Published in final edited form as:
Clin Cancer Res. 2014 September 1; 20(17): 4598–4612. doi:10.1158/1078-0432.CCR-13-3380.

Large-scale characterization of DNA methylation changes in human gastric carcinomas with and without metastasis

Zhaojun Liu[1,†], **Jun Zhang**[1,2,†], **Yanhong Gao**[1,†], **Lirong Pei**[3], **Jing Zhou**[1], **Liankun Gu**[1], **Lianhai Zhang**[4], **Budong Zhu**[5], **Naoko Hattori**[6], **Jiafu Ji**[4], **Yasuhito Yuasa**[7], **Wooho Kim**[8], **Toshikazu Ushijima**[6], **Huidong Shi**[3], and **Dajun Deng**[1]

[1]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Division of Etiology, Peking University Cancer Hospital and Institute, Fu-Cheng-Lu, No.52, Haidian District, Beijing, 100142, China

[4]Department of Surgery, Peking University Cancer Hospital and Institute, Fu-Cheng-Lu, No.52, Haidian District, Beijing, 100142, China

[5]Department of Oncology, Peking University Cancer Hospital and Institute, Fu-Cheng-Lu, No.52, Haidian District, Beijing, 100142, China

[2]Shihezi University School of Medicine, Shihezi, 832000, China

[3]GRU Cancer Center, Georgia Regents University, 1411 Laney Walker Blvd, Augusta, GA 30912, USA

[6]Division of Epigenetics, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan

[7]Department of Molecular Oncology, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan

[8]Department of Pathology, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul, 110-799, Korea

Abstract

Purpose—Metastasis is the leading cause of death for gastric carcinoma (GC). An epigenetic biomarker panel for predicting GC metastasis could have significant clinical impact on the care of GC patients. The main purpose of this study is to characterize the methylation differences between GCs with and without metastasis.

Experimental Design—Genome-wide DNA methylation profiles between 4 metastatic and 4 non-metastatic GCs and their surgical margins (SM) were analyzed using methylated-CpG island

Correspondence: Dajun Deng, M.D., Peking University Cancer Hospital and Institute, 100142, China. Phone: +8610-88196752; Fax: +8610-88122437; or Huidong Shi, Ph.D., GRU Cancer Center, Georgia Regents University, Augusta, GA 30912, USA; hshi@gru.edu or dengdajun@bjmu.edu.cn.
[†]Equal contribution

Disclosure of potential conflicts of interest: The authors have declared that no competing interests exist.

Data and materials availability
The methylation array data has been deposited into the Gene Expression Omnibus under accession number GSE47724. Data files S1–S3 are available upon request.

amplification with microarray. The methylation states of 73 candidate genes were further analyzed in GC patients in a discovery cohort (*n*=108) using DHPLC, bisulfite-sequencing, and MethyLight. The predictive values of potential metastasis-methylation biomarkers were validated in GC patient cohorts in China (*n*=330), Japan (*n*=129), and Korea (*n*=153).

Results—The GC genome showed significantly higher proportions of hypomethylation in the promoter and exon-1 regions, as well as increased hypermethylation of intragenic fragments when compared to SMs. Significant differential methylation was validated in the CGIs of 15 genes (Ps<0.05) and confirmed using bisulfite-sequencing. These genes included *BMP3*, *BNIP3*, *CDKN2A*, *ECEL1*, *ELK1*, *GFRA1*, *HOXD10*, *KCNH1*, *PSMD10*, *PTPRT*, *SIGIRR*, *SRF*, *TBX5*, *TFPI2*, and *ZNF382*. Methylation changes of *GFRA1*, *SRF* and *ZNF382* resulted in up- or down-regulation of their transcription. Most importantly, the prevalence of *GFRA1*, *SRF*, and *ZNF382* methylation alterations was consistently and coordinately associated with GC metastasis and the patients' overall survival throughout discovery and validation cohorts in China, Japan and Korea.

Conclusion—Methylation changes of *GFRA1*, *SRF*, and *ZNF382* may be a potential biomarker set for prediction of GC metastasis.

Keywords

gastric cancer; metastasis; DNA methylation

Introduction

Gastric carcinoma (GC) is the second leading cause of cancer death throughout the world (1). Global statistics showed that in 2008 alone, nearly 989,000 people were diagnosed with GC, and approximately 464,000 people died from this disease (2). Currently, GC prognosis is primarily determined based on the clinical data and pathological stages of patients at the time of diagnosis and treatment (3). However, successful management of GC patients is still hampered by the lack of highly sensitive and specific biomarkers capable of predicting prognosis and likelihood of metastasis. Epigenetic alterations, including aberrant DNA methylation changes, may play an important role in gastric carcinogenesis as indicated by the increased hypermethylation of tumor suppressor genes in GC patients (4–6). Given their important functions in cancer initiation and progression, methylation changes are being investigated as potential biomarkers for the early detection of cancers, the prediction of cancer progression, and the prediction of chemotherapeutic sensitivity (7).

Recent advances in high-throughput technologies have significantly expanded our capability of interrogating genome-wide DNA methylation changes in cancer (6, 8, 9). Methylated-CpG island amplification with microarray (MCAM) is one of the most powerful tools available for displaying differential methylation related to pathogenesis (10). A number of DNA methylome studies have been reported in a variety of primary cancers, including GC. However, few studies have been conducted to vigorously validate the methylation changes of the candidate genes at the single molecule level in numerous tumor samples (6, 11). Therefore, despite the long list of differentially methylated genes in GC patients, a promising DNA methylation biomarker has not yet reached to the clinical utility.

In the present study, genome-wide DNA methylation analysis using the MCAM assay was performed in GCs (10). A large number of differentially methylated regions were identified between GCs and their corresponding surgical margin (SM). In addition, differential methylation profiles between metastatic and non-metastatic GCs were identified. Most importantly, the methylation status of promoter CpG islands (CGIs) from 73 candidate genes was characterized using denatured high performance liquid chromatography (DHPLC) in 48 pairs of gastric samples from GC and non-cancer patients (12). The predictive values of three potential metastasis-related candidates were further validated in multiple cohorts from China, Japan, and Korea following the Reporting Prognostic Tumor Marker Study (REMARK) guidelines. We demonstrated that the methylation status of *GFRA1, SRF* and *ZNF382* could be used as potential synergistic biomarkers for the prediction of GC metastasis.

Materials and Methods

Patient characteristics and sample collection

A total of 504 patients with GC from 3 academic medical centers in China, Japan, and Korea were included in this study. The study was approved by the local Institution Review Boards (IRB) at each institution, and all patients were given written informed consent unless the IRB permitted a waiver. The 2003 UICC-TNM (tumor-node-metastasis) system was used for the classification of GCs (13). 330 Chinese GC inpatients that underwent surgical treatment at Peking University Cancer Hospital & Institute between 1999 and 2006 were enrolled in the discovery and validation cohorts based on the following criteria: a) availability of frozen, fresh GC and SM samples; b) follow-up available for at least 5 years; c) falls into the proper pathological TNM (pTNM) stages as described in the results section. In the validation cohort from Korea, 153 GC inpatients that received surgical treatment were selected from Seoul National University Hospital during 2004 with a follow-up of at least 3 years. Paraffin-embedded SM samples were used in the Korea study. The validation cohort from Japan included 78 GC inpatients that acquired surgical treatment between 1995 and 2002 with a follow-up of at least 5 years, as well as an additional 79 GC patients between 2010 and 2011 who did not have survival data. The SM samples were not available for these Japanese patients. GCs were classified as cardiac or non-cardiac in terms of location (14). Patients with pre-operative chemotherapy were not included in the discovery or independent validation cohorts. Normal/gastritis biopsies (NorG) from 56 outpatients at Peking University Cancer Hospital were used as the cancer-free controls.

Study design

The discovery patient cohort from Peking University Cancer Hospital consisted of 54 randomly selected patients with non-metastatic GCs and 54 matched patients with distant metastatic GCs. Among them, 8 paired GC and the corresponding SM samples from patients with or without distant and lymph metastasis were analyzed using MCAM on a customized Agilent promoter array. The clinical and histological features of these 8 patients can be found in Supplementary Table S1. The remaining GC and SM samples from 100 patients were used for the characterization of 73 CGIs using DHPLC and bisulfite clone sequencing. The methylation states of the three most promising candidate CGIs were analyzed in three

analogous independent validation cohorts from China (*n*=222), Japan (*n*=129), and Korea (*n*=153). The overall study design is outlined in Supplementary Fig. S1. Genomic DNA was isolated using phenol/chloroform extraction.

Cell lines and culture

MKN74 cell line was kindly provided by Dr. Yasuhito Yuasa at Tokyo Medical and Dental University in 2010; RKO cell line, from Dr. Guoren Deng, at University California in San Francisco in 2001; AGS, by Dr. Chengchao Shou in 2009, HeLa and MGC803, by Dr. Yang Ke in 2004, at Peking University Cancer Hospital. All cells were grown in monolayer in appropriate medium supplemented with 10% FBS and maintained at 37°C in humidified air with 5% CO_2. These cell lines were tested and authenticated by Beijing JianLian Genes Technology Co., LTD before they were used in this study. STR patterns were analyzed using Goldeneye™20A STR Identifiler PCR Amplification Kit. Gene Mapper v3.2 software (ABI) was used to match the STR pattern with the online databases of National Platform of Experimental Cell Resources for Sci-Tech for MGC803 cell and the American Type Culture Collection (ATCC) for other cells.

Genome-wide analysis of DNA methylation in GC tissues using MCAM

Genomic DNA (2 μg) from 8 pairs of fresh GC and SM samples was analyzed using the MCAM approach (10). Briefly, genomic DNA was digested consecutively with *SmaI* and *XmaI* which cut unmethylated and methylated CCCGGG sites, respectively. The *XmaI* digestion produces sticky ends that can be ligated to linkers, while *SmaI* digestion results in blunt ends that are unable to be ligated to linkers. The ligation-mediated PCR products from GC and SM samples were purified and labeled with Alexafluor647 or 555, respectively, using the Bioprime Plus Array CGH Indirect Genomic Labeling kit (Invitrogen, Carlsbad, CA) according to the manufacturers' instructions. The labeled DNA was co-hybridized to a custom designed Agilent Oligonucleotide Array, and the slides were washed and scanned as described previously (15). Data was extracted using the Feature Extraction Tool (Agilent Technologies, Santa Clara, CA) and exported for further analysis. The custom designed Agilent oligonucleotide array was designed using Agilent eArray service (https://earray.chem.agilent.com/earray). The array consisted of approximately 99,028 probes (44–60mers) that covered 29,879 *in silico SmaI*-digested DNA fragments (>60bp and <2000bp) in the human genome. The probes were tiled within each fragment with 100bp spacing. The methylation states of 6177 genes were determined using this custom methylation array.

Microarray data normalization and probe/gene selection

The raw array data was processed and normalized by the Beijing CO-FLY Bioinformatic Company. Background model adjustment was carried out using the minimum normalization algorithm. Systematic differences between arrays were normalized using the quantile method as described (16, 17). The methylation array data, as well as the probe information, have been deposited into the Gene Expression Omnibus under accession number GSE47724.

The mean intensity of the normalized array hybridization (methylation) signal of each probe for sex-related chromosomes and autosomes in the SM samples from 4 males and 4 females (Supplementary Fig. S2A–C) were analyzed. As expected, the intensities of 784 of 2390 X

chromosome-probes (32.8%) were significantly higher in the female samples than the male samples (Student's *t*-test, $P<0.05$); in contrast, 35 of 87 Y chromosome-probes (40.2%) were significantly higher in the male samples when compared to the female samples. These sex-specific differences were only observed in 1250 of 96,550 (1.3%) probes in the 22 autosomes. These results confirmed that the quality of the normalized data is sufficient to differentiate sex-specific DNA methylation and suitable for studying GC- or metastasis-related methylation changes.

The methylation signal ratio ([GC]/[SM]) was calculated for each array probe. The Student's paired *t*-test ($P<0.01$) was used to identify the differentially methylated probes between GC and SM samples from the 8 patients analyzed. The Mann-Whitney *U*-test ($P<0.029$) was used to identify the metastasis-specific differentially methylated probes between the 4 metastatic and 4 non-metastatic GC patients. The methylation ratio data including the adjusted *P*-values for each probe are included in the data file S1.

The difference between GC-related hypermethylated and hypomethylated probes was calculated for each sliding window (sequence or region) using 51 probe-matched fragments, which included the target probe along with 25 probes both upstream and downstream of the target. Probes near the centromeres and telomeres of each chromosome were not included due to the absence of the 25 upstream or downstream probes. The numerical differences for 99K probes were charted to display the detailed regional methylation trend (or net methylation signal) for the corresponding chromosome arm.

Identification of differentially methylated candidate genes

In order to identify GC and metastasis-specific differentially methylated candidate genes for further evaluation, the promoter and exon-1 regions were focused on due to their known inverse correlation to epigenetic repression of gene transcription. The differentially methylated probes in these regions were defined as the top-100 probes and used in hierarchical clustering analysis and preparation of a heatmap, when their *P*-values were less than 0.05, and their absolute mean difference-values were within the top 100. Candidate genes were selected from these GC- or metastasis-related probes according to their function information in the public databases.

Hot-start PCR and DHPLC analysis

CpG-free universal primer sets and bisulfite-modified DNA (18) were used to amplify the genes of interest. The PCR reaction mixture (30 μl) included 20 ng DNA template, 0.15 mmol/L dNTP, 0.15 μmol/L of each primer, and 0.9 U of HotStart *Taq* DNA polymerase (Qiagen GmbH, Hilden, Germany). The PCR products were then analyzed quantitatively by DHPLC using the WAVE® DNA Fragment Analysis System (12, 19). PCR products of hypermethylated and hypomethylated genes were separated using a DNASep® analytical column (Transgenomic) at the corresponding partial denaturing temperature as listed in the supplementary materials and methods). *M.SssI*-methylated genomic DNA, obtained from blood samples, was used as a positive control. A sample containing a methylated PCR product peak was defined as methylation-positive and used to calculate methylation positive rate [ratio of methylation-positive sample number to total sample number]. The peak areas

corresponding to the methylated and unmethylated PCR products were used to calculate the percentage of methylated copies [proportion of hypermethylated copies = methylation-peak area/ total peak area] for each gene analyzed.

MethyLight

The methylation states of *GFRA1*, *SRF*, and *ZNF382* were determined using the MethyLight assays. Gene-specific probes labeled with 6FAM and TAMRA were employed to quantify the relative copy number of methylated alleles compared to the *COL2A1* control (20). The sequences of the primer set and gene-specific probes can be found in the supplementary materials and methods.

Statistical analysis

The SPSS 16.0 *Trend*-test and Pearson's Chi-square test were used to analyze the difference in methylation frequency between GC and SM samples and between metastatic and non-metastatic GC samples. Student's paired *t*-test, Kruskal-Wallis *H*-test, and One-Way ANOVA were used to identify differentially methylated regions between the different groups of samples. The Mann-Whitney *U*-test and Student's *t*-test were used to analyze the association between the percentage of methylated copies and the clinicopathological features. All statistical tests were two-sided, and $P<0.05$ was considered statistically significant. The cutoff value was calculated according to the receiver operating characteristic (ROC) curve using the percentage of methylated copies to predict GC metastasis. The log-rank test was used to compare survival time between groups. Cox-proportional hazards models were used to identify independent predictors of survival (month) with adjustment for relevant clinical covariates. Functional annotation of the differentially methylated regions was performed using EpiExplorer (21).

Results

Genome-wide analysis of GC-related differential DNA methylation

In order to identify differentially methylated genes related to GC development and metastasis, genome-wide DNA methylation analysis was conducted in 8 pairs of GC and SM samples using the MCAM assay utilizing a 99K custom-designed Agilent oligonucleotide microarray as described above (10). Through this method, 9860 probes in 4047 genes were identified with significant methylation differences between the 8 GC and 8 SM samples (Paired *t*-test, $P<0.01$). Of the differentially methylated probes, 4177 showed hypermethylation (42%; [GC]>[SM]), while the remaining probes were hypomethylated (58%; [GC]<[SM]) (Data file S1). Nearly half of the hypomethylated probes (49%) were found to be within a $10^{2\sim3}$ bp region of the transcription start site (TSS), while 42% of the hypermethylated probes were within a $10^{3\sim4}$ bp region of the TSS (Supplementary Fig. S2D; $P<0.0000001$). When compared to the hypermethylated probes, the hypomethylated probes showed a considerably higher GC content than the hypermethylated ones (*Median*, 0.68 vs. 0.50), indicating the hypomethylation lies mainly in typical CGIs (Supplementary Fig. S2E). The promoter and exon-1 regions showed significantly higher proportions of hypomethylation to hypermethylation (26.8% vs. 23.4% for the promoter, $P=1.2\times10^{-4}$; 13.8% vs. 3.0% for exon-1, $P=5.5\times10^{-75}$) in GCs compared to SMs. The opposite trend was

seen in the intragenic regions, which showed a significantly lower proportion of hypomethylation to hypermethylation (32.1% vs. 41.5%, $P=5.4\times10^{-22}$) (Fig. 1A). A heatmap displaying the top-100 differentially methylated probes between GCs and SMs in the promoter and exon-1 regions is provided in Fig. 1B.

Most GC-related differentially methylated probes were clustered in specific chromosomal regions, especially sub-telomeric regions (Data File S2). Although the presence of *SmaI*/*XmaI* restriction sites primarily determined the distribution patterns of probes with GC-related methylation changes, certain chromosomal locations showed increased hypomethylation with little to no overlapping hypermethylation. After being normalized with respect to the probe density, chromosomes 7, 8, and 20 were clearly shown to harbor multiple long-range hypermethylated domains. In contrast, most regions in chromosomes 3, 4, 14, 15, and 18 were found to be more favorable to long-range hypomethylation (Fig. 1C).

Genome-wide analysis of GC metastasis-related differential DNA methylation

Among the 8 pairs of GC and SM samples analysed, half were metastatic GCs and the other half were sex-, age-, location-, and differentiation-matched non-metastatic control GCs (Supplementary Table S1). The MCAM analysis identified 8553 probes that were differentially methylated between the metastatic and non-metastatic GC groups (Mann-Whitney *U*-test, $P<0.029$). Among these metastasis-related candidate probes, 623 probes corresponded to 480 genes that overlapped with the GC-related genes identified above. A heatmap displaying the top 100 metastasis-related, differentially methylated probes is provided in Fig. 1D.

Identification in 15 GC-related aberrantly methylated genes

From the list of differentially methylated CGIs, 63 candidate genes were selected for further analysis based on their known functions and statistical significance of differential methylation signals between metastatic and non-metastatic GCs or between GCs and SMs (Supplementary Table S2). Ten known tumor-related genes that were not included in the oligonucleotide array were also selected as complementary and control genes for the validation study. The CGIs of these 73 genes were amplified using CpG-free primer sets. The bisulfite-PCR products were then analyzed using DHPLC to quantify the methylation levels of these CGIs in the 8 paired GC and SM samples (Fig. 2A; Data file S3). Differential methylation was observed in 37 CGIs between the 8 pairs of samples (Supplementary Table S2, underlined). The methylation levels of these 37 CGIs were further examined in additional 40 pairs of GC and SM samples, as well as 56 NorG samples. Significant differential methylation between paired GC/SM and NorG samples was observed in 15 CGIs (Ps<0.05; Table 1). The number of samples with hypermethylated CGIs in the promoter and exon-1 of *BMP3*, *BNIP3*, *ECEL1*, *HOXD10*, *KCNH1*, *PSMD10*, *PTPRT*, *SRF*, *TBX5*, *TFPI2*, and *ZNF382* gradually increased from the NorG → SM → GC samples (Trend or Chi-square test, Ps<0.040). These results suggest that hypermethylation of these 11 genes may play significant roles in GC development. Furthermore, the GC samples showed a significantly higher percentage of hypermethylated *CDKN2A* and *GFRA1* (Ps<0.050) and significantly lower levels of methylation in *ELK1* and *SIGIRR* when compared to the SM samples.

The positive rate of methylation in *CDKN2A* and *PSMD10* was significantly higher in the GC and SM samples than was seen in the NorG samples. In contrast, the positive-rate and proportion of methylated *ELK1* and *GFRA1* in the NorG samples were strikingly higher than in the GC and SM samples, indicating that hypomethylation of these genes occurs in gastric carcinogenesis as field effects. Furthermore, the positive rates of *BNIP3*, *KCNH1*, and *ZNF382* methylation in the GC samples were more than 3-times higher than the SM and NorG samples (29% vs. 7~4%, 42% vs. 4~14%, and 69% vs. 18~23%, respectively). Based on this information, these genes are most likely involved in GC-specific methylation changes.

The methylation states of these CGIs were further confirmed using traditional bisulfite sequencing. The bisulfite sequencing results were consistently in agreement with the DHPLC analysis (Fig. 2B and Supplementary Fig. S3). In addition, quantitative MethyLight assays using fresh or formalin-fixed paraffin-embedded tissue samples further validated the DHPLC results (Supplementary Fig. S4A; Spearman test, $P<0.020$). To understand if methylation changes in these CGIs affect gene expression, the mRNA levels of *SRF*, *ZNF382*, and *GFRA1* were analyzed in matched tissue samples using quantitative RT-PCR. The qRT-PCR results showed that mRNA expression of all three genes was inversely correlated with the prevalence of methylation in their CGIs (Supplementary Fig. S4B; Spearman test, $P<0.050$).

Confirmation of GC metastasis-related DNA methylation markers

Among the above 48 pairs of GC and SM samples, 24 pairs were from patients with lymphatic and distant metastasis, and 24 pairs were from sex-, age-, location-, and GC differentiation grade-matched patients without metastasis. Thus, the methylation states of these 15 CGIs were further analyzed to determine if they are associated with GC metastasis. DHPLC results showed that the methylation states of the *BMP3*, *GFRA1*, *SRF*, and *ZNF382* CGIs were significantly different between metastatic and non-metastatic GC samples. The proportion of methylated *BMP3* and *GFRA1* was lower in the metastatic GC samples than the non-metastatic GC samples (*median*, 1.8% vs. 5.9%; 8.6% vs. 38.6%; Mann-Whitney *U*-test, $Ps<0.040$). The positive rate of *SRF* and *ZNF382* methylation was also lower in the metastatic GC samples than the non-metastatic GC samples (4% vs. 33%; 54% vs. 79%, $P=0.020/0.066$). Therefore, the relationship between GC metastasis and methylation of these four CGIs was tested in additional GC and SM samples obtained from Chinese patients ($n=50\sim60$). When these samples were taken together as a discovery cohort, the relationship between GC lymph/distant metastasis and the methylation changes in *GFRA1*, *SRF*, and *ZNF382* was statistically significant (Table 2); however, such an association was not observed for *BMP3* (data not shown).

To investigate whether the methylation status of the three potential biomarkers mentioned above had an impact on overall survival, Kaplan-Meier analysis was performed on each gene individually. Results showed that the overall survival of GC patients with *GFRA1* or *ZNF382* methylation-high (cut-off value: percentage of methylated copies >26.4% for *GFRA1* or 1.3% for *ZNF382*) or *SRF* methylation-positive was elongated when compared to methylation-low or methylation-negative patients in the discovery cohort (log-rank test,

P=0.068, Fig. 3A; *P*=0.010, Fig. 3B; *P*=0.001, Fig. 3C, respectively). Sub-stratification analysis revealed that *SRF* methylation was only correlated with overall survival in patients with non-cardiac GCs (*P*<0.033) but not with cardiac GCs (*P*=0.146). Therefore, only non-cardiac GC patients were included in the survival analysis in the following *SRF* methylation validation cohorts.

The predictive value of these methylation markers for GC metastasis was further confirmed using three independent validation cohorts in China (*n*=222), Japan (*n*=129), and Korea (*n*=153). Because the proportion of both methylated and unmethylated alleles of CGIs can be quantitatively and simultaneously determined using DHPLC, this method was consistently used to detect the methylation levels within these CGIs in freshly-frozen gastric samples from Chinese and Japanese patients. However, MethyLight was used to analyze the paraffin-embedded samples from the Korean patients, as fresh samples were not available. Results from these cohorts showed that the methylation positive rates of *GFRA1*, *SRF*, and *ZNF382* were inversely and significantly correlated with pTNM stage and lymph metastasis in all three cohorts (Table 3). Kaplan-Meier analysis also showed that the overall survival of GC patients with higher methylation levels of *GFRA1* and *SRF* CGIs was consistently longer than those without methylation of these two genes across all three validation cohorts (Fig. 3A and 3B). However, correlation between *ZNF382* methylation and overall survival of GC patients was not statistically significant in all three validation cohorts (Fig. 3C). These results indicate that *ZNF382* methylation may be a weak GC metastasis biomarker when compared with *GFRA1* and *SRF* methylation.

In addition, after adjustment for age, sex, differentiation, location, pTNM stage, and vascular embolus, *GFRA1* or *SRF* methylation was still an adequate prognostic indicator in multivariate analysis among all patients in these validation cohorts (hazard ratios [HR]=0.543 or 0.395; 95%CI [0.304–0.938] or [0.165–0.945]; *n*=300 or 452).

Sub-stratification analysis showed that the overall survival of stage-I&II GC patients with methylated *SRF* was significantly elongated when compared to *SRF* methylation-negative patients in all four cohorts (HR=0.357, 95%CI [0.164–0.778], *n*=198). Similar difference was also observed for *GFRA1* or *ZNF382* methylation-high, but not statistically significant (HR=0.608 or 0.498, 95%CI [0.336–1.099] or [0.243–1.023], *n*=173 or 167). Among GC patients from Korea whose histological types of GCs were available, *GFRA1* methylation-high was significantly associated with low-risk of metastasis of both intestinal- and diffuse-types of GCs (positive rate: 82.4% and 76.5% for non-metastatic GCs; 43.2% and 43.1% for metastatic GCs, *Ps*<0.05). *GFRA1* methylation-high was also significantly correlated with longer overall survival of diffuse-type GC patients (HR=0.482, 95%CI [0.247–0.938], *n*=67). However, *ZNF382* methylation-high was significantly associated with low-risk of metastasis of intestinal-type GCs (93.8% vs. 62.5%, *P*=0.036), but not diffuse-type GCs.

Synergic analysis of three methylation markers

In order to investigate if a combination of the methylation markers (*GFRA1*, *SRF*, and *ZNF382*) has a synergistic effect on predicting GC metastasis, the merged data was reanalyzed in the above 4 patient cohorts. As expected, the number of patients with one or more methylated genes among the three-gene panel was significantly decreased in GC

samples with lymph/distant metastasis (Fig. 3D left; linear-trend test, $P<0.00001$; one gene vs. two genes, $P=0.046$). The sensitivity and specificity of 2~3 positive-methylation changes of 3 genes for detection of non-metastatic GCs were 60% and 67%, respectively. The positive and negative predictive values were 57% and 69%. In addition, multivariate analysis also showed that the number of combined methylation changes of *GFRA1*, *SRF*, and *ZNF382* was an independent predictor of overall survival for GC patients ($n=246$) after adjusting for the pTNM stage, GC location, differentiation, vascular embolus, age, and sex (HR=0.734; 95%CI [0.562–0.958]) (Fig. 3D right). The pTNM stage and GC location were also independent survival factors (HR=3.608; 95%CI [2.648–4.917] and HR=2.723; 95%CI [1.608–4.613], respectively). These results suggest that using a combination of this three-gene panel may function as a synergic biomarker set for predicting GC prognosis.

GFRA1, SRF, and ZNF382 expression changes in gastric carcinogenesis

The protein expression of the three genes in the paired GC and SM samples in both regular tissue sections and tissue microarray (TMA) were analyzed using the IHC assay as described in the supplementary methods (22). IHC analysis revealed that GFRA1 expression was predominantly observed in the cytoplasm of stromal cells, especially in the vessel cells in GCs (Supplementary Fig. S5A). Among 38 pairs of IHC-informative cases, the proportion of GCs with strong GFRA1-staining was significantly higher than SMs (24/38 vs. 12/38, $P<0.01$). Among 28 pairs of informative cases, the proportion of GCs with strong ZNF382-staining in epithelial cells was lower than SMs (4/28 vs. 11/28, $P<0.07$) (Supplementary Fig. S5B). Statistically significant association was not observed between GFRA1 (or ZNF382) staining and clinical parameters, such as invasion, lymph metastasis, embolus, differentiation, and overall survival. SRF-staining was only observed in the nucleus of some stromal fibroblasts and smooth muscle cells in both regular GC and SM sections (Supplementary Fig. S5C). Therefore, SRF expression was not further examined using TMA.

Discussion

Over a four-year period, a comprehensive epigenetic biomarker discovery and validation study involving over 500 patient samples from three large academic medical centers in China, Japan, and Korea had been conducted. The biomarker discovery effort started off with a genome-wide analysis of differentially methylated genes between metastatic and non-metastatic GCs in a small number of patient samples. The microarray-based methylation profiling identified a large number of GC-specific and metastasis-specific candidate genes that were differentially methylated. From the list of differentially methylated genes, a step-by-step elimination process identified a 15-gene panel associated with GC/metastasis-specific DNA methylation changes. The 15 genes were validated using multiple independent methods from a discovery cohort of GC patient samples. Finally, a methylation biomarker-set consisting of *GFRA1*, *SRF*, and *ZNF382* was validated for the prediction of GC metastasis and patients' overall survival in four cohorts from China, Japan, and Korea. This novel epigenetic biomarker set may be used in the decision making process for personalized post-operational therapy. To our knowledge, this is the first such study which specifically focuses on the metastasis of gastric cancer.

A large number of genome-wide DNA methylation studies have been reported for many different tumor types in recent years (6, 10, 11). However, most of the studies failed to perform large-scale and in-depth follow-up studies to validate the candidate genes discovered through the genome-wide analyses. As a result, few methylation markers have been developed from the large number of DNA methylation studies published so far. The present study represents the most comprehensive and quantitative characterization of DNA methylation biomarkers in GC to date. Moreover, the three methylation biomarkers associated with GC metastasis and patients' survival were validated not only in multiple cohorts, but also in freshly frozen and paraffin-embedded samples using several independent methods such as DHPLC and MethyLight. The vigorous testing performed in this study ensures the high reliability and feasibility of these novel biomarkers in different clinical settings.

It has been previously reported that 2,540 of 17,800 tested genes are differentially expressed between 80 pairs of GC and SM samples. Furthermore, it was found that there are four times as many upregulated genes in GCs than there are downregulated genes (1983 vs. 557; GSE27342) (23). Therefore, the frequent DNA hypomethylation in the promoter and exon-1 regions of the GC methylome observed in this study may account for the prevalent increase in gene expression. In fact, an increasing number of studies have reported reactivation of proto-oncogenes by DNA hypomethylation in several cancers (24–26).

Long-range epigenetic silencing and large epigenetic structures have been reported in different cancers (27–29). Differential long-range hypermethylation and hypomethylation trends may be related to cancer/tissue-specific DNA methylation (11, 30). In the present study, it was found that chromosomes 7, 8, and 20 appeared more favorable for long-range hypermethylation (or amplification of methylated-regions). In contrast, chromosomes 3, 4, 14, 15, and 18 had an affinity for long-range hypomethylation (or deletion of methylated-regions). Further studies are warranted to determine which of these long-range hypermethylated and hypomethylated- regions are GC-specific changes and which are changes across cancer-types.

Most of the 15 aberrantly methylated genes identified in GCs are involved in cell proliferation, differentiation, apoptosis, adhesion, and embryonic development (Supplementary Table S2). Previous reports have demonstrated that silencing of *BMP3*, *BNIP3*, *CDKN2A*, *HOXD10*, *TFPI2*, and *ZNF382* via methylation correlates with both the development and progression of cancers (12, 19, 31–34). Similar associations were also observed in GC samples used in the present study. Methylation changes of *KCNH1*, *PSMD10*, and *SRF* in cancer tissues have not previously been reported. Furthermore, *BNIP3*, *KCNH1*, and *ZNF382* methylation levels were more than 3-times higher in the GC samples than in SM and NorG samples. It is needed to study whether methylation of these genes may affect their expression states in gastric carcinogenesis. In addition, though *TBX5* and *ELK1* methylation is not associated with GC metastasis, the overall survival of GC patients with methylated *TBX5* or *ELK1* was longer than those without methylation (P=0.017 or 0.003; data not shown). Because some methylation changes may occur in both GC and SM samples from cancer patients, more GC-related methylation changes could potentially be identified if the NorG samples were used as the normal stomach reference.

Among three genes identified with GC development- and metastasis-related methylation changes, *GFRA1* is a cell surface GDNF/neurturin receptor and a tyrosine kinase that is normally expressed in the nervous system and kidney. However, this gene is over-expressed in gut neural crest stem cells and in many cancers (35–41). The present study provides the first evidence that hypomethylation of *GFRA1* CGIs may account for its overexpression in cancers. *SRF* is a master regulator of myogenesis and multiple cellular processes including cell proliferation and migration. Furthermore, *SRF* is known to play important roles in the epithelial-mesenchymal transition and experimental invasion through cancer and stromal cells (42–48). The present study shows, for the first time, that methylation in the exon-1 region of its CGIs may epigenetically inactivate *SRF* transcription. Most importantly, we found that *SRF* methylation was correlated with overall survival in non-cardiac GC patients, but not in cardiac GC patients. It is well known that *H. pylori* infection increases risk of non-cardiac GC, but not cardiac GC (49). The incidence of cardiac GC is also gradually increased in Western countries coincided with a decrease in prevalence of *H. pylori* infection, (50). Therefore, whether *H. pylori* infection contributes to *SRF* methylation and its biological subsequence warrants future study.

GFRA1 and *SRF* are two crucial genes in the GDNF-GFRA-RET-RAS-MEK-ERK-ELK-SRF pathway involved in cell migration and cancer invasion (37, 41, 46–48). Therefore, epigenetic alterations of *GFRA1* and *SRF* may play important roles in GC metastasis through modulating this important pathway. *ZNF382* is a candidate tumor suppressor gene, and its methylation is associated with GC development (34). However, its link with cancer metastasis has not previously been reported. In the present study, it was found that the methylation status of *GFRA1*, *SRF* or *ZNF382* was consistently and significantly associated with GC metastasis and patients' overall survival in multiple cohorts from different populations, suggesting that they may be used as potential biomarkers for predicting GC metastasis and prognosis. Most importantly, the combination of the three markers was not only identified as an independent survival factor, but also as a strong synergistic biomarker set helping to distinguish metastatic GCs from non-metastatic GCs. The TMA analysis of GFRA1 and ZNF382 from 40 GC patients failed to demonstrate statistically significant association of their protein expression with clinicopathological parameters and overall survival of these patients; however, up-regulation of GFRA1 protein and down-regulation of ZNF382 was indeed observed in the GCs compared to SMs, which is in agreement with hypo- and hyper-methylation of *GFRA1* and *ZNF382* observed in GCs. Our results suggest that DNA methylation analysis mighty be a more suitable diagnostic tool than IHC for these genes. To further prove the clinical utility of this marker panel on early prediction for GC metastasis, a prospective followup study among non-metastatic GC patients is being conducted.

In conclusion, through a comprehensive and collaborative epigenetic biomarker discovery effort, we have demonstrated that the DNA methylation changes of *GFRA1*, *SRF*, and *ZNF382* were coordinately associated with GC metastasis and overall patient survival, and this three gene panel has potential to be used as a synergistic biomarker set capable of improving the prognosis and treatment for GC patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support: This work is supported by Natural Science Foundation of China (A3 Foresight Program No. 30921140311 and 31261140372), National Basic Research Program of China (973 Program 2011CB504201 and 2010CB529300) to DD, National Institute of Health, United States (Grant CA 134304) to HS. HS is a Georgia Cancer Coalition Distinguished Cancer Scholar.

We thank Mr. James Wilson and Mr. Austin Shull at Georgia Regents University for English language editing. We also thank Dr. Yu Sun and Mr. Xiang Zheng for the TMA preparation and IHC assays.

References

1. Parkin D. Global cancer statistics in the year 2000. Lancet Oncol. 2001; 2:533–43. [PubMed: 11905707]
2. Ferlay J, Shin H, Bray F, Forman D, Mathers C, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 2010; 127:2893–917. [PubMed: 21351269]
3. Hohenberger P, Gretschel S. Gastric cancer. Lancet. 2003; 362:305–15. [PubMed: 12892963]
4. Ushijima T, Sasako M. Focus on gastric cancer. Cancer Cell. 2004; 5:121–5. [PubMed: 14998488]
5. Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypermethylation profile of human cancer. Cancer Research. 2001; 61:3225–9. [PubMed: 11309270]
6. Zouridis H, Deng N, Ivanova T, Zhu Y, Wong B, Huang D. Methylation subtypes and large-scale epigenetic alterations in gastric cancer. Sci Transl Med. 2012; 4:156ra140.
7. Deng D, Liu Z, Du Y. Epigenetic alterations as cancer diagnostic, prognostic, and predictive biomarkers. Adv Genet. 2010; 71:125–76. [PubMed: 20933128]
8. Esteller M. Dormant hypermethylated tumour suppressor genes: questions and answers. J Pathol. 2005; 205:172–80. [PubMed: 15643671]
9. Grady W, Willis J, Guilford P, Dunbier A, Toro T, Lynch H, et al. Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. Nat Genet. 2000; 26:16–7. [PubMed: 10973239]
10. Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, et al. Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet. 2007; 3:2023–36. [PubMed: 17967063]
11. Sproul D, Kitchen RR, Nestor CE, Dixon JM, Sims AH, Harrison DJ, et al. Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. Genome Biol. 2012; 13:R84. [PubMed: 23034185]
12. Deng DJ, Deng GR, Smith MF, Zhou J, Xin HJ, Powell SM, et al. Simultaneous detection of CpG methylation and single nucleotide polymorphism by denaturing high performance liquid chromatography. Nucleic Acids Research. 2002; 30:13E. [PubMed: 11752242]
13. Sobin LH. TNM, sixth edition: new developments in general concepts and rules. Semin Surg Oncol. 2003; 21:19–22. [PubMed: 12923912]
14. Shah MA, Khanin R, Tang L, Janjigian YY, Klimstra DS, Gerdes H, et al. Molecular classification of gastric cancer: a new paradigm. Clin Cancer Res. 2011; 17:2693–701. [PubMed: 21430069]
15. Bennett LB, Schnabel JL, Kelchen JM, Taylor KH, Guo J, Arthur GL, et al. DNA hypermethylation accompanied by transcriptional repression in follicular lymphoma. Genes Chromosomes Cancer. 2009; 48:828–41. [PubMed: 19530241]
16. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003; 19:185–93. [PubMed: 12538238]

17. Wang D, Zhang Y, Huang Y, Li P, Wang M, Wu R, et al. Comparison of different normalization assumptions for analyses of DNA methylation data from the cancer genome. Gene. 2012; 506:36–42. [PubMed: 22771920]
18. Eads, C.; Laird, P. Combined bisulfite restriction analysis (COBRA). In: Mills, K.; Ramsahoye, B., editors. DNA Methylation Protocols (Methods in Molecular Biology, Vol200). Totowa, New Jersey: Humana Press; 2002. p. 53-70.
19. Luo D, Zhang B, Lv L, Xiang S, Liu Y, Ji J, et al. Methylation of CpG islands of p16 associated with progression of primary gastric carcinomas. Lab Invest. 2006; 86:591–8. [PubMed: 16534497]
20. Widschwendter M, Siegmund KD, Müller HM, Fiegl H, Marth C, Muller-Holzner E, et al. Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. Cancer Res. 2004; 64:3807–13. [PubMed: 15172987]
21. Halachev K, Bast H, Albrecht F, Lengauer T, Bock C. EpiExplorer: live exploration and global analysis of large epigenomic datasets. Genome Biol. 2012; 13:R96. [PubMed: 23034089]
22. Sun Y, Li JY, He JS, Zhou LX, Chen K. Tissue microarray analysis of multiple gene expression in intestinal metaplasia, dysplasia and carcinoma of the stomach. Histopathology. 2005; 46:505–14. [PubMed: 15842632]
23. Cui J, Chen Y, Chou WC, Sun L, Chen L, Suo J, et al. An integrated transcriptomic and computational analysis for biomarker identification in gastric cancer. Nucleic Acids Res. 2011; 39:1197–207. [PubMed: 20965966]
24. Wu H, Chen Y, Liang J, Shi B, Wu G, Zhang Y, et al. Hypomethylation-linked activation of PAX2 mediates tamoxifen-stimulated endometrial carcinogenesis. Nature. 2005; 438:981–7. [PubMed: 16355216]
25. Pattani KM, Soudry E, Glazer CA, Ochs MF, Wang H, Schussel J, et al. MAGEB2 is activated by promoter demethylation in head and neck squamous cell carcinoma. PLoS One. 2012; 7:e45534. [PubMed: 23029077]
26. Fornari F, Milazzo M, Chieco P, Negrini M, Marasco E, Capranico G, et al. In hepatocellular carcinoma miR-519d is up-regulated by p53 and DNA hypomethylation and targets CDKN1A/p21, PTEN, AKT3 and TIMP2. J Pathol. 2012; 227:275–85. [PubMed: 22262409]
27. Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, et al. Increased methylation variation in epigenetic domains across cancer types. Nat Genet. 2011; 43:768–75. [PubMed: 21706001]
28. Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, et al. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. Nat Genet. 2012; 44:40–6. [PubMed: 22120008]
29. Coolen MW, Stirzaker C, Song JZ, Statham AL, Kassir Z, Moreno CS, et al. Consolidation of the cancer genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity. Nat Cell Biol. 2010; 12:235–46. [PubMed: 20173741]
30. Costello JF, Frühwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, et al. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nat Genet. 2000; 24:132–8. [PubMed: 10655057]
31. Zou H, Harrington JJ, Shire AM, Rego RL, Wang L, Campbell ME, et al. Highly methylated genes in colorectal neoplasia: implications for screening. Cancer Epidemiol Biomarkers Prev. 2007; 16:2686–96. [PubMed: 18086775]
32. Sugita H, Iida S, Inokuchi M, Kato K, Ishiguro M, Ishikawa T, et al. Methylation of BNIP3 and DAPK indicates lower response to chemotherapy and poor prognosis in gastric cancer. Oncol Rep. 2011; 25:513–8. [PubMed: 21152877]
33. Hibi K, Goto T, Kitamura YH, Yokomizo K, Sakuraba K, Shirahata A, et al. Methylation of TFPI2 gene is frequently detected in advanced well-differentiated colorectal cancer. Anticancer Res. 2010; 30:1205–7. [PubMed: 20530429]
34. Cheng Y, Geng H, Cheng SH, Liang P, Bai Y, Li J, et al. KRAB zinc finger protein ZNF382 is a proapoptotic tumor suppressor that represses multiple oncogenes and is commonly silenced in multiple carcinomas. Cancer Res. 2010; 70:6516–26. [PubMed: 20682794]

35. Cacalano G, Fariñas I, Wang LC, Hagler K, Forgie A, Moore M, et al. GFRalpha1 is an essential receptor component for GDNF in the developing nervous system and kidney. Neuron. 1998; 21:53–62. [PubMed: 9697851]
36. Wiesenhofer B, Stockhammer G, Kostron H, Maier H, Hinterhuber H, Humpel C. Glial cell line-derived neurotrophic factor (GDNF) and its receptor (GFR-alpha 1) are strongly expressed in human gliomas. Acta Neuropathol. 2000; 99:131–7. [PubMed: 10672319]
37. Frisk T, Farnebo F, Zedenius J, Grimelius L, Höög A, Wallin G, et al. Expression of RET and its ligand complexes, GDNF/GFRalpha-1 and NTN/GFRalpha-2, in medullary thyroid carcinomas. Eur J Endocrinol. 2000; 142:643–9. [PubMed: 10822229]
38. Iwahashi N, Nagasaka T, Tezel G, Iwashita T, Asai N, Murakumo Y, et al. Expression of glial cell line-derived neurotrophic factor correlates with perineural invasion of bile duct carcinoma. Cancer. 2002; 94:167–74. [PubMed: 11815973]
39. Enomoto H, Hughes I, Golden J, Baloh RH, Yonemura S, Heuckeroth RO, et al. GFRalpha1 expression in cells lacking RET is dispensable for organogenesis and nerve regeneration. Neuron. 2004; 44:623–36. [PubMed: 15541311]
40. Esseghir S, Reis-Filho JS, Kennedy A, James M, O'Hare MJ, Jeffery R, et al. Identification of transmembrane proteins as potential prognostic markers and therapeutic targets in breast cancer by a screen for signal sequence encoding transcripts. J Pathol. 2006; 210:420–30. [PubMed: 17054309]
41. Gil Z, Cavel O, Kelly K, Brader P, Rein A, Gao SP, et al. Paracrine regulation of pancreatic cancer cell invasion by peripheral nerves. J Natl Cancer Inst. 2010; 102:107–18. [PubMed: 20068194]
42. Treisman R. Identification of a protein-binding site that mediates transcriptional response of the c-fos gene to serum factors. Cell. 1986; 46:567–74. [PubMed: 3524858]
43. Miano J, Long X, Fujiwara K. Serum response factor: master regulator of the actin cytoskeleton and contractile apparatus. Am J Physiol Cell Physiol. 2007; 292:C70–81. [PubMed: 16928770]
44. Vartiainen M, Guettler S, Larijani B, Treisman R. Nuclear actin regulates dynamic subcellular localization and activity of the SRF cofactor MAL. Science. 2007; 316:1749–52. [PubMed: 17588931]
45. Busche S, Kremmer E, Posern G. E-cadherin regulates MAL-SRF-mediated transcription in epithelial cells. J Cell Sci. 2010; 123:2803–9. [PubMed: 20663922]
46. Connelly J, Gautrot J, Trappmann B, Tan D, Donati G, Huck WT, et al. Actin and serum response factor transduce physical cues from the microenvironment to regulate epidermal stem cell fate decisions. Nat Cell Biol. 2010; 12:711–18. [PubMed: 20581838]
47. Psichari E, Balmain A, Plows D, Zoumpourlis V, Pintzas A. High activity of serum response factor in the mesenchymal transition of epithelial tumor cells is regulated by RhoA signaling. J Biol Chem. 2002; 277:29490–5. [PubMed: 12039949]
48. Medjkane S, Perez-Sanchez C, Gaggioli C, Sahai E, Treisman R. Myocardin-related transcription factors and SRF are required for cytoskeletal dynamics and experimental metastasis. Nat Cell Biol. 2009; 11:257–68. [PubMed: 19198601]
49. Kim JY, Lee HS, Kim N, Shin CM, Lee SH, Park YS, et al. Prevalence and clinicopathologic characteristics of gastric cardia cancer in South Korea. Helicobacter. 2012; 17:358–68. [PubMed: 22967119]
50. Devesa SS, Fraumeni JF Jr. The rising incidence of gastric cardia cancer. J Natl Cancer Inst. 1999; 91:747–9. [PubMed: 10328099]

Statement of translational relevance

Gastric carcinoma (GC) is the second leading cause of cancer deaths in the world, with many occurring in East Asia. To identify DNA methylation biomarkers for prediction of GC metastasis, scientists and oncologists from China, USA, Japan, and Korea have carried out a five-year collaborative study to profile differential methylation patterns in metastatic and non-metastatic GCs and perform an in-depth characterization of methylation changes in the CpG islands of 73 candidate genes. From this study, we established a methylation biomarker-set composed of three genes GFRA1, SRF, and ZNF382 that could be used to synergistically predict GC metastasis and patients' overall survival from multiple patient cohorts in China, Japan, and Korea. The established marker set will be a useful clinical tool for decision-making on personalized post-operational therapy that is currently not available.

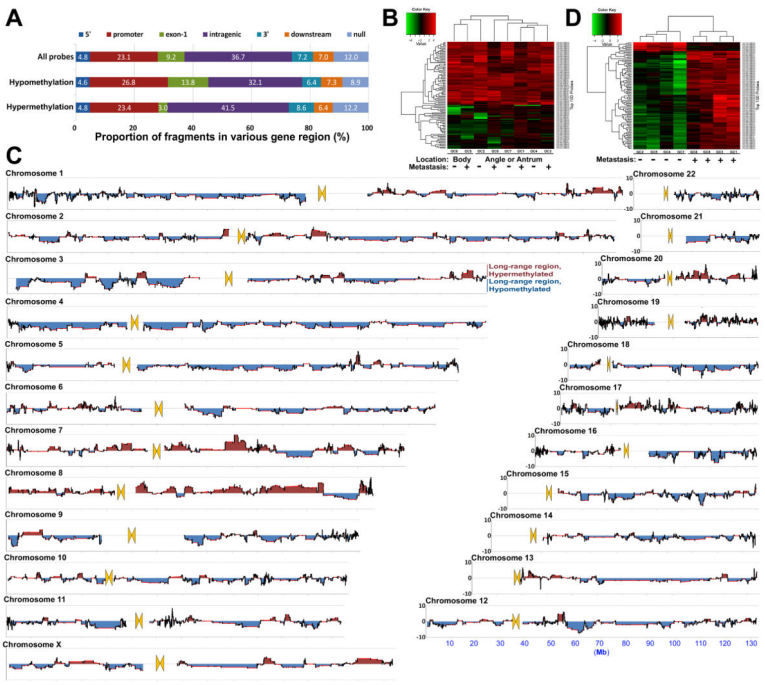

Figure 1. Distribution of probes with significant GC-related differential methylation changes in the human GC genome

(**A**) More hypomethylation was observed in the promoter and exon-1 regions, whereas more hypermethylation was observed in the gene body region. (**B**) Heatmap of the top-100 probes with differential methylation changes between GC and SM samples in supervised analysis. (**C**). Patterns of detailed regional methylation trends for each chromosome arm in GCs are displayed. The regional methylation value represents the average value of normalized methylation signal ratios between 8 GCs and 8 paired SMs for each sliding window (sequence or region) covering 51 probe-matched fragments. The long-range hypermethylated and hypomethylated regions are indicated with deep-red and blue color, respectively. Double-triangle indicates centromere. (**D**) Heatmap of the top-100 probes with differential methylation changes between metastatic samples (marked with “+”) and non-metastatic GC samples (marked with “−”) in supervised analysis.

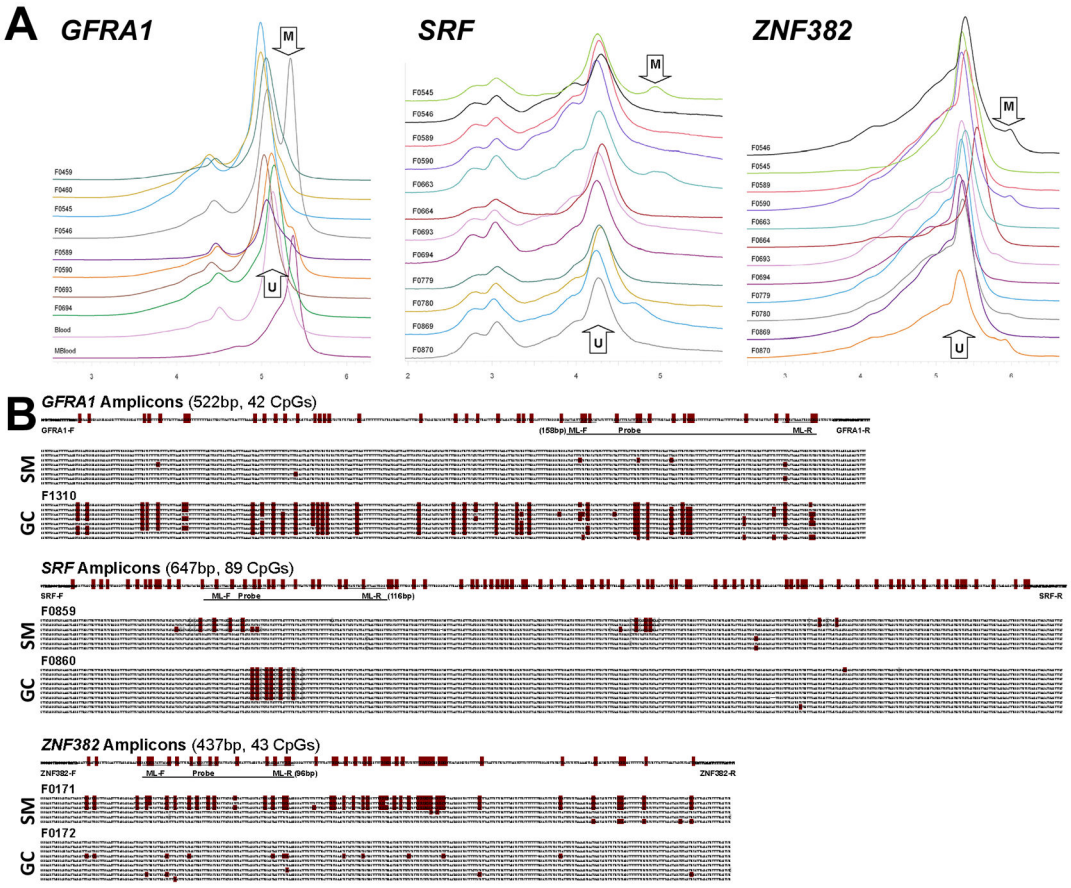

Figure 2. DNA methylation of *GFRA1*, *SRF*, and *ZNF382* in GC samples

(**A**) Representative DHPLC chromatograms of bisulfite PCR amplicons of *GFRA1*, *SRF*, and *ZNF382* CGIs, respectively. The hypermethylated (M) and hypomethylated (U) PCR products of each gene in the 8 pairs of GC and SM samples were separated with the DNASep analytical column at partial denaturing temperature as described in the method section. The peak areas corresponding to the methylated and unmethylated PCR products were used to calculate the percentage of methylated copies [proportion of hypermethylated copies = methylation-peak area/ total peak area] for each gene analyzed. (**B**) Representative bisulfite clone sequencing results of *GFRA1*, *SRF*, and *ZNF382* in the representative GC and paired SM samples. The dark red dots indicate methylated CpG sites. Locations of the primer sets and probes used in the MethyLight assays are also illustrated.

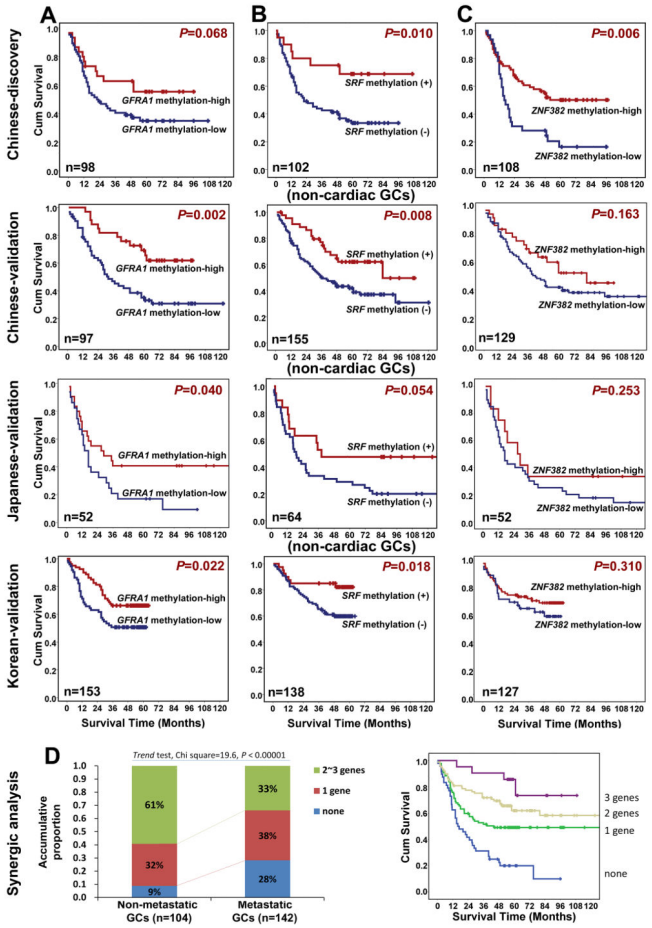

Figure 3. Kaplan-Meier survival curves of GC patients with different *GFRA1*, *SRF*, and *ZNF382* methylation states

(**A**, **B**, **C**) *GFRA1* and *ZNF382* methylation-high and *SRF* methylation-positive in GC or SM tissues were good survival factors with statistical significance for GC patients in the Chinese-discovery cohorts, Chinese, Japanese, and Korean validation cohorts. (**D**) Synergistic analysis of three methylation markers. Distribution of the number of patients with methylation changes in one to three genes (*GFRA1*, *SRF*, and *ZNF382*) in metastatic and non-metastatic GC groups. The number of patients with one more differentially methylated genes in non-metastatic GCs was significantly higher than metastatic GCs (left chart). The more number of genes associates with differential methylation, the longer of overall survival of GC patients will be (right chart).

Table 1

Prevalence of CGI methylation in gastric mucosa samples containing various pathological changes from GC patients and non-cancerous control patients

CpG islands of genes	Pathological changes	Methylation-positive rate *		Percentage of methylated copies in the methylation-positive samples	
		Positive rate (%)	*P*-value	*Median* [25% ~ 75%]	*P*-value
BMP3	GC	74/102 (72.5)	<0.001 *1*	8.6 (4.0~29.3)	<0.001 *5*
	SM	36/102 (35.3)	<0.001 *2*	1.9 (1.0~5.4)	
	NorG	3/48 (6.3)		0.2 (0.1~0.9)	
BNIP3	GC	17/58 (29.3)	<0.001 *1*	7.3 (3.6~16.4)	
	SM	4/58 (6.9)		15.3 (6.0~27.8)	
	NorG	2/45 (4.4)		20.5	
CDKN2A	GC	12/91 (13.2)	<0.001 *3*	4.12 (0.3~13.7)	0.043 *6*
	SM	15/91 (16.5)	<0.001 *2*	0.45 (0.2~1.3)	
	NorG	1/46 (2.2)		0.63	
ECEL1	GC	47/58 (81.0)	<0.001 *1*	13.9 (2.3~26.7)	0.003 *7*
	SM	26/58 (44.8)		0.0 (0.0~4.9)	
	NorG	15/42 (35.7)		0.0 (0.0~4.5)	
ELK1	GC	43/48 (89.6)		56.0 (29.3~82.1)	0.001 *5*
	SM	43/48 (89.6)		68.3 (47.9~100.0)	
	NorG	43/43 (100.0)		75.9 (67.7~100.0)	
GFRA1	GC	59/98 (60.2)		5.4 (0.0~59.2)	0.002 *7*
	SM	46/98 (46.9)	0.003 *2*	0.0 (0.0~12.6)	0.000 *8*
	NorG	35/48 (72.9)		44.4 (0.0~74.5)	0.017 *9*
HOXD10	GC	36/48 (75.0)	0.024 *1*	16.1 (0.8~21.7)	0.012 *7*
	SM	30/48 (62.5)		10.4 (0.0~15.0)	
	NorG	15/30 (50.0)		11.0 (0.0~61.8)	

CpG islands of genes	Pathological changes	Methylation-positive rate *		Percentage of methylated copies in the methylation-positive samples	
		Positive rate (%)	*P*-value	*Median* [25% ~ 75%]	*P*-value
KCNH1	GC	20/48 (41.7)	<0.001 *4*	1.0 (0.4~3.0)	0.005 *7*
	SM	2/48 (4.2)		0.3	
	Nor	6/44 (13.6)		17.1 (13.7~29.0)	
PSMD10	GC	19/48 (39.6)	0.011 3	33.9 (18.5~45.3)	
	SM	17/48 (35.4)	0.023 *2*	36.2 (11.6~45.8)	
	NorG	2/22 (9.1)		64.1	
PTPRT	GC	42/58 (72.4)	<0.001 *1*	10.6 (0.0~28.0)	0.009 *7*
	SM	20/58 (34.5)		0.0 (0.0~11.0)	
	NorG	4/21 (19.0)		0.0 (0.0~0.0)	
SIGIRR	GC	27/48 (56.3)	0.001 *4*	18.9 (0.0~30.0)	0.023 *7*
	SM	42/48 (87.5)	<0.004 *2*	23.1 (18.1~30.0)	
	NorG	29/47 (61.7)		13.5 (0.0~24.6)	
SRF	GC	30/102 (29.4)	0.030 *1*	10.7 (2.8~18.4)	
	SM	20/102 (19.6)		13.5 (6.7~36.2)	
	NorG	4/31 (12.9)		2.1 (1.0~9.6)	
TBX5	GC	45/58 (77.6)	0.032 *1*	30.8 (20.0~48.1)	
	SM	36/58 (62.1)		25.7(14.1~38.1)	
	NorG	11/21 (52.4)		11.6 (7.9~37.2)	
TFPI2	GC	38/58 (65.5)	<0.001 *1*	25.7 (0.0~32.0)	<0.001 *5*
	SM	16/58 (27.6)		0.0 (0.0~15.3)	
	NorG	4/47 (8.5)		0.0 (0.0~0.0)	
ZNF382	GC	75/108 (69.4)	<0.001 *1*	4.5 (2.0–11.8)	0.002 *6*
	SM	25/108 (23.1)		1.9 (0.7~3.5)	
	NorG	10/56 (17.9)		3.9 (0.9~7.5)	

[*] the ratio between the number of methylation-positive sample and the number of total tested sample;

[1] *trend*-test;

[2/3] SM/GC vs. NorG, Chi-square test;

[4] GC vs. SM, Chi-square test;

[5] GC vs. SM vs. NorG, Kruskal-Wallis test;

[6] GC vs. SM, Mann-Whitney *U*-test;

[7] GC vs. SM, paired t-test;

[8] SM vs. NorG, Mann-Whitney *U*-test;

[9] NorG vs. GC, Mann-Whitney *U*-test.

Table 2

SRF, *ZNF382*, and *GFRA1* methylation prevalence comparison in SM and GC samples from Chinese patients in the discovery cohort with various clinicopathological characteristics

Clinicopathological features	*SRF* methylation positive rate (%)		*ZNF382* methylation positive rate (%)		*GFRA1* methylation positive rate (%)		Percentage of methylated- *GFRA1* copies (%) *14*	
	SM	GC	SM	GC	SM	GC	SM	GC
Age								
<60	10/49 (20.4)	13/49 (26.5)	9/52 (17.3)	36/52 (69.2)	21/48 (43.8)	29/48 (60.4)	10.9 (7.5–54.3)	49.1 (5.4–62.4)
≥60	10/53 (18.9)	17/53 (32.6)	16/56 (28.6)	39/56 (69.6)	25/50 (50.0)	30/50 (60.0)	20.2 (6.5–44.5)	55.8 (17.3–85.7)
Sex								
Male	11/70 (15.7)	22/70 (31.4)	19/70 (27.1)	48/70 (68.6)	29/66 (43.9)	36/66 (54.5)	37.4 (7.1–63.7)	53.8 (18.7–84.3)
Female	9/32 (28.1)	8/32 (25.0)	6/38 (15.8)	28/38 (73.7)	17/32 (53.1)	23/32 (71.9)	8.8 (6.5–29.0)	41.4 (6.4–61.9)
Location								
Cardiac	4/19 (21.1)	8/29 (27.6)	4/20 (20.0)	12/20 (60.0)	11/17 (64.7)	10/17 (58.8)	34.8 (10.9–43.8)	52.5 (20.4–70.5)
Non-cardiac	16/83 (19.3)	22/83 (26.5)	21/88 (23.9)	64/88 (72.7)	35/81 (43.2)	49/81 (60.5)	9.6 (6.9–47.3)	51.9 (8.6–70.5)
Different.								
Well/ Mod.	6/35 (17.1)	11/34 (32.4)	7/31 (22.6)	**21/31 (67.7)**	10/28 (35.7)	16/28 (57.1)	28.9 (5.5–71.4)	**62.1 (49.8–96.8)**
Poor	12/62 (19.4)	16/63 (25.4)	18/77 (23.4)	**18/77 (23.4)** *9*	35/64 (54.7)	40/64 (62.5)	12.5 (7.1–43.8)	**42.4 (6.9–62.6)** *15*
Vascular embolus								
No	16/50 (32.0)	17/50 (34.0)	14/53 (26.4)	41/53 (77.4)	21/44 (47.7)	25/44 (56.8)	12.9 (7.0–41.9)	61.0 (35.5–90.7)
Yes	3/50 (6.0) *1*	12/50 (24.0)	11/52 (21.2)	33/52 (63.5)	22/49 (44.9)	32/49 (65.3)	22.8 (6.7–45.7)	32.1 (5.3–61.1) *16*
pTNM stage								
I–II	14/45 (31.1)	18/45 (40.0)	13/47 (27.7)	37/47 (78.7)	23/42 (54.8)	24/42 (57.1)	12.9 (5.5–43.8)	61.4 (33.9–89.8)
III–IV	5/57 (8.8) *2*	15/57 (26.3)	11/61 (18.0)	39/61 (63.9)	23/56 (41.1)	35/56 (62.5)	20.2 (7.9–47.3)	41.4 (5.7–61.9) *17*
Local invasion								
$T_{1–2}$	7/19 (37.0)	7/19 (36.8)	4/19 (21.1)	13/19 (68.4)	11/20 (55.0)	10/20 (50.0)	16.2 (6.0–43.1)	59.9 (13.8–92.1)
T_3	11/60 (18.3)	20/61 (32.8)	16/64 (25.0)	47/64 (73.4)	25/55 (45.5)	45/55 (81.8)	12.5 (6.6–63.7)	43.4 (8.9–69.8)
T_4	2/23 (8.7) *3*	3/22 (13.6) *6*	5/25 (20.0)	16/25 (64.0)	10/23 (43.5)	14/23 (60.9)	17.9 (7.6–45.7)	50.9 (14.2–61.3)
Lymph metastasis								
N_0	**16/55 (29.1)**	19/56 (33.9)	**18/58 (31.0)**	**45/58 (77.6)**	**30/55 (54.5)**	32/55 (58.2)	11.0 (5.8–42.0)	**60.6 (30.7–89.8)**

Clinicopathological features	*SRF* methylation positive rate (%)		*ZNF382* methylation positive rate (%)		*GFRA1* methylation positive rate (%)		Percentage of methylated- *GFRA1* copies (%) *14*	
	SM	GC	SM	GC	SM	GC	SM	GC
N_{1-3}	**4/47 (8.5)** *4*	11/46 (23.9)	**7/50 (14.0)** *7*	**31/50 (62.0)** *10*	**16/43 (37.2)** *13*	27/43 (62.8)	39.1 (8.0–87.2)	**22.8 (5.7–61.9)** *18*
Distant metastasis								
M_0	**16/51 (31.4)**	17/51 (33.3)	**17/54 (31.5)**	**43/54 (79.6)**	26/49 (53.1)	29/49 (59.2)	12.7 (5.8–44.1)	60.3 (31.2–85.2)
M_1	**4/51 (7.8)** *5*	13/51 (25.5)	**8/54 (14.8)** *8*	**33/54 (61.1)** *11*	20/49 (40.8)	30/49 (61.2)	28.8 (8.4–66.9)	32.1 (6.2–66.2)
(Total)	20/102 (19.6)	30/102 (29.4)	25/108 (23.1)	76/108 (70.4) *12*	46/98 (46.9)	59/98 (60.2)	14.5 (7.0~45.7)	51.9 (10.4~69.8) *19*

1/4/5 Fisher's test, P=0.002/ 0.029/ 0.005;

3 Trend test, P=0.025;

2/7/8/9/11/13 Chi-square test, P=0.004/ 0.036/ 0.040/ 0.001/ 0.035/ 0.028;

6 trend test, P=0.069;

10 Chi-square test, P=0.077;

12 SM vs. GC, P<0.001;

14 *median* (25%–75% range) for methylation positive samples;

15/16/17/18 Mann-Whitney *U*-test, P=0.026/ 0.012/ 0.015/ 0.038;

19 Paired t-test, SM vs. GC, P=0.002

Table 3

Comparison of *SRF*, *ZNF382*, and *GFRA1* methylation positive rates in GC patients with various clinicopathological characteristics in the Chinese, Japanese and Korean validation cohorts*

Clinicopathological features	Positive rate of *SRF* methylation (%)			Positive rate of *ZNF382* methylation-high (%)			Positive rate of *GFRA1* methylation-high (%)		
	Chinese	Japanese	Korean	Chinese	Japanese	Korean	Chinese	Japanese	Korean
Cut-off value**	None	None	None	>3.2%	>31.4%	>2.7%	>39.5%	>35.3%	None
Age									
<60	31/101 (30.7)	11/47 (23.4)	24/77 (31.2	16/62 (25.8)	8/41 (19.5)	40/65 (61.5)	12/44 (27.3)	19/41 (46.3)	34/74 (45.9)
≥60	34/121 (28.1)	12/31 (38.7)	21/75 (28.0)	20/67 (29.9)	27/88 (30.7)	43/62 (69.4)	21/53 (39.6)	40/88 (45.5)	33/79 (58.2)
Sex									
Male	47/164 (28.7)	12/48 (25.0)	30/112 (26.8)	30/102 (29.4)	27/89 (30.3)	57/91 (62.6)	23/67 (34.3)	40/89 (44.9)	56/106 (52.8)
Female	18/58 (31.0)	11/30 (36.7)	15/40 (37.5)	6/27 (22.2)	8/40 (20.0)	26/36 (72.2)	10/30 (33.3)	19/40 (47.5)	24/47 (51.1)
Location									
Cardiac	20/67 (29.9)	4/14 (28.6)	5/14 (35.7)	14/40 (35.0)	NA	16/21 (76.2)	12/32 (37.5)	NA	11/26 (42.3)
Non-cardiac	45/155 (29.0)	19/64 (29.7)	40/138 (29.0)	22/88 (25.0)	NA	67/106 (63.2)	21/67 (31.3)	NA	69/127 (54.3)
Different.									
Well/ Mod.	19/54 (35.2)	4/15 (26.7)	20/67 (29.9)	11/30 (36.7)	10/31 (32.3)	35/48 (72.9)	14/29 (48.3)	17/31 (54.8)	49/96 (51.0)
Poor	42/157 (26.8)	19/63 (30.2)	25/84 (29.8)	25/94 (26.6)	22/92 (23.9)	46/77 (59.7)	19/65 (29.2)	39/92 (42.4)	31/57 (57.6)
Vascular embolus									
No	52/155 (33.5)	19/63 (30.2)	11/34 (32.4)	21/63 (33.3)	NA	21/31 (67.7)	0/3 (0.0)	NA	23/34 (67.6)
Yes	11/62 (17.7) *1*	4/15 (26.7)	34/118 (28.8)	14/60 (23.3)	NA	61/95 (64.2)	33/94 (35.1)	NA	57/119 (47.9) *13*
pTNM stage									
I–II	19/52 (36.5)	10/17 (58.8)	32/85 (37.6)	17/33 (51.5)	17/35 (48.6)	53/78 (67.9)	22/36 (61.1)	24/35 (68.6)	52/84 (61.9)
III–IV	44/167 (26.3)	13/61 (21.3) *3*	13/67 (19.4) *6*	19/96 (19.8) *7*	14/78 (17.9) *9*	30/49 (61.2)	10/59 (16.9) *7*	30/78 (38.5) *3*	28/69 (40.6) *7*
Local invasion									
$T_{1–2}$	16/35 (45.7) *2*	13/33 (39.4)	35/109 (32.1)	9/18 (50.0) *8*	22/62 (35.5) *10*	57/85 (67.1)	12/21 (57.1)	35/62 (56.5) *12*	58/100 (58.0)
T_3	33/133 (24.8)	10/36 (27.8)	9/38 (23.7)	18/77 (23.4)	11/59 (18.6)	22/35 (62.9)	16/58 (27.6)	21/59 (35.6)	21/44 (47.7)
T_4	15/52 (28.8)	0/9 (0.0) *4*	1/5 (20.0)	8/33 (24.2)	2/8 (25.0)	4/7 (57.1)	4/16 (25.0) *11*	3/8 (37.5)	1/9 (11.1) *14*
Lymph metastasis									

Clinicopathological features	Positive rate of *SRF* methylation (%)			Positive rate of *ZNF382* methylation-high (%)			Positive rate of *GFRA1* methylation-high (%)		
	Chinese	**Japanese**	**Korean**	**Chinese**	**Japanese**	**Korean**	**Chinese**	**Japanese**	**Korean**
Cut-off value**	**None**	**None**	**None**	**>3.2%**	**>31.4%**	**>2.7%**	**>39.5%**	**>35.3%**	**None**
N_0	14/42 (33.3)	**7/9 (77.8)**	17/44 (38.6)	**18/32 (56.3)**	**15/28 (53.6)**	**38/46 (82.6)**	**22/33 (66.7)**	**21/28 (75.0)**	**35/50 (70.0)**
$N_{1–3}$	51/180 (28.3)	**16/69 (23.2)** *5*	28/108 (25.9)	**18/97 (18.6)** *7*	**20/101 (19.8)** *7*	**45/81 (55.6)** *5*	**11/64 (17.2)** *7*	**38/101 (37.6)** *3*	**45/103 (43.7)** *5*
Distant metastasis									
M_0	65/222 (29.3)	19/55 (34.5)	40/134 (29.9)	36/129 (28.4)	35/129 (27.1)	59/88 (67.0)	33/97 (34.0)	59/129 (42.4)	**68/113 (60.2)**
M_1	-	4/23 (17.4)	5/18 (27.8)	-	-	24/39 (61.5)	-	-	**12/40 (30.0)** *5*
(Total)	65/222 (29.3)	23/78 (29.5)	45/152 (29.6)	36/129 (28.4)	35/129 (27.1)	83/127 (65.4)	33/97 (34.0)	59/129 (45.7)	80/153 (52.3)

*The methylation states of three tested genes in frozen samples [from Chinese (SM) and Japanese (GC)] were analysed by DHPLC; in fixed paraffin samples [from Korean (SM)], by MethyLight; In addition, 33, 20, and 13 patients with pre-operative chemotherapy were included in the Chinese testing cohorts for *SRF*, *GFRA1*, and *ZNF382* methylation, respectively. Significant differences in the methylation positive rates were not observed between patients with and without pre-operative chemotherapy.

**The cutoff value is calculated according to ROC curve (not shown) when more than half of samples are methylation-positive.

1/3/5/6/7/9/13 Chi-square test, P=0.020/ 0.003/ 0.002/ 0.014/ 0.001/ 0.040/ 0.042;

2/8/10/12 Chi-square test, $pT_{1–2}$ vs. $pT_{3–4}$, P=0.018/ 0.042/ 0.040/ 0.019;

4/11/14 Trend test, P=0.028/ 0.028/ 0.011