health threat world-wide (1). In an effort to identify novel biomarkers that can be used as molecular probes in the understanding of gastric cancer, we have previously performed 86 cases of global gene expression profiling and 30 microarray comparative genomic hybridization (CGH) analyses of Korean gastric cancer patient tissues, both of which were performed on complementary DNA (cDNA) microarrays containing 17,000 probes, and reported on the selection of genes whose variations in expression levels and/or copy number alteration are correlated with clinical parameters (2-5). In the gene expression profiling, we initially reported on the 94 genes differentially expressed between the tumors and non-tumorous gastric tissues and, in a further report, presented the selection of 22 biomarker candidates from 94 genes for gastric cancer (3,4).

DNA microarray technology, due to its ability to screen tens of thousands of genes in an unbiased way, has been most successfully applied to biomedical research field in identifying cancer-related genes and molecular pathways. However, use of various types of microarray formats, the quantity and quality of specimens used and especially the non-strict use of statistical measures often times led to reporting of genes, whose expression pattern sometimes cannot be reproduced by a third party. One of the ways to overcome this problem was suggested to be to use a stricter gene selection standard, including lower p-values in the statistical analyses (6).

In this study, as a way of validating the gene selection, we applied the previously identified 22 in the prediction of tissue types in an independent public microarray test set. A detailed analysis of one of the down-regulated genes, GPX3, is also presented.

Materials and methods

Public microarray data acquisition and data analysis. A public gastric cancer gene expression-profiling dataset composed of 22 gastric cancers and 8 non-cancerous gastric tissues (7), performed on Affymetrix GeneChip Human Full Length Array HuGeneFL, was downloaded from Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo). A set of 22 genes we previously showed that could distinguish tumor and non-tumorous gastric tissues was imported to the dataset. Eleven probes representing 10 common genes were present. A t-test was further performed to select a smaller group of genes differentially expressed within the 11 probes at a significance level of $p<0.001$.

Correspondence to: Dr Sanghwa Yang, Cancer Metastasis Research Center, Yonsei University College of Medicine, Seoul 120-752, Korea
E-mail: yangsanghwa@gmail.com

*Contributed equally

Key words: gastric cancer, biomarker, gene expression profiling, methylation

Table I. Oligonucleotide primers used for the RT-PCR.

Gene	Forward primer (5'-3')	Reverse primer (5'-3')	Tm[a]
GPX3	CTGCTTTCCCTGCTCCTG	GTCCATCTTGACGTTGCTGA	52.9/53.1
SST	CTGATCCGCGCCTAGAGTT	CAAGGGTCTCGCTGAAGACT	54.2/53.2
CHGA	CACAGCGGTTTTGAAGATGA	CTGGGAGTGCTCCTGTTCTC	53.3/53.4
KLF4	ACCCTGGGTCTTGAGGAAGT	TTCTGGCAGTGTGGGTCATA	53.7/53.4
LIPF	GTTGTGTTTTTGCAGCATGG	CCCCAGTCATAAGCTTGGAA	53.7/53.9
ACTB	CTCTTCCAGCCTTCCTTCCT	CACCTTCACCGTTCCAGTTT	53.8/53.7

[a]Tm, melting temperatures for forward and reverse primers, respectively. All primer sets were designed to amplify ~600 base pair-fragment of each gene in primer 3 (38).

Cell culture and 5-aza-dC treatment. Gastric cancer cells, AGS, MKN-45, NCI-N87, SNU-1, SNU-484, KATOIII, SNU601, SNU719, MKN1 and MKN74 were from Korean Cell Line Bank (Seoul, Korea). Cells were maintained in RPMI-1640 (Sigma, St. Louis, MO) supplemented with 10% heat-inactivated fetal bovine serum, 100 μg/ml streptomycin and 50 U/ml penicillin at 37˚C in a humidified incubator containing 5% CO_2. For demethylation experiment, $5x10^5$ cells were grown for 24 h before 5-aza-dC (Sigma) treatment. 5-aza-dC in PBS was filtered and treated to the final concentration of 1.6 μM for 3 days with media and 5-aza-dC changes at every 24 h. PBS-treated cells were grown alongside as untreated controls.

Reverse transcription-PCR. Total RNAs were isolated from 5-aza-dC-treated and -untreated control cells using Allprep DNA/RNA Mini Kit (Qiagen GmbH, Hilden, Germany). To produce cDNAs, 10 μg of total RNA were incubated with 4 μg of oligo-dT primer in a total volume of 13.4 μl at 65˚C for 15 min and was cooled on ice. Superscript II (400 U) (Invitrogen, Carlsbad, CA), reaction buffer, DTT and dNTPs mix were added as recommended by the manufacturer to the final volume of 25 μl and incubated for 90 min at 42˚C. The cDNAs were purified and eluted in 100 μl distilled H_2O using QIAquick PCR Purification Kit (Qiagen). For the PCR, 10 μl of the cDNA reaction mixture was used in a final 50 μl PCR cycling consisted of 95˚C for 5 min, 30 cycles of 94˚C for 30 sec, 48˚C for 30 sec, and 72˚C for 60 sec, followed by a final incubation at 72˚C for 10 min in MJ Research PTC-200 apparatus (MJ Research, Watertown, MA). Oligonucleotide primers used for the PCR are listed in Table I.

Bisulfite sequencing for confirmation of promoter methylation. To extract the promoter sequence for GPX3, the genomic sequence surrounding GPX3 promoter and coding region was obtained from Ensembl Genome Browser (www.ensembl.org). GPX3 gene is almost 8.5 kbs in genomic sequence length, with 5 exons that encode a 226-residue soluble protein. DNA sequence of 2,090 bps in length (chr5: 150,379,194-150,381,283) containing a 5' upstream sequence, transcription initiation site, first exon of the coding sequence and part of the first intron sequence for GPX3 was used to predict and build CpG island map, oligonucleotide primers for bisulfite sequencing in MethPrimer (8). Independent confirmation of the 2,090-bp sequence for its sequence and identification of transcription start site was made by comparing the Ensemble sequence with GPX3 promoter sequence (5'-cttgaaaggtggc tgggagcgccggacacctcagacggacggtggccagGGATCAGGCAG; capital letters indicate the transcription start site) from the eukaryotic promoter database (EPD) (9). For demonstration of CpG methylation in GPX3 promoter sequence, 500 ng of genomic DNAs were used for C→T(U) conversion following a suggested protocol in EZ DNA Methylation™ Kit (Zymo Research, Orange, CA) and were used as a template for amplification of a 535-bp fragment in a PCR reaction containing a pair of GPX3 bisulfite sequencing primers (5'-GTATTTTG GAGTTAAAAGAGGAAGG-3' and 5'-ACAAAAAAAATT CAAAAAATCTCTC-3') designed in Meth Primer (8). PCR products were cloned into pGEM-T easy TA cloning vector (Promega, Madison, WI), amplified in JM109, and the purified plasmid DNAs were sequenced using either SP6 or T7 primer.

Immunohistostaining of GPX3 on tissue microarray. Immunostaining of two tissue microarrays, each containing 59 normal or 59 matched cancerous stomach tissues, respectively, was performed at SuperBioChips (Seoul, Korea) using rabbit polyclonal antibody to glutathione peroxidase 3 (Novus Biologicals, Littleton, CO). Detailed clinical information is shown in Table II. Following deparaffinization and hydration, antigen retrieval was performed by heating in a microwave (≤700 W) for 15 min (3 times, 5 min each) in 0.01 M citrate buffer (pH 6.0). Endogenous peroxidase activity was blocked by 3% hydrogen peroxide, followed by non-specific protein blocking. Then slides were treated with primary antibody (1:750) for 1 h. Biotinylated anti-rabbit IgG was used as a secondary antibody and was incubated for 30 min. For detection, VectaStain Elite ABC alkaline phosphatase (Vector laboratories, Burlingame, CA) was added for 30-min incubation, followed by treatment with DAB substrate for 2 min. A counterstaining was performed using Mayer's hematoxylin. Staining intensity was scored according to weighted-histoscore method (10); tumor cell intensity was divided into 0, 1, 2 and 3. Histoscores were calculated from the sum of (1 x % of cells staining weakly positive) + (2 x % of cells staining moderately positive) + (3 x % of cells staining strongly positive), with a maximum of 300. Histoscores were categorized into a 4-point system; negative (histoscore = 0),

Table II. Clinical information for cancer samples on tissue microarray.

Category	Class	p-value[a]	p-value[b]	p-value[c]
	Non-tumor (n=59) Gastric cancer (n=58)	0.006	-	-
Age min = 39 max = 72 median = 60	<60 (n=27) ≥60 (n=31)	0.150	0.735	0.295
Sex	female (n=19) male (n=39)	0.300	0.647	0.804
Differentiation[d]	moderate (n=11) poor (n=23) well (n=11)	0.045	0.074	0.068
Lauren	diffuse (n=33) intestinal (n=25)	0.480	0.064	0.414
Stage	I (n=24) II (n=13) III (n=10) IV (n=11)	0.017	0.022	<0.000
GPX3 expression	negative, weak (n=33) moderate, strong (n=25)	-	-	0.001

[a]p-values refer to the significance of differential GPX3 expressions calculated from immunostaining histoscores. For a two-group test, a t-test was used and for 3 or more-group comparison, ANOVA was employed. [b]p-values from the Chi-square test denote the statistical significance of GPX3 expression between clinical classes in each category. [c]p-values from Cox regression analysis denote the statistical significance of survival between clinical classes in each category. Classes in each category are compared with its relation to survival. [d]Differentiation is shown for 45 tubular adenocarcinoma tissues. Of the 59 cancer tissues, the others included signet ring cell carcinoma (n=8), mucinous adenocarcinoma (n=4), papillary adenocarcinoma (n=1) and undifferentiated carcinoma (n=1).

weak (1-100), moderate (101-200) and strong (201-300) staining. Inter-class correlation coefficients were calculated to confirm consistency between two observations. Of 59 sets of samples stained in two slides, 58 cancers and 59 normal samples produced successful staining. For statistical comparison of GPX3 protein expression, a t-test or a one way analysis of variance (ANOVA) was employed. For comparison of survival of patients based on GPX3 expression, patients were grouped into 'negative-weak' or 'moderate-strong' staining groups. Survival curves were compared using log-rank (Mantel-Cox) test. The result is presented as % of patient survivals in the form of Kaplan-Meier survival curve. Chi-square was applied to test the relationship between GPX3 expression and other clinical parameters, including tumor stages, differentiation and Lauren classification. SPSS 13.0 (SPSS, Chicago, IL) was used for the statistical analyses. Multivariate analyses were performed using a Cox regression model to identify independent disease-free survival factors. $p<0.05$ was considered significant.

Results

Outline of gene selection. The overall approach that led to the selection of 8 genes differentially expressed in gastric cancer is schematically shown in Fig. 1 and can be explained in two parts. The first half depicts the microarray-based genomic approach, from which the selection of 22 genes had been reported (3,4,11). The second half summarizes the result shown in the current report, in which we employed *in silico* analysis to reduce the number of genes from 22 to 8.

Validation of gastric cancer biomarker selection in an independent public microarray dataset. When the 22 genes were imported into the public Affymetrix gene expression-profiling dataset containing 22 primary human advanced gastric cancer and 8 non-cancerous gastric tissues (7), 11 probes for 10 genes were commonly present. Of these, 8 genes showed a significant difference in expression levels between the 2 groups in a t-test at a significance level of $p<0.001$ (Table III). The relative expression levels between cancer and non-cancerous tissues are well-maintained in both of the cDNA microarray and the Affymetrix microarray dataset (Fig. 2A). In the following hierarchical clustering, cancer and non-cancerous tissues are clearly distinguished, except that one of the cancer samples, GSM51781, is tightly clustered with normal samples (Fig. 2B). To examine this irregular clustering of one sample in a statistically more stringent manner, the class prediction algorithm was employed as was done previously (3). The

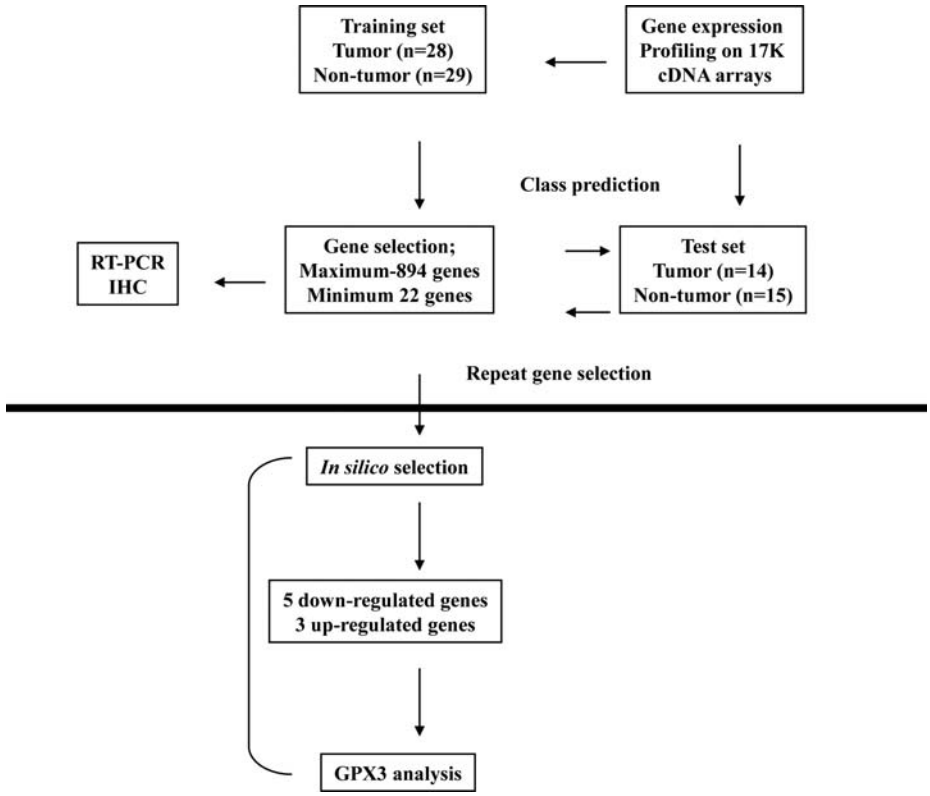

Figure 1. Gastric cancer biomarker selection scheme. A horizontal gray line distinguishes previous gene expression profiling of gastric cancer at the top and the current report. Gene expression profiling analyses of gastric cancer, including gene selection, cross-validation in a training set and class prediction of the test set, which led to the selection of a minimum set of 22 genes that can distinguish tumor and non-tumor were reported before (3,4). In the current report, an *in silico* selection, in which an effort was made to decrease the gene number to a smaller size was made by applying the 22 genes to a public gastric cancer microarray dataset performed on Affymetrix format (7) (details in text and in Fig. 2).

Table III. Genes showing significant difference in expression levels between normal gastric tissues and cancer tissues in an Affymetrix gene chip dataset.

Symbol	Common name	Cancer	Normal	F.C.[a]	p-value
GPX3	Glutathione peroxidase 3	98.1	309.4	-3.2	2.0E-07
SST	Somatostatin	18.6	272.2	-14.3	1.4E-05
CHGA	Chromogranin A	40.1	305.4	-7.7	<1E-07
KLF4	Krüppel-like factor 4	18.7	130.7	-7.15	1.0E-06
LIPF	Lipase, gastric	99.6	4800.3	-50.0	1.0E-06
SERPINH1	Serpin peptidase inhibitor, clade H , member 1	468.8	204.3	+2.2	6.8E-05
THY1	Thy-1 cell surface antigen	145.2	78.7	+1.8	0.00031
INHBA	Inhibin, ß A	79.7	25.1	+3.1	0.00033

[a]F.C. refers to a fold-change. Symbols '-' and '+' refer to down- and up-regulation in cancer, respectively.

same incorrectly clustered cancer sample, GSM51781, was predicted to be non-cancerous tissue in the cross-validation, when the compound covariate predictor, diagonal linear discriminant analysis or support vector machines were utilized in the BRB ArrayTool (data not shown). Therefore, the 8 genes show at least 96% efficiency in correctly predicting the cancer versus non-cancerous gastric tissues. The 8 genes are composed of five down-regulated genes (GPX3, SST, CHGA, KLF4 and LIPF) and three up-regulated genes (SERPINH1, THY1 and INHBA) in gastric cancer. Of these,

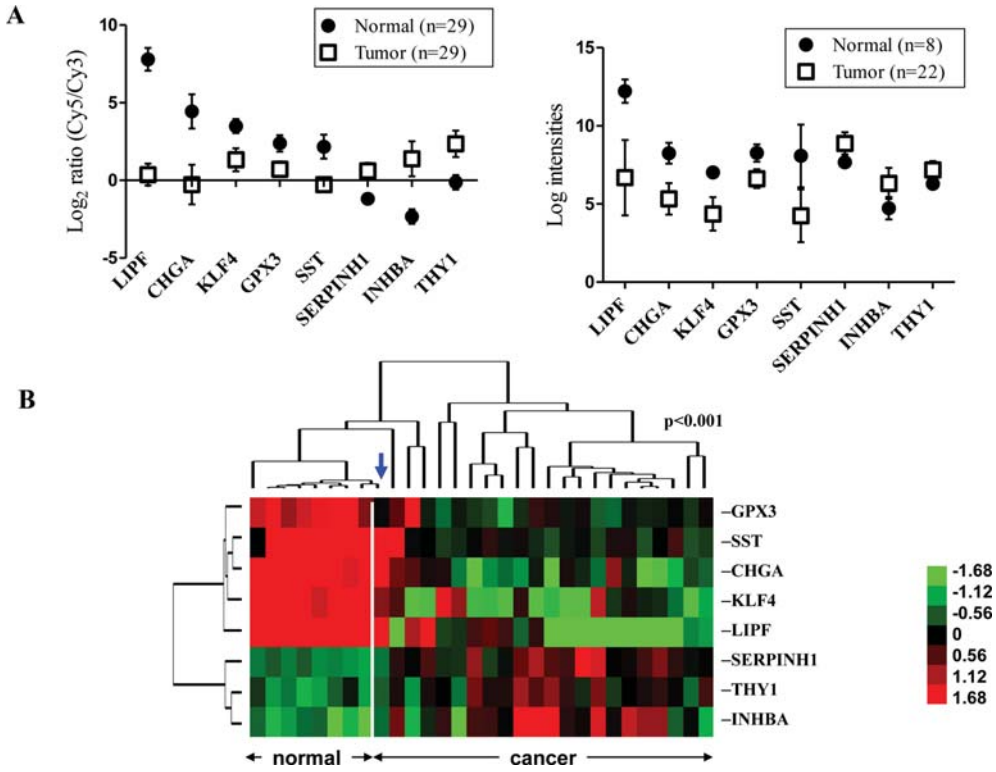

Figure 2. Expression levels of 8 genes in gastric cancer patients. (A) Gene expression levels in 58 gastric tissues (3,4), from which the 8 genes are selected and another set of 30 tissues from public microarray data (7) are shown as average gene expression ratios or fluorescence, respectively. SD is shown as a bar on each box or circle in the plot. (B) A hierarchical clustering of the public gastric cancer data (7) based on the expression level of 8 genes. Non-tumor and tumor samples are represented in red and blue, respectively. A tumor sample closely clustered to the normal samples is shown with an arrow at the top. A scale bar indicating color change and expression log ratio is on the right.

the protein products of GPX3, SST, CHGA, LIPF and INHBA are secretory. Transcriptional activator KLF4 is located in nucleus, whereas, SERPINH1 resides in endoplasmic reticulum and THY1 is connected to the cell membrane through a glycosylphosphatidylinositol (GPI) linker.

KLF4, GPX3, SST and LIPF are down-regulated due to promoter methylation in gastric cancer cells. We tested if any of the five repressed genes was showing its pattern of down-regulation due to gene promoter methylation, one of the key epigenetic mechanisms of gene suppression in tumors (12). Five gastric cancer cell lines were subjected to the treatment with a demethylating agent 5-aza-dC and the isolated total RNAs (Fig. 3A) were used for the gene-specific RT-PCRs. Of the five repressed genes in gastric cancer, four (KLF4, GPX3, SST, LIPF) showed increased expression following 5-aza-dC treatment in at least one of the gastric cancer cell lines, implying that the promoter regions of these four genes are likely to be methylated *in vivo*, leading to the gene repression (Fig. 3B).

Down-regulation of GPX3 in multiple types of cancers. We tested how the gene expression pattern observed in the current report is related to more wide variety of cancers, using GPX3 as an example. We downloaded another public microarray dataset by Su *et al*, which is a profiling of 174 tissues representing 10 different cancers (13), and surveyed the expression pattern of GPX3 (Fig. 4A). GPX3 is highly and consistently down-regulated in bladder, breast, colorectal, gastro-esophagus and prostate cancers. In lung and ovarian cancers, GPX3 shows a gradient of expression levels over the tissue samples, implying high intra-cancer heterogeneity in expression level. GPX3 is most highly expressed in carcinomas of kidney and liver, in which still a high heterogeneity is observed. Next, we examined the GPX3 expression patterns in various microarray datasets containing cancer and healthy controls. Specifically, GPX3 is down-regulated in cervical, thyroid, head and neck and lung cancers compared with corresponding non-cancerous tissues (Fig. 4B). In melanoma, GPX3 shows significant changes in expression pattern from normal to benign nevi, where GPX3 is up-regulated in benign nevi. GPX3 is then down-regulated again in the transition from benign nevi to malignant melanoma.

Tissue microarray analysis of GPX3 expression in gastric cancer. To examine the down-regulation of GPX3 at a protein level in clinical samples, immunohistostainings of a pair of tissue microarrays containing gastric cancer tissues and corresponding non-cancerous gastric tissues were performed (Fig. 5A). GPX3 antibody staining intensities were converted into histoscores (Materials and methods). In a t-test of histoscores, GPX3 showed a significant repression in cancer (n=58) than in non-cancerous (n=59) tissues (p=0.006, Fig. 5B). When the cancers were grouped based on GPX3 antibody staining intensities, there were significant differences in patient survival based on GPX3 expression; relatively high level ('moderate' and 'strong' combined, n=25) GPX3 expression led to longer survival of the patients than the low

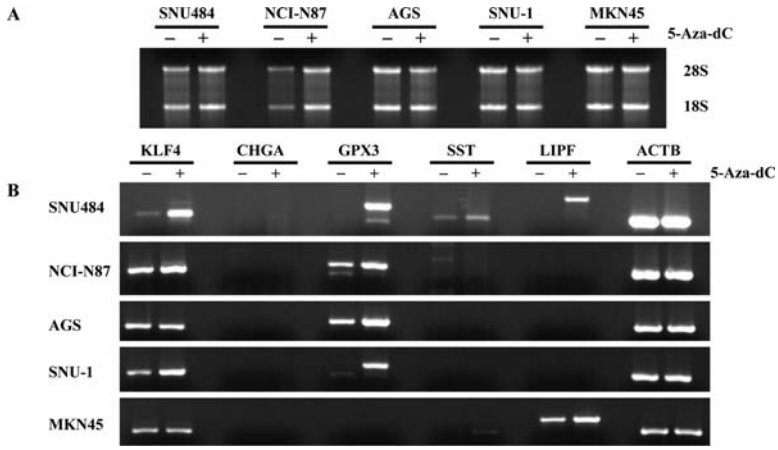

Figure 3. Elevated expression of KLF4, GPX3, SST and LIPF after 5-aza-dC treatment in gastric cancer cells. (A) To test the possibility of the gene repression by methylation, total RNAs were purified before and after 5-aza-dC treatment to five gastric cancer cells. (B) Gene-specific PCR primers were designed to amplify the region of five genes. The primers (Table III) were designed to amplify the PCR products of 600 base pairs in length.

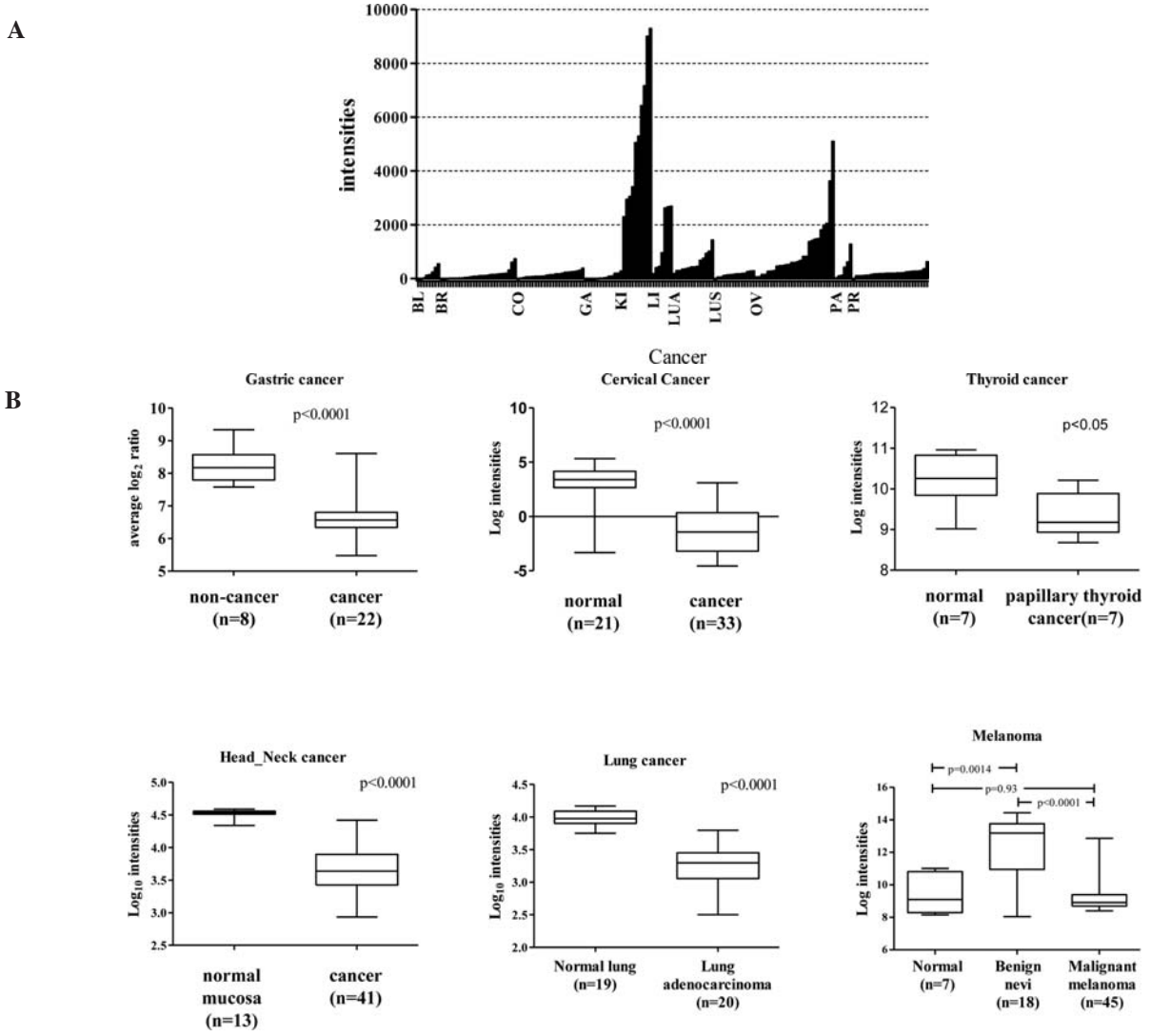

Figure 4. Expression pattern of GPX3 in cancers. (A) To examine the relative expression level of GPX3 in different cancers, molecular classification of a human cancer dataset containing BL (bladder/ureter), BR (breast), CO (colon), GA (gastroesophagus), KI (kidney), LI (liver), LUA (lung adenocarcinomas), LUS (lung squamous cell carcinoma), OV (ovary), PA (pancreas) and PR (prostate) (13). Expression levels are shown as probe intensities from Affymetrix U95a GeneChip. (B) Expression levels in cancer and corresponding normal tissues from microarray studies are shown for gastric (7), cervical (33), thyroid (dataset record GDS1732 from GEO, gene expression omnibus, N.I.H.), head and neck (34), lung cancer (35) and melanoma (36) are shown in relation to the corresponding non-tumor tissues. Group labels (non-tumor, normal, tumor, lung adenocarcinomas) used in the original articles or datasets are used in this figure without changes.

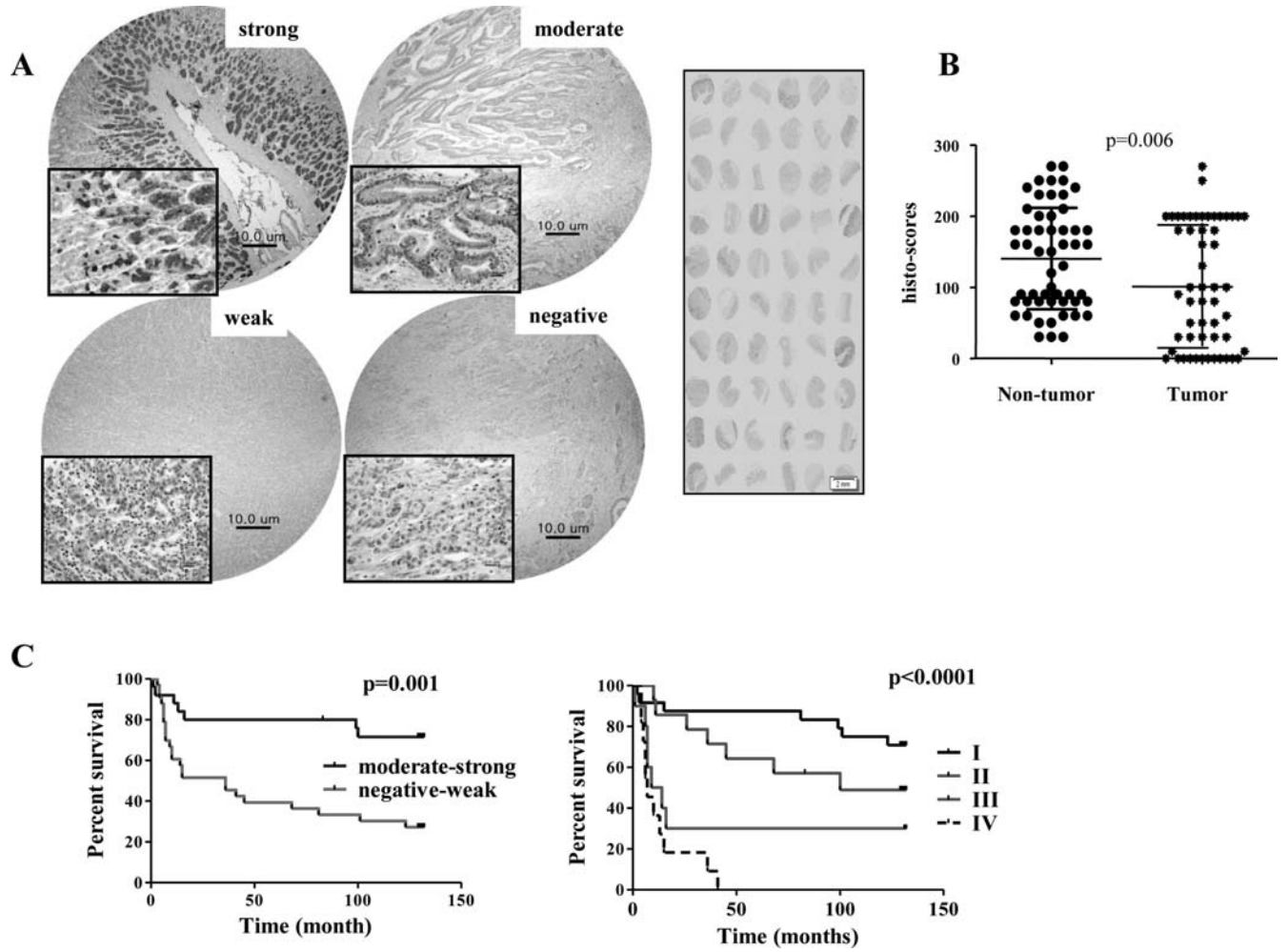

Figure 5. GPX3 protein expression analysis on tissue microarrays. (A) Representative images of GPX3 antibody staining for strong to negative signals. The overall picture of a tissue microarray is shown on the right in a box. (B) A box plot for histoscores in gastric cancer and non-tumor gastric tissues. (C) Kaplan-Meier survival curves for patients grouped based on GPX3 expression and tumor stages (left). As a reference, survival curves for patients grouped by tumor stages and Lauren classification are also shown.

level ('weak' and 'negative' combined, n=33) (Fig. 5C). As a control, tumor stage also had a significant effect on patient survival, whereas other clinical parameters including Lauren classification were not major prognostic factors (Table II). Chi-square test showed that there is a significant dependency between tumor stage and high GPX3 expression (p=0.022), implying that as gastric cancer progresses, there is a concomitant reduction in GPX3 expression.

GPX3 promoter methylation and down-regulation in gastric cancer cells. To test GPX3 promoter methylation in detail, GPX3 expressions were measured in additional gastric cancer cells. In all, GPX3 showed significant increases in mRNA expression after 5-aza-dC treatment in 9 cancer cells, implying that the GPX3 is generally under epigenetic regulation in gastric cancer cells (Fig. 6A). Next, to directly confirm the promoter methylation, bisulfite sequencing was employed for AGS, MKN74, MKN-45, SNU-1 and SNU-719 cells (Fig. 6B and Materials and methods). In all cases, majority of the CpGs in DNA sequences extending from around the transcription site to the first exon, where the majority of the CpG sites reside, are highly methylated.

Discussion

The development of diagnostic tools, based on the molecular markers, which are less invasive and less expensive, has the potential to aid in the early detection of gastric cancer, together with the currently used endoscopy, therefore, leading to a better chance for appropriate treatment (14). In addition, these bio-markers serve as critical tools in understanding carcinogenesis at a molecular level.

The current report is a continuation of the genomic analyses of gastric cancer and we describe the selection of 8 genes showing significantly altered gene expression pattern in gastric cancer compared with non-cancerous gastric tissues. Collectively, the 8-gene signature can correctly predict the cancer and non-cancerous gastric tissues in the cross-validation in a cDNA microarray dataset and was able to correctly predict 29 of 30 cases in an independent Affymetrix dataset (Fig. 2). We did not apply the class prediction approach to develop the 8-gene signature as a diagnostic measure per se. Rather it was used as a method for reliable gene selection. In terms of molecular functions, the 8 genes represent a response to oxidative stress (GPX3),

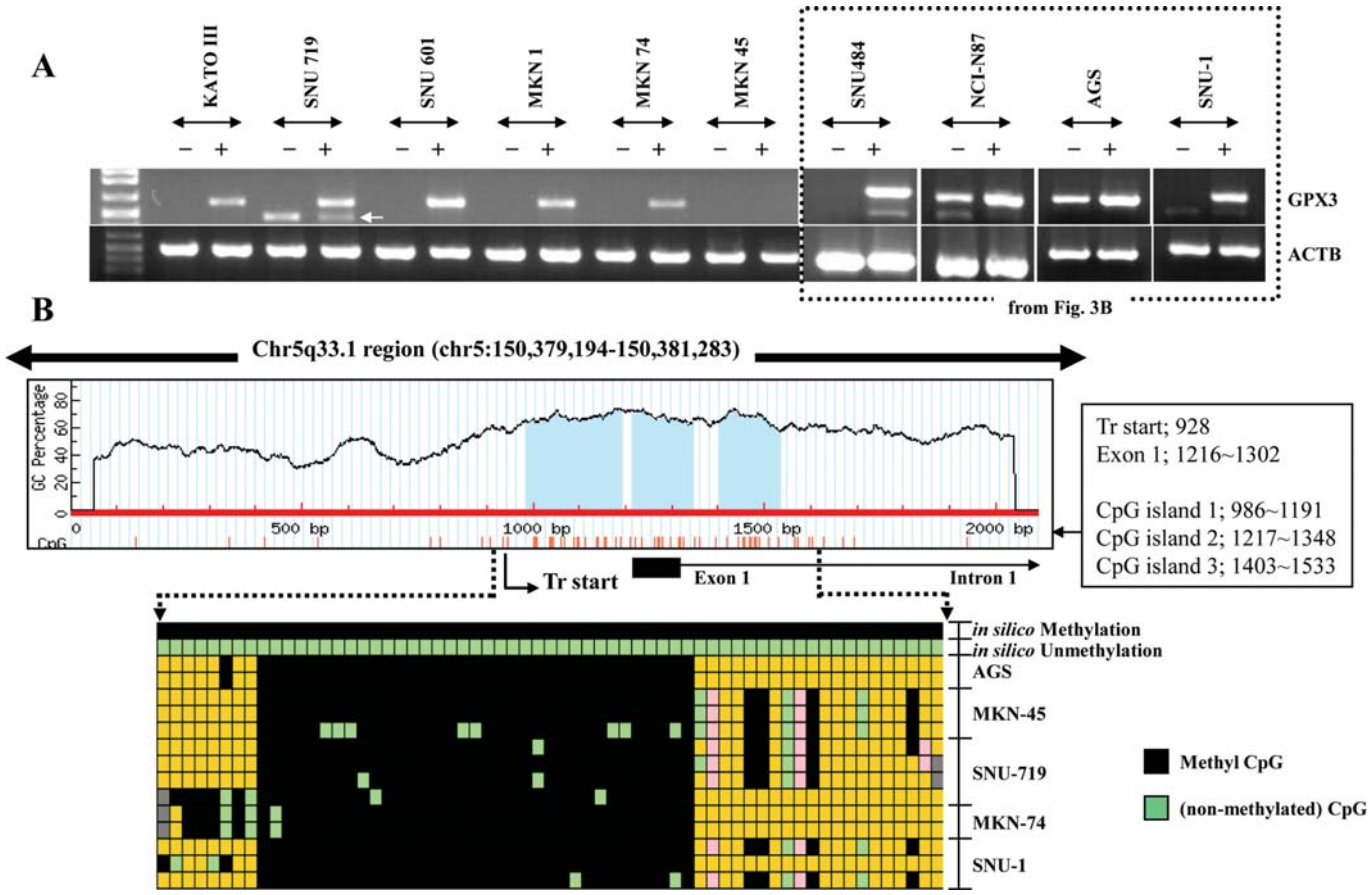

Figure 6. GPX3 repression by promoter methylation. (A) RT-PCR of GPX3 before and after 5-aza-dC treatment for gastric cancer cells. Arrow under GPX3 band in SNU719 cells indicates non-specific PCR products. PCR products for KATOIII, SNU719, SNU601, MKN1 and MKN74 were resolved on a 1% agarose gel. GPX3 RT-PCRs from SNU484, NCI-N87, AGS and SNU-1 in a dotted box are borrowed from Fig. 3B and shown here again for a parallel comparison. (B) Three predicted CpG islands from GPX3 promoter (Materials and methods) are shown in blue color in the GPX3 genomic DNA sequence map. Transcription start site ('Tr start') and the first exon for GPX3 are also shown. Detailed information covering the CpG islands is shown in the box at right. The DNA sequence of 2,090 bp in length shown in horizontal red bar corresponds to Chr5q33.1 region (chr5: 150,379,194-150,381,283, the number representing the bases from p term of chromosome 5), surrounding the upstream and part of exon 1 and intron 1 of GPX3 are analyzed in detail in this figure and in (C). The chromosomal sequence (chr5: 150,379,194-150,381,283) is the actual chromosomal location for GPX3 gene retrieved from Ensembl. (C) Bisulfite sequencing and data analysis. C→T-conversed sequences were used as input to map the CpG methylation pattern in CpGviewer (37). The 2090-bp fragment used to predict and map the promoter sequence and its artificially C→T-conversed sequences were used as *in silico*-methylated (black) and -unmethylated (light green) reference sequences, respectively. Independent *E. coli* clones were picked for plasmid DNA isolation for DNA sequencing and analyses for AGS, SNU719, MKN-74, SNU-1 gastric cancer cells. The dark box represents methylated CpG, which therefore did not undergo C→T conversion by bisulfite treatment, whereas unmethylated CpG is transformed in TpG by C→T conversion by bisulfite.

regulation of cell proliferation, migration or differentiation (SST, KLF4), lipid metabolic process (LIPF), response to heat shock, binding to unfolded protein (SERPINH1), Rho GTPase activator activity/angiogenesis/cell-cell adhesion, migration (THY1) and negative regulation of cell cycle and growth (INHBA).

Pubmed and public microarray database searches confirm the altered expression for these genes in several cancer models, for example methylation and down-regulation of GPX3 (14-17), SST (18-21) and KLF4 (22-25) in cancer compared with corresponding non-cancerous counterparts. Likewise, the up-regulation of THY1 in colon cancer microarray study together with real-time reverse transcription-PCR validation has been reported before (26), but there is also a report on the down-regulation of THY1 in metastatic nasopharyngeal carcinoma (27). Up-regulations of INHBA in pancreatic cancer (28) and oral tongue squamous cell carcinoma (OTSCC) (29) and of SERPINH1 in ductal adenocarcinomas (30,31) compared with corresponding matching non-tumor tissues are also in line with the expression pattern of these genes in gastric cancer as is shown in this report.

The mechanisms underlying the differential expression and signal transductions following the expressions for these genes in gastric cancer have not yet been intensively studied. As an example of a detailed study of one of the selected genes, GPX3 down-regulation due to promoter methylation is probably the first report in gastric cancer, following similar results in prostate cancer (21) and in Barrett's-related adenocarcinomas (32). We have shown through an immunohistostaining study of tissue microarrays that GPX3 is expressed at a significantly lower protein level in normal gastric tissues than in gastric cancer. Yet, its wide variations in expression within cancers can lead to the difference in survival rate of cancer patients. GPX3 expression in gastric

cancer seems not to be dependent on any of the clinicopathological parameters, including tumor stage and Lauren's classification. Rather it is only distinctly repressed in tumor in general than in non-tumor gastric tissues. This is expected since, in our case it was selected from the comparison of non-cancerous gastric tissues and heterogeneous (tumor stage, intestinal, diffuse types combined) gastric cancer tissues. It is comparable to the case in prostate cancer, where the difference in survival of the patients based on GPX3 expression was not associated with Gleason's grade, but with the relapse of prostate cancer (14). We have shown that GPX3 is down-regulated in gastric cancer primarily due to promoter methylation. Overexpression of GPX3 and identification of its transcriptional targets by genomic and proteomic approach in gastric cancer cells are underway.

In conclusion, eight genes showing statistically significant differences in expression level between gastric cancer tissues and non-cancerous gastric tissues and also possessing a predictive ability for the two tissue types are reported. Individually, they can serve as molecular tools in studying gastric cancer.

Acknowledgements

This study was supported by the Faculty Research Grant of Yonsei University College of Medicine for 2008 (No. 6-2008-0169). This work was also supported by National Research Foundation of Korea Grant funded by the Korean Government (2009_0076922).

References

1. Parkin D M, Bray F and Ferlay J and Pisani P: Global cancer statistics, 2002. CA Cancer J Clin 55: 74-108, 2005.
2. Yang S: Gene amplifications at chromosome 7 of the human gastric cancer genome. Int J Mol Med 20: 225-231, 2007.
3. Yang S and Chung H: Novel biomarker candidates for gastric cancer. Oncol Rep 19: 675-680, 2008.
4. Yang S, Shin J, Park KH, Jeung H-C, Rha SY, Noh SH, Yang WI and Chung HC: Molecular basis of the differences between normal and tumor tissues of gastric cancer. Biochim Biophys Acta 1772: 1033-1040, 2007.
5. Yang SH, Seo MY, Jeong HJ, Jeung H-C, Shin J, Kim SC, Noh SH, Chung H and C and Rha SY: Gene copy number change events at chromosome 20 and their association with recurrence in gastric cancer patients. Clin Cancer Res 11: 612-620, 2005.
6. Dupuy A and Simon RM: Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. J Natl Cancer Inst 99: 147-157, 2007.
7. Hippo Y, Taniguchi H, Tsutsumi S, Machida N, Chong J-M, Fukayama M and Kodama T and Aburatani H: Global gene expression analysis of gastric cancer by oligonucleotide microarrays. Cancer Res 62: 233-240, 2002.
8. Li L-C and Dahiya R: Methprimer: designing primers for methylation pcrs. Bioinformatics 18: 1427-1431, 2002.
9. Schmid CD, Perier R and Praz V and Bucher P: EPD in its twentieth year: towards complete promoter coverage of selected model organisms. Nucl Acids Res 34: D82-D85, 2006.
10. Witton C, Hawe S and Cooke T and Bartlett J: Cyclooxygenase 2 (Cox2) expression is associated with poor outcome in ER-negative, but not ER-positive, breast cancer. Histopathology 45: 47-54, 2004.
11. Yang S, Jeung H-C, Jeong HJ, Choi YH, Kim JE, Jung J-J, Rha SY, Yang WI and Chung HC: Identification of genes with correlated patterns of variations in DNA copy number and gene expression level in gastric cancer. Genomics 89: 451-459, 2007.
12. Esteller M: Epigenetics in cancer. N Engl J Med 358: 1148-1159, 2008.
13. Su AI, Welsh J B, Sapinoso LM, Kern SG, Dimitrov P, Lapp H, Schultz PG, Powell SM, Moskaluk CA, Frierson HF Jr and Hampton GM: Molecular classification of human carcinomas by use of gene expression signatures. Cancer Res 61: 7388-7393, 2001.
14. Yu YP, Yu G, Tseng G, Cieply K, Nelson J, Defrances M, Zarnegar R and Michalopoulos G and Luo J-H: Glutathione peroxidase 3, deleted or methylated in prostate cancer, suppresses prostate cancer growth and metastasis. Cancer Res 67: 8043-8050, 2007.
15. Lee O-J, Schneider-Stockz R, McChesney PA, Kuester D, Albert Roessnerz, Vieth M, Moskalukb CA and El-Rifai W: Hypermethylation and loss of expression of glutathione peroxidase-3 in Barrett's tumorigenesis. Neoplasia 7: 854-861, 2005.
16. Lodygin D, Epanchintsev A, Menssen A, Diebold J and Hermeking H: Functional epigenomics identifies genes frequently silenced in prostate cancer. Cancer Res 65: 4218-4227, 2005.
17. Schmutzler C, Mentrup B, Schomburg L, Hoang-Vu C, Herzog V and Köhrle J: Selenoproteins of the thyroid gland: expression, localization and possible function of glutathione peroxidase 3. Biol Chem 388: 1053-1059, 2007.
18. Mori Y, Cai K, Cheng Y, Wang S, Paun B, Hamilton JP, Jin Z, Sato F, Berki AT, Kan T, Ito T, Mantzur C, Abraham JM and Meltzer SJ: A genome-wide search identifies epigenetic silencing of somatostatin, tachykinin-1, and 5 other genes in colon cancer. Gastroenterology 131: 797-808, 2006.
19. Colucci R, Blandizzi C, Ghisu N, Florio T and Del Tacca M: Somatostatin inhibits colon cancer cell growth through cyclooxygenase-2 downregulation. Br J Pharmacol 155: 198-209, 2008.
20. Jin Z, Mori Y, Hamilton J, Olaru A, Sato F, Yang J, Ito T, Kan T, Agarwal R and Meltzer S: Hypermethylation of the somatostatin promoter is a common, early event in human esophageal carcinogenesis. Cancer 112: 43-49, 2008.
21. Ongenaert M, Wisman G, Volders H, Koning A, van der Zee A and van Criekinge W and Schuuring E: Discovery of DNA methylation markers in cervical cancer using relaxation ranking. BMC Med Genomics 1: 57, 2008.
22. Wei D, Gong W, Kanai M, Schlunk C, Wang L, Yao JC, Wu T-T, Huang S and Xie K: Drastic down-regulation of Krüppel-like factor 4 expression is critical in human gastric cancer development and progression. Cancer Res 65: 2746-2754, 2005.
23. Wei D, Kanai M, Huang S and Xie K: Emerging role of KLF4 in human gastrointestinal cancer. Carcinogenesis 27: 23-31, 2006.
24. Yasunaga J-I, Taniguchi Y, Nosaka K, Yoshida M, Satou Y, Sakai T, Mitsuya H and Matsuoka M: Identification of aberrantly methylated genes in association with adult T-cell leukemia. Cancer Res 64: 6002-6009, 2004.
25. Zhao W, Hisamuddin I, Nandan M, Babbin B, Lamb N and Yang V: Identification of krüppel-like factor 4 as a potential tumor suppressor gene in colorectal cancer. Oncogene 23: 395-402, 2004.
26. Chiu S-T, Hsieh F-J, Chen S-W, Chen C-L, Shu H-F and Li H: Clinicopathologic correlation of up-regulated genes identified using cdna microarray and real-time reverse transcription-PCR in human colorectal cancer. Cancer Epidemiol Biomarkers Prev 14: 437-443, 2005.
27. Lung HL, Bangarusamy DK, Xie D, Cheung AKL, Cheng Y, Kumaran MK, Miller L, Liu ET-B, Guan X-Y, Sham JS, Fang Y, Li L, Wang N, Protopopov AI, Zabarovsky ER, Tsao SW, Stanbridge EJ and Lung ML: Thy1 is a candidate tumour suppressor gene with decreased expression in metastatic nasopharyngeal carcinoma. Oncogene 24: 6525-6532, 2005.
28. Cao D, Hustinx SR, Sui G, Bala P, Sato N, Martin S, Maitra A, Murphy KM, Cameron JL, Yeo CJ, Kern SE, Goggins M, Pandey A and Hruban RH: Identification of novel highly expressed genes in pancreatic ductal adenocarcinomas through a bioinformatics analysis of expressed sequence tags. Cancer Biol Ther 3: 1081-1089, 2004.
29. Ye H, Yu T, Temam S, Ziober B, Wang J, Schwartz J, Mao L, Wong D and Zhou X: Transcriptomic dissection of tongue squamous cell carcinoma. BMC Genomics 9: 69, 2008.
30. Iacobuzio-Donahue C, Maitra A, Shen-Ong G L, van Heek T, Ashfaq R, Meyer R, Walter K, Berg K, Hollingsworth MA, Cameron JL, Yeo CJ, Kern SE, Goggins M and Hruban RH: Discovery of novel tumor markers of pancreatic cancer using global gene expression technology. Am J Pathol 160: 1239-1249, 2002.

31. Maitra A, Iacobuzio-Donahue C, Rahman A, Sohn T, Argani P, Meyer R, Yeo C, Cameron J, Goggins M, Kern S, Ashfaq R, Hruban R and Wilentz R: Immunohistochemical validation of a novel epithelial and a novel stromal marker of pancreatic ductal adenocarcinoma identified by global expression microarrays: Sea urchin fascin homolog and heat shock protein 47. Am J Clin Pathol 118: 52-59, 2002.
32. Peng DF, Razvi M, Chen H, Washington K, Roessner A, Schneider-Stock R and El-Rifai W: DNA hypermethylation regulates the expression of members of the mu-class glutathione S-transferases and glutathione peroxidases in Barrett's adenocarcinoma. Gut 58: 5-15, 2009.
33. Scotto L, Narayan G, Nandula S, Arias-Pulido H, Subramaniyam S, Schneider A, Kaufmann A, Wright J, Pothuri B and Mansukhani M: Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: Potential role in progression. Genes Chromosomes Cancer 47: 755-765, 2008.
34. Ginos MA, Page GP, Michalowicz BS, Patel KJ, Volker SE, Pambuccian SE, Ondrey FG, Adams GL and Gaffney PM: Identification of a gene expression signature associated with recurrent disease in squamous cell carcinoma of the head and neck. Cancer Res 64: 55-63, 2004.
35. Stearman R, Dwyer-Nield L, Zerbe L, Blaine S, Chan Z, Bunn P, Johnson G, Hirsch F, Merrick D and Franklin W: Analysis of orthologous gene expression between human pulmonary adenocarcinoma and a carcinogen-induced murine model. Am J Pathol 167: 1763-1775, 2005.
36. Talantov D, Mazumder A, Yu J, Briggs T, Jiang Y, Backus J, Atkins D and Wang Y: Novel genes associated with malignant melanoma but not benign melanocytic lesions. Clin Cancer Res 11: 7234-7242, 2005.
37. Carr IM, Valleley EMA, Cordery SF, Markham AF and Bonthron DT: Sequence analysis and editing for bisulphite genomic sequencing projects. Nucl Acids Res 35: e79, 2007.
38. Rozen S and Skaletsky H: Primer3 on the www for general users and for biologist programmers. Methods Mol Biol 132: 365-386, 2000.