[52] S.A. Vokes, A.P. McMahon, Hedgehog signaling: iguana debuts as a nuclear gatekeeper, Curr. Biol. 14 (2004) R668–670.
[53] F.E. Domann, J.C. Rice, M.J. Hendrix, B.W. Futscher, Epigenetic silencing of maspin gene expression in human breast cancers, Int. J. Cancer 85 (2000) 805–810.
[54] M. van Engeland, G.M. Roemen, M. Brink, M.M. Pachen, M.P. Weijenberg, A.P. de Bruine, J.W. Arends, P.A. van den Brandt, A.F. de Goeij, J.G. Herman, K-ras mutations and RASSF1A promoter methylation in colorectal cancer, Oncogene 21 (2002) 3792–3795.
[55] K.J. Wagner, W.N. Cooper, R.G. Grundy, G. Caldwell, C. Jones, R.B. Wadey, D. Morton, P.N. Schofield, W. Reik, F. Latif, E.R. Maher, Frequent RASSF1A tumour suppressor gene promoter methylation in Wilms' tumour and colorectal cancer, Oncogene 21 (2002) 7277–7282.
[56] A. Dobrovic, D. Simpfendorfer, Methylation of the BRCA1 gene in sporadic breast cancer, Cancer Res. 57 (1997) 3347–3350.
[57] M. Esteller, J.M. Silva, G. Dominguez, F. Bonilla, X. Matias-Guiu, E. Lerma, E. Bussaglia, J. Prat, I.C. Harkes, E.A. Repasky, E. Gabrielson, M. Schutte, S.B. Baylin, J.G. Herman, Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors, J. Natl. Cancer Inst. 92 (2000) 564–569.
[58] X. Xu, M.D. Gammon, Y. Zhang, T.H. Bestor, S.H. Zeisel, J.G. Wetmur, S. Wallenstein, P.T. Bradshaw, G. Garbowski, S.L. Teitelbaum, A.I. Neugut, R.M. Santella, J. Chen, BRCA1 promoter methylation is associated with increased mortality among women with breast cancer, Breast Cancer Res. Treat. 115 (2009) 397–404.
[59] M. Wei, T.A. Grushko, J. Dignam, F. Hagos, R. Nanda, L. Sveen, J. Xu, J. Fackenthal, M. Tretiakova, S. Das, O.I. Olopade, BRCA1 promoter methylation in sporadic breast cancer is associated with reduced BRCA1 copy number and chromosome 17 aneusomy, Cancer Res. 65 (2005) 10692–10699.

Gastric Cancer (2012) 15:382–388
DOI 10.1007/s10120-011-0126-1

ORIGINAL ARTICLE

Identification of gastric cancer risk markers that are informative in individuals with past *H. pylori* infection

Sohachi Nanjo · Kiyoshi Asada · Satoshi Yamashita · Takeshi Nakajima · Kazuyuki Nakazawa · Takao Maekita · Masao Ichinose · Toshiro Sugiyama · Toshikazu Ushijima

Received: 25 August 2011 / Accepted: 26 November 2011 / Published online: 12 January 2012

Abstract

Background Epigenomic damage induced by *Helicobacter pylori* infection is accumulated in gastric mucosae before the development of malignancy. In individuals without current *H. pylori* infection, DNA methylation levels of specific CpG islands (CGIs) are associated with gastric cancer risk. Because risk estimation in individuals with past infection is clinically important, we here aimed to identify the risk markers that reflect epigenomic damage induced by *H. pylori* infection, and that are informative in these individuals.

Methods Gastric mucosae were obtained from 55 gastric cancer patients (GC-Pt) (21 with current infection and 34 with past infection) and 55 healthy volunteers (HV) (7 never-infected, 21 with current infection, and 27 with past infection). Hypermethylated CGIs were searched for by methylated DNA immunoprecipitation-CGI microarray, and methylation levels were analyzed by quantitative methylation-specific polymerase chain reaction (PCR).

Results By microarray analysis of a pool of three samples from GC-Pt with past infection and another pool of samples from HV with past infection, 15 hypermethylated CGIs in the former pool were isolated. Seven of them had significantly higher methylation levels in GC-Pt with past infection ($n = 10$) than in HV with past infection ($n = 10$) ($P < 0.001$). In a validation cohort (21 GC-Pt with past infection and 14 HV with past infection), the seven new markers had large areas under the receiver-operating characteristic curves (0.78–0.84) and high odds ratios (12.7–36.0) compared with two currently available markers (0.60–0.65, 5.0–5.7).

Conclusions We identified seven novel gastric cancer risk markers that are highly informative in individuals with past infection.

Keywords Carcinogenesis · DNA methylation · Gastric cancer · *Helicobacter pylori*

Electronic supplementary material The online version of this article (doi:10.1007/s10120-011-0126-1) contains supplementary material, which is available to authorized users.

S. Nanjo · K. Asada · S. Yamashita · T. Ushijima (✉)
Division of Epigenomics, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
e-mail: tushijim@ncc.go.jp

S. Nanjo · T. Sugiyama
Third Department of Internal Medicine, University of Toyama, Toyama, Japan

T. Nakajima
Gastrointestinal Endoscopy Division, National Cancer Center Hospital, Tokyo, Japan

K. Nakazawa · T. Maekita · M. Ichinose
Second Department of Internal Medicine, Wakayama Medical University, Wakayama, Japan

Introduction

Early detection of cancer is critically important to reduce its morbidity and mortality, and early detection can be achieved by identifying individuals at high risk of developing cancers. In the risk estimation of gastric cancers, a history of *Helicobacter pylori* infection, which increases gastric cancer risk 2.2- to 21-fold [1–4], plays the major role, but the vast majority of individuals with a history of *H. pylori* infection do not develop gastric cancers. Also, gene polymorphisms associated with gastric cancers have been identified, and they have been shown to confer odds ratios (ORs) mostly between 1.0 and 2.0 [5, 6]. To obtain

clinically useful risk markers, we have to develop markers that are informative even in individuals with a history of *H. pylori* infection and that confer higher ORs.

Recently, we showed that *H. pylori* infection induces epigenomic damage, especially aberrant DNA methylation, in gastric mucosae [7]. DNA methylation levels of specific CpG islands (CGIs) were very high in the gastric mucosae of individuals with active *H. pylori* infection irrespective of gastric cancer risk, and decreased to certain levels after *H. pylori* was eradicated [8]. Importantly, these methylation levels in individuals without active *H. pylori* infection were correlated with gastric cancer risk [7, 9]. It is considered that aberrant DNA methylation is induced both in gastric stem cells and in non-stem cells, that methylation induced in stem cells will remain even after *H. pylori* eradication, and that methylation levels in individuals without current *H. pylori* infection reflect gastric cancer risk (degree of the epigenetic field defect) [10].

The correlation between methylation levels and gastric cancer risk has been analyzed in individuals without current *H. pylori* infection [7, 9]. Based on the data in our previous study [7], currently available methylation risk markers, *FLNc* and *THBD*, have ORs of 4.2–7.0 to detect gastric cancer patients (GC-Pt) among such individuals. However, individuals without current *H. pylori* infection indeed consist of never-infected individuals and those with past infection, and risk estimation is important in individuals with past infection.

In this study, we aimed to identify gastric cancer risk markers that reflect epigenomic damage induced by *H. pylori* infection, and that are informative in individuals with past infection.

Materials and methods

Tissue samples and determination of *H. pylori* infection status

Fifty-five healthy volunteers (HV) with endoscopic findings of no malignancy were recruited, with written informed consents, on the occasion of a gastric cancer screening program, with the approval of the institutional review board. Fifty-five GC-Pt who had undergone curative endoscopic submucosal dissection (ESD) of a well-differentiated adenocarcinoma in the non-cardia according to the Japanese classification of gastric carcinoma [11] were also recruited, with written informed consents, with the approval of the Institutional Review Board. Gastric mucosae were collected by endoscopic biopsy of the antrum. The biopsy specimens were frozen in liquid nitrogen immediately after biopsy, and stored at −80°C until DNA extraction. High molecular weight DNA was extracted by the phenol/chloroform method.

Current *H. pylori* infection was analyzed by a serum anti-*H. pylori* IgG antibody test (SRL, Tokyo, Japan) in HV and by urea breath test (Otsuka Pharmaceutical, Tokushima, Japan) in GC-Pt. Also, the presence of current or past *H. pylori* infection was detected by the endoscopic presence of atrophic gastritis in the antrum, because atrophic change induced by *H. pylori* infection arises in the antrum in 83% of individuals with *H. pylori* infection [12] and remains in all individuals who have had *H. pylori* eradication therapy [13]. "Never-infected individuals" were defined as those who were negative for *H. pylori* analysis and did not have atrophic gastritis in the antrum. "Individuals with current infection" were defined as those who were positive for *H. pylori* analysis. "Individuals with past infection" were defined as those who were negative for *H. pylori* analysis and had atrophic gastritis in the antrum.

Methylated DNA immunoprecipitation-CGI microarray analysis

Methylated DNA immunoprecipitation (MeDIP)-CGI microarray analysis was performed as previously described [14, 15]. Briefly, 5 μg of genomic DNA was immunoprecipitated with an anti-5-methylcytidine antibody (Diagnode, Liége, Belgium), and the precipitated DNA and the input DNA were labeled with cyanin (Cy) 5 and Cy3, respectively. A human CGI oligonucleotide microarray (Agilent Technologies, Santa Clara, CA, USA) was hybridized with the labeled probes and scanned with an Agilent G2565BA microarray scanner (Agilent Technologies). Scanned data were processed with Feature Extraction Software Version 9.1 (Agilent Technology) and Agilent G4477AA ChIP Analytics 1.3 software. The signal of a probe was converted into a "Me value", which represented the methylation level as a value from 0 (unmethylated) to 1 (methylated). Differentially methylated regions were detected by comparison between the Me values of two samples, and data were visualized in the UCSC Genome Browser (http://genome.ucsc.edu/) on NCBI36/hg18 assembly (National Center for Biotechnology Information, Bethesda, MD, USA).

Sodium bisulfite modification and quantitative methylation-specific polymerase chain reaction

Fully methylated DNA and fully unmethylated DNA were prepared by methylating genomic DNA with *SssI* methylase (New England Biolabs, Beverly, MA, USA) and by amplifying genomic DNA with the GenomiPhi amplification system (GE Healthcare, Buckinghamshire, UK), respectively. Bisulfite modification was performed using 1 μg of *Bam*HI-digested genomic DNA, and the modified

DNA was suspended in 40 μl of Tris–ethylenediamine tetraacetic acid (EDTA) buffer [16]. An aliquot of 2 μl of sodium bisulfite-treated DNA was used in one reaction of quantitative methylation-specific polymerase chain reaction (PCR; qMSP).

qMSP was performed using primer sets specific to methylated and unmethylated sequences (Supplementary Table 1), SYBR® Green I (BioWhittaker Molecular Applications, Rockland, ME, USA), and an iCycler Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, USA). The number of molecules in a sample was determined by comparing its amplification with those of standard DNA that contained known numbers of molecules (10^1–10^9 molecules). Standard DNA was prepared by purifying the PCR products using the Wizard SV Gel and PCR Clean-Up System (Promega, Fitchburg, WI, USA). The methylation level was calculated as the fraction of methylated (M) molecules in the total number of DNA molecules (number of M molecules + number of unmethylated molecules). The percentage of methylated reference (PMR) was calculated as the fraction of the methylated reference {(number of M molecules in a sample)/(number of *Alu* repeat sequences in a sample)}/{(number of M molecules in *SssI*-treated DNA)/(number of *Alu* repeat sequences in *SssI*-treated DNA)} [17].

Statistical analysis

Differences in mean methylation levels or PMR were analyzed by the Student's *t*-test. The receiver-operating characteristic (ROC) curve was drawn, and the area under the curve (AUC) and OR were analyzed by binomial distribution and binomial logistic regression analysis, respectively. All the analysis was performed using PASW statistics (SPSS, Chicago, IL, USA), and the results were considered significant when *P* values of less than 0.05 were obtained by two-sided tests.

Results

Isolation of hypermethylated CGIs in GC-Pt compared with HV in individuals with past *H. pylori* infection

A pool of three samples from HV with past infection and another pool of three samples from GC-Pt with past infection were analyzed by MeDIP-CGI microarray analysis. CGIs that were hypermethylated in the latter group compared with the former group were selected as follows: (1) Me value in the latter pool was higher than that in the former pool by 0.2 or more, (2) Me value in the former pool was lower than 0.4, and (3) criteria (1) and (2) were satisfied in three consecutive probes. A total of 15 CGIs were isolated by these criteria (Table 1), and representative data around CGI #5 are shown in Fig. 1.

From the 15 CGIs, those differentially methylated in a screening set, which consisted of 10 HV with past infection and 10 GC-Pt with past infection, were searched for by evaluating PMRs by qMSP (Supplementary Table 2). Seven CGIs (#1 to #7; Table 1), distributed on various chromosomes, were methylated at significantly higher

Table 1 CGIs identified by MeDIP-CGI microarray

CGI no.	Gene symbol	Name	Chromosomal position	Location around a gene
#1	*EMX1*	Empty spiracles, homeobox 1	2p13.2	Intron 1
#2	*miR663*	MicroRNA 663	20p11.1	Overlap
#3	*NKX6-1*	NK6, homeobox 1	4q21.23	Intron 1
#4	*OTP*	Orthopedia homeobox	5q13.3	Downstream
#5	*OPLAH*	5-Oxoprolinase (ATP-hydrolysing)	8q24.3	Downstream
#6	*CYP1B1*	Cytochrome P450, family 1, subfamily B, polypeptide 1	2p22.2	Exon 1
#7	*NEFM*	Neurofilament, medium polypeptide	8p21	Exon 1
#8	*PMF1*	Polyamine-modulated factor 1	1q22	Intron 1
#9	*BDNF*	Brain-derived neurotrophic factor	11p14.1	Intron 1
#10	*SSTR5*	Somatostatin receptor 5	16p13.3	Promoter
#11	*MYO1D*	Myosin ID	17q11.2	Intron 1
#12	*CAMK2N2*	Calcium/calmodulin-dependent protein kinase II inhibitor 2	3q27.1	Promoter
#13	*GATA4*	GATA binding protein 4	8p23.1	Promoter
#14	*NFATC1*	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1	18q23	Promoter
#15	*ANKRD9*	Ankyrin repeat domain 9	14q32.31	Exon 1

CGI CpG island, *MeDIP* methylated DNA immunoprecipitation

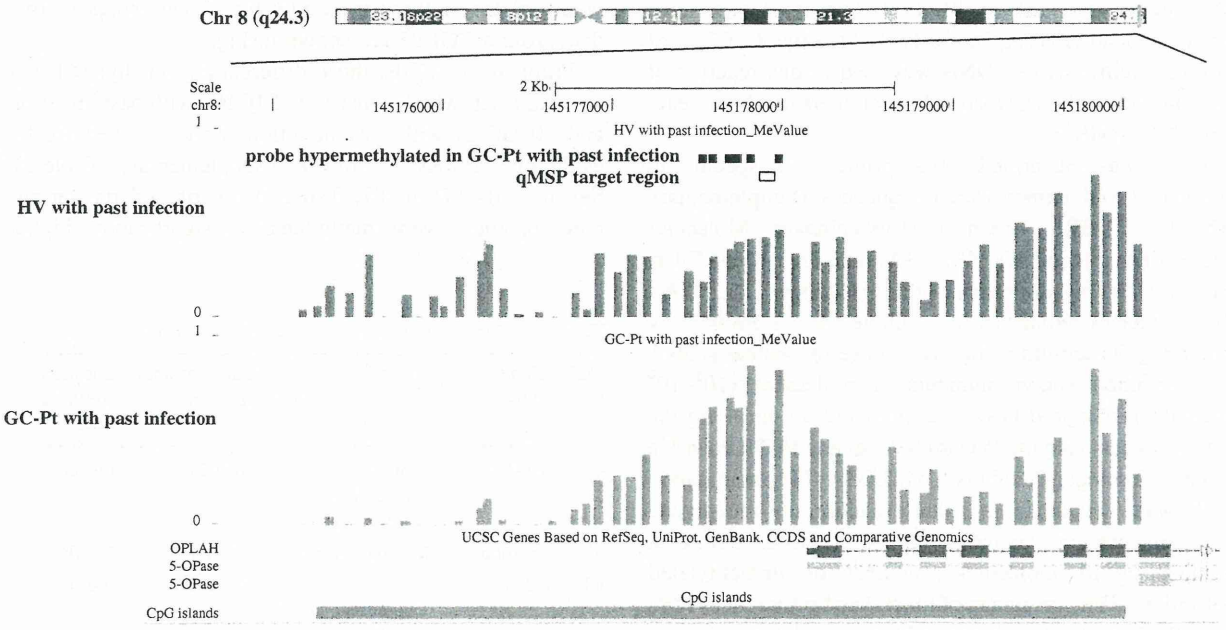

Fig. 1 Data of methylated DNA immunoprecipitation-CpG island (MeDIP-CGI) microarray analysis in the genomic region around CGI #5. Methylation levels were assessed by Me values, and the Me values of the two pools were visualized by the UCSC Genome Browser (http://genome.ucsc.edu/) for a genomic region (from nt. 145,174,733 to nt. 145,180,586 on chromosome 8 in NCBI36/hg18 assembly). *Vertical bars* show Me values of individual probes. *Closed boxes* above the Me values indicate the differentially methylated probes. Quantitative methylation-specific polymerase chain reaction (*qMSP*) primers were designed in the area shown by the *open box*. *HV* healthy volunteers, *GC-Pt* gastric cancer patients

levels in GC-Pt than in HV ($P < 0.05$). Relative positions against a gene also varied—two CGIs being located in exon 1, two in intron 1, two 300 bp downstream of the annotated end, and one overlapping with *pre-microRNA 663*.

Validation of the usefulness of the seven markers

The usefulness of the seven CGIs was validated by qMSP analysis of an independent set of samples (Fig. 2). The validation set consisted of seven never-infected HV (Group [G] 1), 21 HV with current infection (G2), 14 HV with past infection (G3), 21 GC-Pt with current infection (G4), and 21 GC-Pt with past infection (G5) (Supplementary Table 3). For comparison, two currently available markers (*FLNc* and *THBD*) were also analyzed. In the individuals with past infection (G3 and G5), the seven CGIs had levels that were 2.8-, 1.5-, 3.8-, 2.3-, 2.5-, 1.8-, and 3.8-fold, respectively, higher in G5 than in G3 ($P < 0.01$). *FLNc* tended to have a higher level in G5 than in G3 ($P = 0.087$), but *THBD* did not show any significant difference ($P = 0.341$). These data showed that the methylation levels of all the seven CGIs had the power of cancer risk estimation even in individuals with past infection.

In the HV, methylation levels in G2 were much higher than those in G1 ($P < 0.05$), but those in G3 were lower than those in G2. This observation supported the model that active infection by *H. pylori* induces methylation potently in non-stem cells, in addition to stem cells, and that methylation levels will eventually decrease after *H. pylori* infection has been eradicated. Also, methylation levels in G3 were significantly higher (four of the seven CGIs, $P < 0.05$) or tended to be higher than those in G1. This observation again supported the model that methylation induced in stem cells will remain even after *H. pylori* infection is eradicated.

Power of the seven CGIs as gastric cancer risk markers

AUCs to detect individuals in G5 were calculated using individuals in G3 and G5 (Table 2; Fig. 3). AUCs for the seven CGIs ranged between 0.78 and 0.84 and were significantly larger than 0.5 ($P < 0.01$). In contrast, the AUCs for the two currently available markers were 0.69 (95% CI 0.51–0.87) and 0.65 (95% CI 0.45–0.84), respectively, and were not significantly different from 0.5. Using optimal cut-off values obtained by the ROCs, ORs for the seven CGIs were calculated to be 12.7–36.0 (Table 2). ORs for

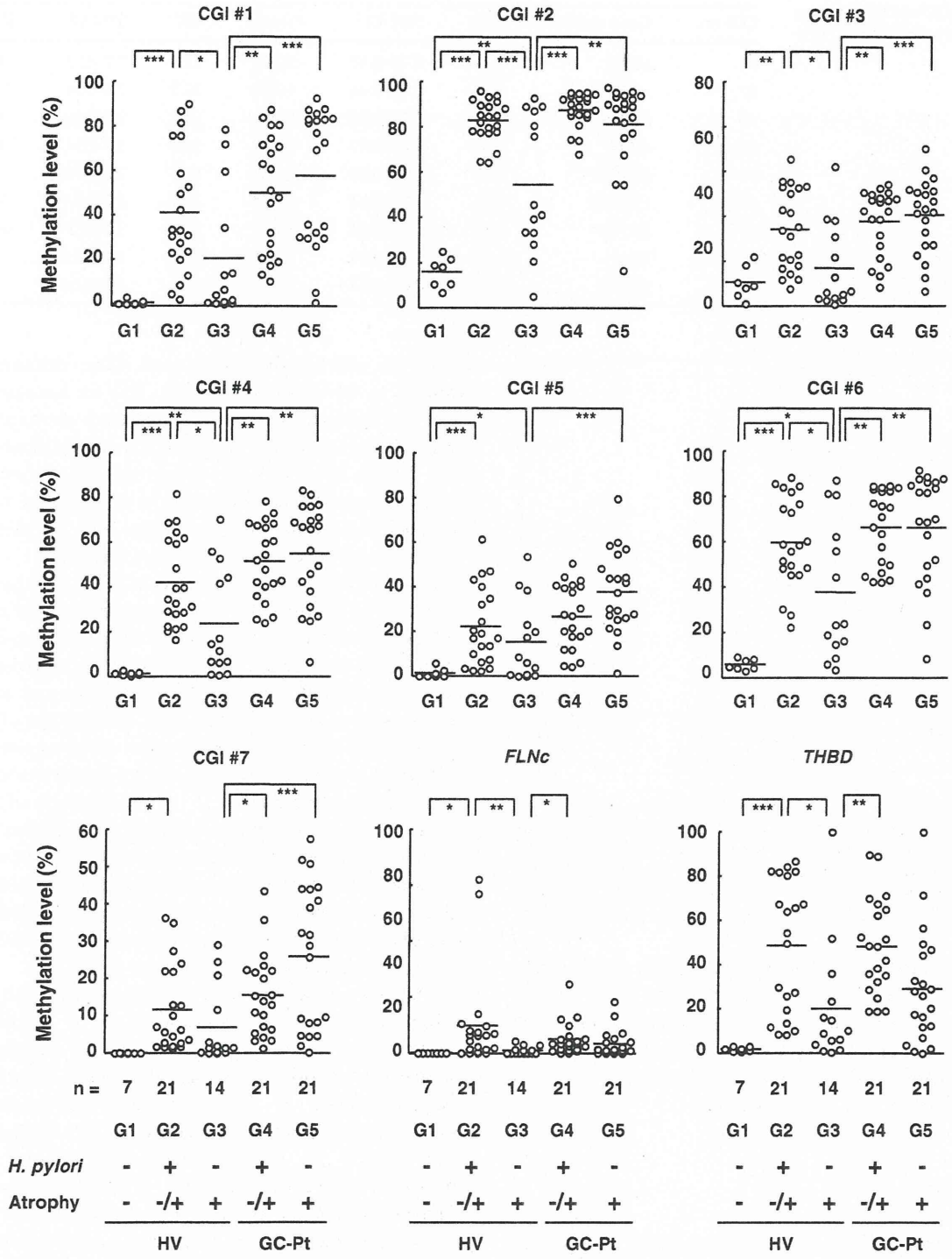

Fig. 2 Methylation levels of the seven CGIs and two currently available markers, *FLNc* and *THBD*, in the validation set. The *horizontal line* represents the mean methylation level in each group. Methylation levels of the seven CGIs in Group 5 (G5) were significantly higher than those in G3 ($P < 0.01$), but there were no significant differences for the two currently available markers. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$

Table 2 AUC and OR for new and currently available markers

CGI no.	Gene symbol	AUC	95% CI	*P* value	OR	95% CI	*P* value
#1	*EMX1*	0.84	0.70–0.97	<0.001	23.8	3.7–153	<0.001
#2	*miR663*	0.78	0.62–0.94	0.006	26.7	2.8–258	0.005
#3	*NKX6-1*	0.84	0.69–0.99	<0.001	15.0	2.8–80.1	0.002
#4	*OTP*	0.83	0.69–0.97	0.001	36.0	3.7–354	0.002
#5	*OPLAH*	0.83	0.69–0.98	0.001	15.6	2.9–83.5	0.001
#6	*CYP1B1*	0.78	0.62–0.94	0.006	12.7	2.1–76.7	0.006
#7	*NEFM*	0.84	0.71–0.98	<0.001	23.8	3.7–153	<0.001
–	*FLNc*	0.69	0.51–0.87	0.055	5.7	1.2–25.9	0.025
–	*THBD*	0.65	0.45–0.84	0.152	5.0	1.1–21.8	0.032

CGI CpG island, *AUC* area under the curve, *CI* confidence interval, *OR* odds ratio

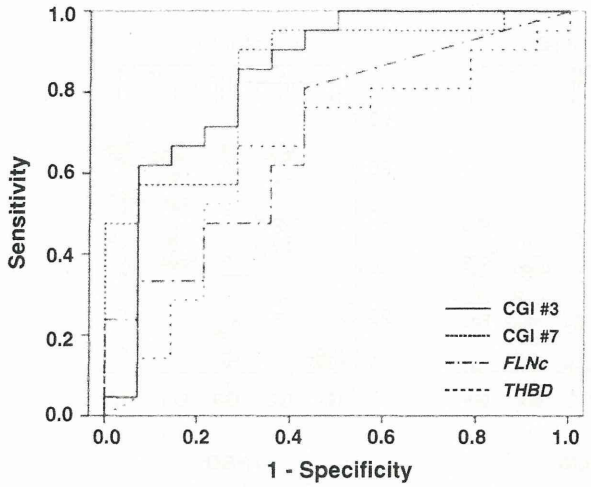

Fig. 3 Receiver-operating characteristic (ROC) curves of CGI #3 and #7, whose AUC values were the largest in the seven CGIs, are shown with those of two currently available markers, *FLNc* and *THBD*. *Black line*, *dotted line*, *dot-and-dash line*, and *dashed line* show ROC curves of CGI #3, #7, *FLNc*, and *THBD*, respectively. The AUC values of CGI #3 and #7 were larger than those of *FLNc* and *THBD*

the two currently available markers, *FLNc* and *THBD*, were 5.7 (95% CI 1.2–25.9) and 5.0 (95% CI 1.1–21.8), respectively. These results clearly showed that the methylation levels of the seven CGIs had greater power than the two currently available markers to estimate gastric cancer risk in individuals with past infection.

Discussion

In the present study, by carrying out genome-wide methylation analysis of gastric cancer patients (GC-Pt) and healthy volunteers (HV), both with past infection, we screened seven gastric cancer risk markers that are highly informative in individuals with past infection. Their usefulness was validated in 35 individuals (21 GC-Pt and 14 age-matched HV). To our knowledge, this is the first study that has evaluated epigenetic gastric cancer risk markers in individuals with past infection, and these markers are expected to be especially useful. This is because the number of individuals with past infection is increasing as more and more people receive *H. pylori* eradication therapy [18], but the usefulness of the current methods for gastric cancer risk estimation, i.e., a combination of the detection of *H. pylori* infection and the serum pepsinogen test, in this population has not been established [18–20].

None of the seven CGIs were located in promoter regions. We analyzed the association between the methylation levels of the seven CGIs and the expression levels of genes close to them, but no association was observed for any of the seven CGIs (data not shown). This was in line with the current knowledge that DNA methylation of only promoter CGIs consistently causes gene silencing, but that methylation of gene bodies may or may not be associated with increased expression [14, 21, 22]. The lack of association between methylation and gene expression supported the hypothesis that the methylation of these seven CGIs reflects the degree of overall epigenomic damage in gastric stem cells, and that the degree of epigenomic damage, and not the change of expression of individual genes, is associated with gastric cancer risk.

Epigenomic damage induced by *H. pylori* infection is one of the major causes of gastric cancer [23–26], but it is not known whether the epigenomic damage is independent of other risk factors. For example, salt intake is a risk factor for gastric cancer [27, 28], and although it does not induce methylation in gastric mucosae by itself in a Mongolian gerbil model [29, 30], it shows synergistic effects with *H. pylori* on cancer development [31]. It is not known yet whether epigenomic damage in the gastric mucosa provides independent information from past salt exposure or whether the exposure is already reflected in methylation levels. Multivariate analysis in a large cohort with a reliable record of history of salt intake will clarify this issue, and might provide a risk marker that complements the epigenetic gastric cancer risk markers.

In conclusion, we identified seven CGIs whose methylation levels are increased after *H. pylori* infection, and

are associated with gastric cancer risk even in individuals with past infection. These seven CGIs are promising candidate markers to estimate gastric cancer risk.

Acknowledgments This study was supported by Grants-in-Aid for Pioneering Basic Research and for the Third-term Comprehensive Cancer Control Strategy from the Ministry of Health, Labour and Welfare, Japan.

References

1. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, et al. *Helicobacter pylori* infection and the development of gastric cancer. N Engl J Med. 2001;345:784–9.
2. Forman D, Webb P, Parsonnet J. *H. pylori* and gastric cancer. Lancet. 1994;343:243–4.
3. Suzuki H, Iwasaki E, Hibi T. *Helicobacter pylori* and gastric cancer. Gastric Cancer. 2009;12:79–87.
4. Ekström AM, Held M, Hansson LE, Engstrand L, Nyrén O. *Helicobacter pylori* in gastric cancer established by CagA immunoblot as a marker of past infection. Gastroenterology. 2001;121:784–91.
5. El-Omar EM, Carrington M, Chow WH, McColl KE, Bream JH, Young HA, et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. Nature. 2000;404:398–402.
6. Loh M, Koh KX, Yeo BH, Song CM, Chia KS, Zhu F, et al. Meta-analysis of genetic polymorphisms and gastric cancer risk: variability in associations according to race. Eur J Cancer. 2009;45:2562–8.
7. Maekita T, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M, et al. High levels of aberrant DNA methylation in *Helicobacter pylori*-infected gastric mucosae and its possible association with gastric cancer risk. Clin Cancer Res. 2006;12:989–95.
8. Nakajima T, Enomoto S, Yamashita S, Ando T, Nakanishi Y, Nakazawa K, et al. Persistence of a component of DNA methylation in gastric mucosae after *Helicobacter pylori* eradication. J Gastroenterol. 2010;45:37–44.
9. Nakajima T, Maekita T, Oda I, Gotoda T, Yamamoto S, Umemura S, et al. Higher methylation levels in gastric mucosae significantly correlate with higher risk of gastric cancers. Cancer Epidemiol Biomarkers Prev. 2006;15:2317–21.
10. Ushijima T. Epigenetic field for cancerization. J Biochem Mol Biol. 2007;40:142–50.
11. Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma—2nd English edition. Gastric Cancer. 1998;1:10–24.
12. Asaka M, Sugiyama T, Nobuta A, Kato M, Takeda H, Graham DY. Atrophic gastritis and intestinal metaplasia in Japan: results of a large multicenter study. Helicobacter. 2001;6:294–9.
13. Ohkusa T, Fujiki K, Takashimizu I, Kumagai J, Tanizawa T, Eishi Y, et al. Improvement in atrophic gastritis and intestinal metaplasia in patients in whom *Helicobacter pylori* was eradicated. Ann Intern Med. 2001;134:380–6.
14. Yamashita S, Hosoya K, Gyobu K, Takeshima H, Ushijima T. Development of a novel output value for quantitative assessment in methylated DNA immunoprecipitation-CpG island microarray analysis. DNA Res. 2009;16:275–86.
15. Takeshima H, Yamashita S, Shimazu T, Niwa T, Ushijima T. The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands. Genome Res. 2009;16:275–86.
16. Kaneda A, Kaminishi M, Sugimura T, Ushijima T. Decreased expression of the seven ARP2/3 complex genes in human gastric cancers. Cancer Lett. 2004;212:203–10.
17. Weisenberger DJ, Campan M, Long TI, Kim MJ, Woods C, Fiala E, et al. Analysis of repetitive element DNA methylation by MethyLight. Nucleic Acids Res. 2005;33:6823–36.
18. Selgrad M, Bomschein J, Rokkas T, Malfertheiner P. Clinical aspects of gastric cancer and *Helicobacter pylori*—screening, prevention, and treatment. Helicobacter. 2010;15:40–5.
19. Kim N, Jung HC. The role of serum pepsinogen in the detection of gastric cancer. Gut Liver. 2010;4:307–19.
20. Mizuno S, Kobayashi M, Tomita S, Miki I, Masuda A, Onoyama M, et al. Validation of the pepsinogen test method for gastric cancer screening using a follow-up study. Gastric Cancer. 2009;12:158–63.
21. Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. Science. 2007;315:1141–3.
22. Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. A human B cell methylome at 100-base pair resolution. Proc Natl Acad Sci USA. 2009;106:671–8.
23. Rashid A, Issa JP. CpG island methylation in gastroenterologic neoplasia: a maturing field. Gastroenterology. 2004;127:1578–88.
24. Oue N, Motoshida J, Yokozaki H, Hayashi K, Tahara E, Taniyama K, et al. Distinct promoter hypermethylation of p16INK4a, CDH1, and RAR-beta in intestinal, diffuse-adherent, and diffuse-scattered type gastric carcinomas. J Pathol. 2002;198:55–9.
25. Shin CM, Kim N, Jung Y, Park JH, Kang GH, Kim JS, et al. Role of *Helicobacter pylori* infection in aberrant DNA methylation along multistep gastric carcinogenesis. Cancer Sci. 2011;101:1337–46.
26. Suzuki H, Tokino T, Shinomura Y, Imai K, Toyota M. DNA methylation and cancer pathways in gastrointestinal tumors. Pharmacogenomics. 2008;9:1917–28.
27. World Cancer Research Fund/American Institute for Cancer Research. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Washington, DC: AICR. 2007.
28. Nutrition and the prevention of chronic diseases. World Health Organ Tech Rep Ser. 2003;916:1–149.
29. Hur K, Niwa T, Toyoda T, Tsukamoto T, Tatematsu M, Yang HK, et al. Insufficient role of cell proliferation in aberrant DNA methylation induction and involvement of specific types of inflammation. Carcinogenesis. 2011;32:35–41.
30. Tatematsu M, Nozaki K, Tsukamoto T. *Helicobacter pylori* infection and gastric carcinogenesis in animal models. Gastric Cancer. 2003;6:1–7.
31. Nozaki K, Shimizu N, Inaba K, Tsukamoto T, Inoue M, Kumagai T, et al. Synergistic promoting effects of *Helicobacter pylori* infection and high-salt diet on gastric carcinogenesis in Mongolian gerbils. Jpn J Cancer Res. 2002;93:1083–9.

Identification of a DNA methylation marker that detects the presence of lymph node metastases of gastric cancers

YASUYUKI SHIGEMATSU[1], TOHRU NIWA[1], SATOSHI YAMASHITA[1], HIROKAZU TANIGUCHI[2], RYOJI KUSHIMA[2], HITOSHI KATAI[3], SEIJI ITO[4], TETSUYA TSUKAMOTO[5], MASAO ICHINOSE[6] and TOSHIKAZU USHIJIMA[1]

[1]Division of Epigenomics, National Cancer Center Research Institute, Chuo-ku, Tokyo 104-0045; Divisions of [2]Pathology and Clinical Laboratory and [3]Gastric Surgery, National Cancer Center Hospital, Chuo-ku, Tokyo 104-0045; [4]Department of Gastroenterology, Aichi Cancer Center Hospital, Chikusa-ku, Aichi 464-8681; [5]Department of Diagnostic Pathology, School of Medicine Fujita Health University, Kutsukake-cho, Toyoake, Aichi 470-1192; [6]Second Department of Internal Medicine, Wakayama Medical University, Kimiidera, Wakayama 641-8509, Japan

Received December 8, 2011; Accepted May 3, 2012

DOI: 10.3892/ol.2012.708

Abstract. The accurate detection of the presence of lymph node metastases (LNM) of gastric cancers (GCs) is useful for the implementation of necessary and sufficient treatment, but current methods of detection are unsatisfactory. In the present study, we focused on DNA methylation markers since they have several advantages, including biological and chemical stability and informativeness even in the presence of contaminating cells. Using three metastatic lymph nodes and three primary GCs without LNM, methylation bead array analyses were performed, which enabled the interrogation of 485,577 CpG sites. A total of 31 CpG sites that were hypermethylated in the metastatic lymph nodes, compared with the GCs without LNM, were isolated. Using primary GCs with and without LNM (28 GCs with LNM and 10 without), their methylation levels were measured using quantitative PCR following treatment with sodium bisulfite or a methylation-sensitive restriction enzyme. Of the genomic regions around the 31 CpG sites, 10 regions demonstrated higher methylation levels in the GCs with LNM compared with the GCs without LNM ($P<0.05$). Finally, the hypermethylation of the 10 regions was validated using another set of samples (129 GCs with LNM and 20 without). Hypermethylation of the region around the cg06436185 CpG site predicted the presence of LNM at a sensitivity of 43% and specificity of 85%. Additionally, the hypermethylation of the region was associated with a poor survival rate among GC patients with LNM. The results of the present study indicated that the methylation status of the region was a promising candidate marker to detect the presence of LNM of GCs and may reflect the malignant potential of GCs.

Correspondence to: Dr Toshikazu Ushijima, Division of Epigenomics, National Cancer Center Research Institute, 1-1 Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
E-mail: tushijim@ncc.go.jp

Key words: DNA methylation, gastric cancer, lymph node, metastasis

Introduction

Gastric cancer (GC) is one of the most prevalent malignancies worldwide and remains a leading cause of cancer-related mortality (1,2). Since the presence of lymph node metastases (LNM) is associated with a significantly poorer prognosis of GC patients (3-5), radical resection with free-margin gastrectomy and extended lymphadenectomy are performed for patients with advanced GC to eradicate LNM (6). Such an aggressive resection of the lymph nodes is associated with higher patient morbidity and/or mortality rates (7-9). Alternatively, the absence of LNM allows for minimally invasive surgery, which provides an improved quality of life following treatment. Therefore, the accurate detection of LNM is useful for the implementation of necessary and sufficient treatment.

To detect the presence of LNM, much effort has been made in the fields of imaging and molecular markers. Imaging modalities, including computed tomography (CT), endoscopic ultrasonography (EUS) and ^{18}F-fluorodeoxyglucose positron emission tomography (FDG-PET) are used in clinical practice. However, the sensitivities of these modalities are 77.2, 82.8 and 71%, respectively, and the specificities are 78.3, 74.2 and 74%, respectively (10-13). Moreover, these imaging modalities are almost powerless to detect micrometastases (14,15). With regard to molecular markers, analyses that targeted specific RNA and protein expression have been made. Although a number of these markers were associated with the presence of LNM of GCs (16-19), their utility has not been confirmed by independent studies. Therefore, genome-wide or comprehensive analysis of molecular markers for LNM of GCs is required and validation of the utility of the markers is essential for clinical application.

As a molecular marker, DNA methylation is advantageous, as its status is stable even if a cell is placed in different environments (biologically stable) and DNA is chemically stable, even in clinical materials. In addition, DNA methylation profiles are not disturbed by the presence of a small population of contaminating cells. As a strategy, we used metastatic lymph nodes and primary GCs without LNM for genome-wide analysis as cells with the abililty of LNM may constitute only a small population of the cells in primary GCs with LNM. Differences in methylation levels may be extremely small and may not be detected by the analysis between primary GCs with and without LNM. Alternatively, in metastatic lymph nodes, cancer cells are expected to possess the aberrant DNA methylation following clonal selection. Moreover, the methylation levels of appropriate marker CpG sites in the metastatic lymph nodes are expected to be relatively high compared with those in primary GCs with LNM.

In the present study, we aimed to identify CpG sites with a methylation status associated with the presence of LNM of GCs via a genome-wide methylation analysis using metastatic lymph nodes and primary GCs without LNM and to validate the isolated candidate markers.

Materials and methods

Patients, tissue samples and DNA extraction. A total of 187 GC surgical samples were obtained from patients who underwent gastrectomy with extended lymph node dissection (D2) at the National Cancer Center Hospital (Tokyo, Japan) and Aichi Cancer Center Hospital (Aichi, Japan) between 1994 and 2011 with informed consent. A total of three metastatic lymph nodes were obtained from 3 of the 187 patients. No patients had undergone prior chemotherapy or radiotherapy. Prognostic information of 55 GC patients with LNM was available and the mean follow-up period after surgery was 3,024 days. Disease grades were classified according to the 6th edition of the TNM classification by the UICC. Samples were stored at -80˚C and a high molecular weight DNA was extracted using the phenol/chloroform method. The 187 samples were divided into screening (28 GCs with LNM and 10 without) and validation (129 GCs with LNM and 20 without) sets in advance, between which no significant differences in clinicopathlogical data were observed (Table I). This study was conducted with the approval of the Aichi Cancer Center and National Cancer Center.

Genome-wide methylation analysis. Genome-wide screening of differentially methylated CpG sites was performed using an Infinium HumanMethylation450 BeadChip array, which covers 485,577 CpG sites (Illumina, San Diego, CA, USA) (20). Genomic DNA (1 μg) was treated with sodium bisulfite using a Zymo EZ DNA Methylation kit (Zymo Research, Irvine, CA, USA) and the bisulfite-modified DNA was amplified prior to hybridization to the array. The array was scanned with an iScan System (Illumina) and the data were analyzed using GenomeStudio Methylation Module Software (Illumina). A CpG site was considered to be informative if the sum of the signals for methylated and unmethylated sequences at the CpG site was significantly higher (at $P<0.05$) than signals of the negative control probes on the same array. Methylation levels were represented by β values, with a β value of 0 corresponding to no methylation and 1 corresponding to full methylation.

Quantitative methylation-specific PCR (qMSP). Sample DNA was treated with sodium bisulfite and purified as described previously (21). qMSP was performed using real-time PCR with bisulfite-modified DNA and specific primers (Table II, Fig. 1A). A methylation level was expressed as a percentage of the value of methylated DNA reference (PMR) calculated as the [(number of fragments methylated at a target locus in sample/number of the *Alu* sequences in sample)/(number of fragments methylated at a target locus in *Sss*I-treated DNA/ number of the *Alu* sequences in *Sss*I-treated DNA)]x100 (22).

Quantitative PCR following treatment with a methylation-dependent restriction enzyme (qPTMR). A fully unmethylated control was prepared by amplifying human blood genomic DNA with phi29 DNA polymerase (Illustra GenomiPhi HY kit, GE Healthcare, Buckinghamshire, UK) (23). DNA (1 μg) was treated with *Msp*JI (New England Biolabs, Beverly, MA, USA), which cleaves DNA 9 bp downstream from the ^{m}CNNR sequence (24,25), in a 30 μl reaction [4 U of *Msp*JI, 1X NEB buffer 4 (New England Biolabs) and 0.1 mg/ml BSA] at 37˚C for 20 h. Following purification, the DNA was treated with *Msp*JI again and dissolved in TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) at a concentration of 5 ng/μl without purification. Using 1 μl of the solution, quantitative PCR (qPCR) was performed by real-time PCR with primers that encompassed a target *Msp*JI site (Fig. 1B). To normalize the quantity of input DNA, the number of copies of a standard sequence, which may be amplified with a primer pair (5'-TTGCTTGAAGTTTTGTTGCTGTAGT-3' and 5'-AATAAACTCAGTTGTGACATGGACA-3') and contains no *Msp*JI site, was measured by qPCR. A percentage of the value of unmethylated reference (PUR) was calculated as the [(number of fragments at target locus in sample/number of the standard sequence in sample)/(number of fragments at target locus in GenomiPhi-amplified DNA/number of the standard sequences in GenomiPhi-amplified DNA)]x100. For convenience, the methylation level was expressed as 100-PUR.

Statistical analysis. Statistical analyses were conducted using PASW statistics version 18.0.0 (SPSS Japan Inc., Tokyo, Japan). The difference between the mean values of the two groups of samples was evaluated using Welch's t-test. The Fisher's exact test was used to evaluate the significant difference in relative frequency of the phenomena between two independent groups. Survival curves were computed according to the Kaplan-Meier method and the log-rank test was employed to evaluate the level of significant difference. $P<0.05$ was considered to indicate a statistically significant difference.

Results

Genome-wide screening using metastatic lymph nodes and GCs without LNM. To isolate the CpG sites that are hypermethylated specifically in GCs with LNM, genome-wide methylation analysis was performed using metastatic lymph nodes (n=3) and GCs without LNM (n=3) using an Infinium HumanMethylation450 BeadChip array. The samples used for this analysis were prepared from 6 patients in the screening

Table I. Clinicopathological data of sample sets.

	N	Age (years)	P-value	Gender	N	P-value	T stage	N	P-value
Genome-wide analysis set[a]									
Meta (-)	3	72±4	0.17	Male	2	1.0	T1	0	0.51
							T2	1	
				Female	1		T3	1	
							T4	1	
Meta (+)	3	59±13		Male	2		T1	0	
							T2	0	
				Female	1		T3	1	
							T4	2	
Screening set									
Meta (-)	10	69±6	0.13	Male	7	0.53	T1	0	0.17
							T2	1	
				Female	3		T3	6	
							T4	3	
Meta (+)	28	63±11		Male	18		T1	0	
							T2	0	
				Female	10		T3	14	
							T4	14	
Validation set									
Meta (-)	20	63±11	0.71	Male	13	0.6	T1	0	0.14
							T2	3	
				Female	7		T3	8	
							T4	9	
Meta (+)	129	62±10		Male	91		T1	0	
							T2	4	
				Female	38		T3	55	
							T4	70	

[a]This set comprised samples from the screening set.

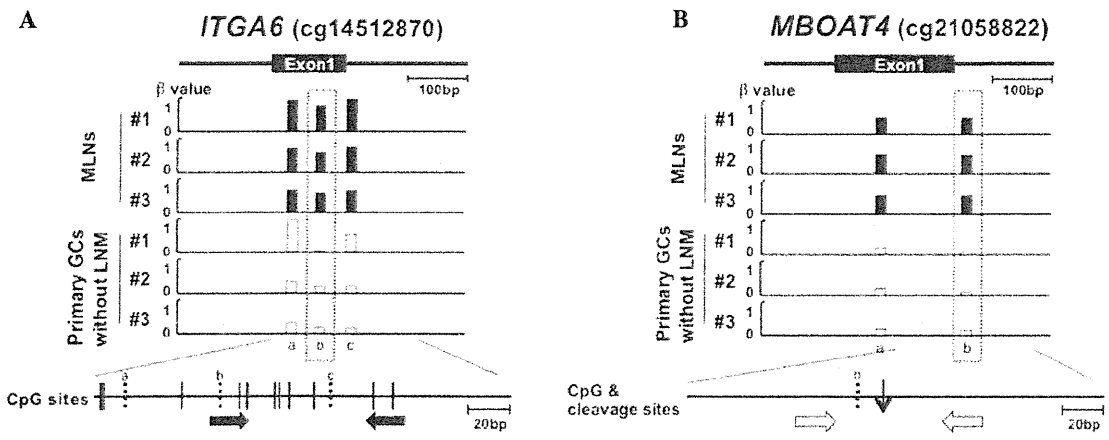

Figure 1. Representative genomic regions around the CpG sites differentially methylated between metastatic lymph nodes and GCs without LNM and primer design in the regions. Below the genomic structure of a region, β values (methylation levels) of the CpG sites carried by Infinium bead array are shown. The differentially methylated CpG site is marked by a rectangle with dotted line. A CpG map is drawn at the bottom, vertical lines (solid and broken lines) indicate CpG sites and broken lines indicate CpG sites whose β values were measured. (A) A region whose methylation level was assessed by qMSP. Primers specific to the methylated sequence (closed arrows) were designed on CpG sites around the differentially methylated sites based on the bisulfite-modified sequence. (B) A region whose methylation level was assessed by qPTMR. Primers (open arrows) were designed to amplify the region encompassing the *Msp*JI-cleaved site (thin vertical arrow) based on the unmodified sequence. GC, gastric cancer; MLN, metastatic lymph nodes; LNM, lymph node metastases; qMSP, quantitative methylation-specific PCR; qPTMR, quantitative PCR following treatment with a methylation-dependent restriction enzyme.

Table II. CpG sites identified by bead-chip array analysis.

No.	Probe name (IlmnID)[a]	Gene symbol	Location (Chr: base)	Relation to CpG island	Position to gene	P-value[b]		Cut-off (YI)	Primer sequences (5'-3')		Annealing temp.	PCR type	Mg^{2+c} (μM)
						Screening	Validation		Forward	Reverse			
1	cg23218354	-	Chr1: 2885244	Island	-	0.05	0.17	10.1 (0.48)	TGGTTTTTATACGGGGGATTTAC	ACTAAACCAAAACGACGATTACG	60	qMSP	1.5
2	cg13239126	*KIAA1026*	Chr1:15256136	-	Body	0.24	-	-	CTCCAGAGAGACAGGCATGGTT	CAAGCCTGACCTTCCCTCTCC	60	qPTMR	1.5
3	cg16112880	*TMEM9*	Chr1:201123745	Island	TSS200	0.41	-	-	CCCGCCCTCTCCTAGCTTCTAT	GGCTGACGTTCCCTTTTCTGGT	63	qPTMR	1.5
4	cg14512870	*ITGA6*	Chr2:173330342	-	Body	0.01	0.07	32.5 (0.44)	TATAGTTGCGATATTATCGTTC	AAACTACCGAAATAAACGCT	51	qMSP	2.5
5	cg09866366	*ABCF3*	Chr3:183903315	Shore	TSS1500	0.34	-	-	TCGTTAGATTACGGGTGTTTC	CAAAACGCATATATAACGATAACG	58	qMSP	2.5
6	cg08812108	-	Chr6:2515318	-	-	0.03	0.24	56.3 (0.44)	AGCGTTGGCGTTAGGTAGGGTAGTTC	CCAAATAACCACCTACGTCTTTACG	63	qMSP	1.5
7	cg06728252	*ABT1*	Chr6:26598149	Island	Body	0.24	-	-	CGCGTAGATCGGTTCGTGAGAC	GCCACGCGCTTAACTATACG	63	qMSP	1.5
8	cg08972588	*TNXB*	Chr6:32014674	-	Body	0.64	-	-	CCTGAGCAAGAATGAGGCCAGA	GGGGACAAGGGGGAGATCACA	65	qPTMR	2.5
9	cg22126965	*COX19*	Chr7:1015501	Shore	TSS1500	0.50	-	-	GGTTTAGAAAGGTTTAGCGAATTGTTC	AACAACCGCAAACAACG	62	qMSP	2.5
10	cg18450582	*DYNC1I1*	Chr7:95546539	-	Body	0.32	-	-	ACCTTGGCCTCTGGATTGTGGA	GCACTGCCTGCCTGAAAGGAGA	64	qPTMR	1.5
11	cg02005782	-	Chr7:105857664	-	-	0.59	-	-	GAAGTCAGCCAGGCATTGGAAG	CCCAGCTGCCTTTCTGATCTCT	65	qPTMR	1.5
12	**cg06436185**	*PRKAG2*	Chr7:151442351	-	Body	0.04	0.03	28.8 (0.24)	ATTTAGTTTTTTGTACGGTTGC	CCCAATAAAACGACGTAACG	55	qMSP	2.5
13	cg21058822	*MBOAT4*	Chr8:30002223	-	TSS200	0.38	-	-	GGCTGTCTCTGGTCTTTTTATC	AGAAAGCCAGTTTTTATTCTGC	61	qPTMR	1.5
14	cg12089032	-	Chr8:72881203	-	-	0.03	0.09	40.6 (0.41)	GCAAGTTAAGGCATCGTAGGAAAGC	GGCAGAGAGGAACAGCTCCTAAG	66	qPTMR	1.5
15	cg23170346	-	Chr8:134863880	-	-	0.95	-	-	CTAGCCACATCCATAGCAGACAGG	CACTCAGCAATGCAAACAGTCTTG	66	qPTMR	1.5
16	cg19878482	*C8orf73*	Chr8:144655026	Shore	TSS200	0.10	-	-	GGAGTTTTTCGGGTTCGGTTTC	CAAAAACCCATTATAAACACGTCCGT	65	qMSP	2.5
17	cg01263942	*DIP2C*	Chr10:695859	-	Body	0.01	0.12	23.2 (0.38)	GTTCGTTATTTGCGTTTTCGTGC	CAACGAAAAAACTCCATAAACCG	59	qMSP	2.5
18	cg03015672	*ARHGAP12*	Chr10:32216066	Shore	5'UTR	0.88	-	-	AGAACAGTGGAGCCGCATGCAA	CCAAAGCAGGCAGTGAAAGCGT	66	qPTMR	1.5
19	cg10326726	*MSMB*	Chr10:51549505	-	TSS200	0.16	-	-	CAACCCTCTGTAAACACTCAAT	TATAGACAGGTACATCCAGGCA	57	qPTMR	2.5
20	cg19864370	-	Chr10:80354592	-	-	0.00	0.29	70.6 (0.69)	GAATAGCTTAGGCCCCTGTCAT	GATAGTGCTAGCCCTTGGGAAT	60	qPTMR	1.5
21	cg03850986	*ABLIM1*	Chr10:116408382	-	Body	0.38	-	-	TGATAAAAATGCTCTGGAATTAG	TGGAGATGTAATGTAGTACACCATA	51	qPTMR	1.5
22	cg25885280	*SHANK2*	Chr11:70760166	-	Body	0.34	-	-	GCGGTGGGGGATTTCTGTAAGGA	GAGCAGGGTGTGCCTTCTCAGGG	68	qPTMR	1.5
23	cg26894278	*CRYL1*	Chr13:21016241	-	Body	0.22	-	-	GTTAAGTTTAAATGGAGCCTTG	TGACAGGATTACAATAAGGCTA	56	qPTMR	1.5
24	cg04339360	*KLF5*	Chr13:73635568	Shore	Body	0.04	0.31	25.4 (0.43)	TAGTCAAGAAAAGAAACCTGTGCAA	TGCCAAACTACCTCAATTCTGTTTA	61	qPTMR	1.5
25	cg16206504	-	Chr13:114917223	Shelf	-	0.02	0.35	35.2 (0.41)	CGAGATTGTAGGCGGTTGTTC	CCTAACTATTACAACAATACCGAACG	63	qMSP	1.5
26	cg14851578	-	Chr14:106187192	Shore	-	0.08	-	-	GGAGTGTGGGTTACGTGTGATTAC	CAATCTCGCCCACTCACG	66	qMSP	1.5
27	cg02990302	*C16orf80*	Chr16:58155189	-	Body	0.04	0.45	56.2 (0.65)	TCCTTTCCTTAGCTCCTTCCAG	AAAAACAGTCGGCTCTTTGTGA	63	qPTMR	1.5
28	cg08292959	*MGAT5B*	Chr17:74878420	Island	Body	0.97	-	-	GGCACCTGCCACTCCATCCG	TGCACTCTGGGCTGTACCACAGTG	63	qPTMR	1.5
29	cg15645685	*PBX4*	Chr19:19730175	Shore	TSS1500	0.26	-	-	CTAATGCTCCCTGCATCCTCAG	TAAACAAGCGAGGTCACTCTTCAGC	64	qPTMR	1.5
30	cg14571622	*NLRP8*	Chr19:56499348	-	3'UTR	0.01	-	-	TGGGGCTTGATTGATCAGTTCC	CCAGGGTTCAAAGCTGAGGTTC	62	qPTMR	1.5
31	cg27050343	*OTC*	ChrX:38211596	-	TSS200	0.15	-	-	AATTTTTGGGTTTAAGTGATTCGTTC	AAAAAAATAATTACTAACCGAACACG	62	qMSP	1.5

[a]IlmnID is a unique CpG site identifier from the Illumina CG database; underlined IlmnID indicates P<0.05 in the screening set; bold IlmnID, P<0.05 in the validation set. [b]Difference was evaluated by Welch's t-test. [c]Final concentration in a PCR analysis. Island, CpG island; Shore, within 2,000 bp of end of CpG island; Shelf, within 2,000 bp of end of Shore; TSS200, within 200 bp upstream of the transcription start site; TSS1500, 200-1,500 bp upstream of the TSS; 5'UTR, in the 5' untranslated region of the gene; Body, within exons or introns of the gene; 3'UTR, in the 3' untranslated region of the gene; YI, maximized Youden index.

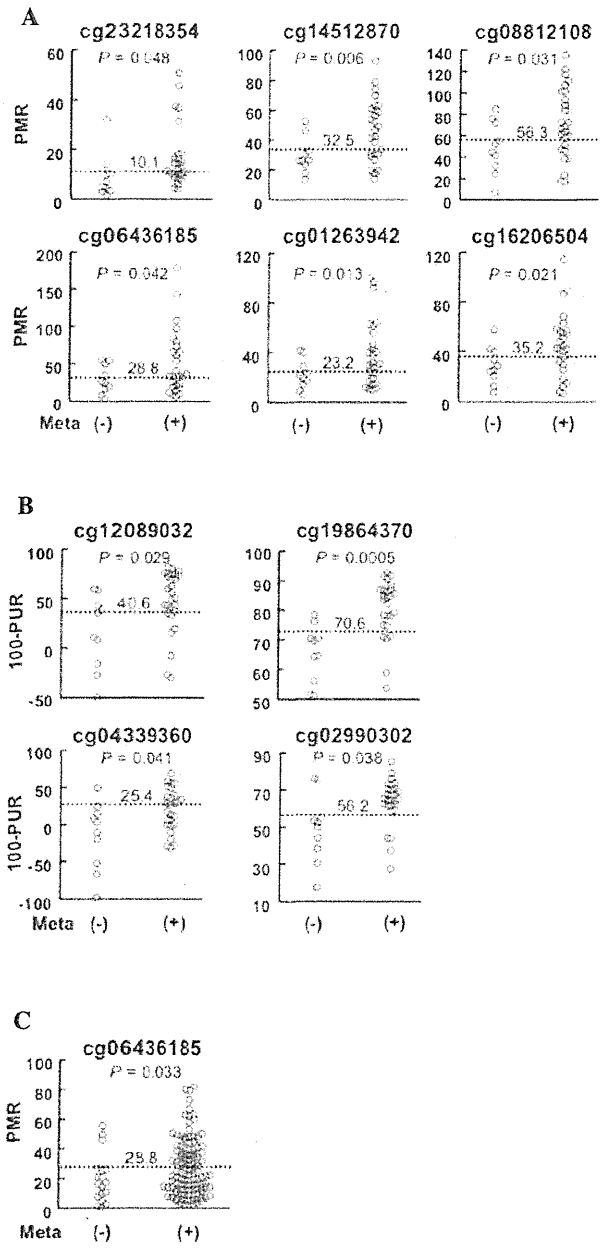

Figure 2. Methylation levels of the candidate genomic regions in primary GCs with and without LNM. Methylation levels were measured by (A) qMSP and (B) qPTMR in the screening sets. The screening set consisted of 10 GCs without LNM and 28 with LNM. (C) Methylation level of the region around cg06436185 in the validation sets was measured by qMSP. The validation set consisted of 20 GCs without LNM and 129 with LNM. Meta (-), GCs without LNM; Meta (+), those with LNM. Horizontal dotted lines are the cut-off methylation levels and the number on the line indicates the value of the level. GC, gastric cancer; LNM, lymph node metastases; qMSP, quantitative methylation-specific PCR; qPTMR, quantitative PCR following treatment with a methylation-dependent restriction enzyme; PUR, percentage of the value of unmethylated reference; PMR, percentage of the value of methylated DNA reference.

set (Table I). The mean number of informative CpG sites was 485,170 (SD 209) in the metastatic lymph nodes and 485,001 (SD 514) in the GCs without LNM (P=0.63). We searched for CpG sites that were highly methylated in the three metastatic lymph nodes [β value > a) 0.6, b) 0.5 and c) 0.4] and hardly

Table III. Association between methylation levels of the genomic region around cg0643618 and clinical characteristics.

		Methylation level		
Parameters	N	Mean	SD	P-value
Age				
≤60	77	32.2	24.0	0.26
>60	110	28.1	25.2	
Gender				
Female	58	34.9	26.4	0.07
Male	129	27.6	23.6	
T category				
T3	83	26.8	27.1	0.07
T4	96	33.6	22.5	

methylated in the three primary GCs without LNM (β value <0.2) and the number of hypermethylated CpG sites was a) 1, b) 31 and c) 209, respectively. To obtain a practicable number of candidate CpG sites, we adopted a cut-off β value of 0.5 and the 31 CpG sites were selected for further analysis (Table II).

Selection of informative candidate genomic regions among primary GCs. Using primary GCs with and without LNM (screening set, Table I), the methylation levels of genomic regions around the 31 CpG sites were measured by qMSP or qPTMR, which are accurate and sensitive enough to detect aberrant DNA methylation in a small population of cells. Of the 31 regions, 10 regions exhibited higher methylation levels in GCs with LNM (1.4- to 1.9-fold) than in those without LNM (Table II and Fig. 2A and B). For each of the 10 genomic regions, a cut-off methylation level was established in order that the Youden index (sensitivity + specificity - 1) would be maximized (Table II and Fig. 2).

Validation of the candidate genomic regions in a different set of samples. To validate the hypermethylation of the 10 candidate genomic regions in GCs with LNM, the methylation levels were analyzed in an independent sample set (validation set, Table I). A region around the cg06436185 CpG site revealed significantly higher methylation levels in GCs with LNM (1.5-fold) than those without (P=0.033, Fig. 2C), whereas the other nine regions were not validated (Table II). The region was located in the gene body of the *PRKAG2* gene and did not belong to a CpG island (Table II). Therefore, it was unlikely that the methylation status of the region around cg06436185 affected the transcription of a gene. Using a cut-off level established in the analysis of the screening set (28.8%), the presence of LNM was detected at a sensitivity of 43% and a specificity of 85%. This result indicated that a methylation level of this region is a candidate marker for the detection of the presence of LNM.

Association between the methylation level of the genomic region around the cg06436185 CpG site and clinicopathological characteristics. Associations between the methylation level of the genomic region around cg06436185

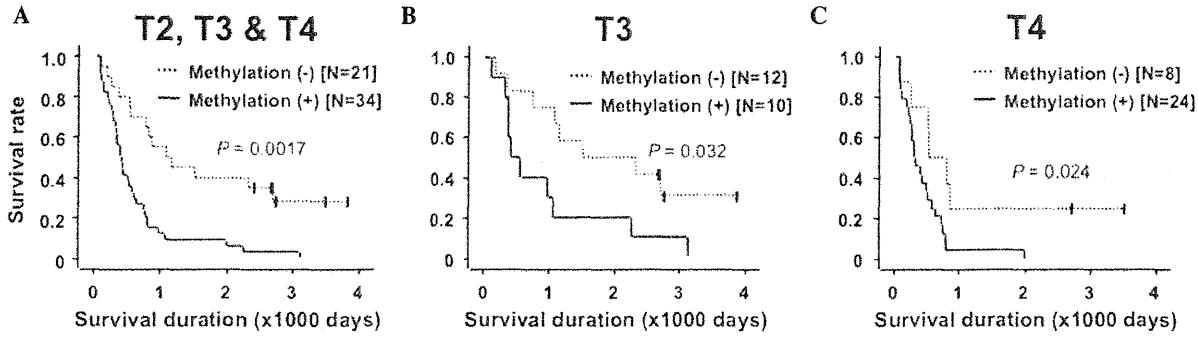

Figure 3. Kaplan-Meier analysis of 55 GC patients with LNM. (A) Overall survival in all the patients with LNM. (B and C) The survival curves of patients categorized into T3 and T4. Methylation (+), GCs with PMR of the region around cg06436185>28.8; Methylation (-), GCs with the PMR<28.8. GC, gastric cancer; LNM, lymph node metastases; PMR, percentage of the value of methylated DNA reference.

and clinicopathological characteristics (age, gender and T category) were analyzed in 157 GC patients with LNM and 30 without LNM. No difference in methylation levels according to age, gender or T category was found (Table III). Using 55 of the 157 GC patients with LNM, whose prognostic information was available (T2, one patient; T3, 22 patients; T4, 32 patients), a correlation between the methylation level and survival rate was analyzed. Patients with high methylation levels (>28.8%; the value used to detect the presence or absence of LNM) had a significantly poorer overall survival rate compared to those with low methylation levels (P=0.0017; Fig. 3A). Since the T category is known to be the major prognostic factor in GC patients (26), patients in the T3 and T4 categories were analyzed separately. In the T3 and T4 subgroups, the patients with high methylation levels demonstrated a significantly poorer overall survival rate than those with low methylation (P=0.032 and 0.024, respectively; Fig. 3B and C). These results revealed that the high methylation level of the genomic region around cg06436185 was associated with an unfavorable prognosis, regardless of the depth of tumor invasion.

Discussion

Using a genome-wide methylation analysis using metastatic lymph nodes and primary GCs without LNM, a genomic region (around cg06436185) whose methylation level in primary GCs was associated with the presence of LNM was successfully identified. Notably, the association was also significant in an independent validation set (P=0.033). Generally, markers isolated by genome-wide analyses need to be validated in a different set of samples due to the overfitting issues caused by multiple testing (27). Even in the present study, 9 of the 10 candidate genomic regions that revealed significant hypermethylation in GCs with LNM in the screening set (P=0.0005-0.048) were not reproduced in the validation set. This observation emphasizes the value of the methylation level of the genomic region around cg06436185. Since it had a sensitivity of 43% and specificity of 85%, the combined use of this novel methylation marker with imaging tools is predicted to improve the diagnostic accuracy of LNM of GCs.

The mean methylation levels of GCs with and without LNM were 18.7 and 27.5%. This small difference is extremely difficult to detect by a genome-wide screening method. Our strategy in the present study was to benefit from the monoclonal growth of cells in metastatic lymph nodes and compare metastatic lymph nodes and GCs without LNM. The methylation levels of the genomic regions around cg06436185 were 13.2 and 54.3%, respectively, in these samples. This relatively significant difference was identified using genome-wide screening, which has a relatively low accuracy in the analysis of methylation levels. Using a more accurate and sensitive method, qMSP, the small difference between GCs with and without LNM (18.7 and 27.5%, respectively) was clearly demonstrated.

A method to measure methylation levels in CpG-poor genomic regions, qPTMR, was developed using a combination of digestion with a methylation-dependent restriction enzyme and qPCR. qPTMR had an error range of 5% in this study. It is difficult to measure methylation levels in CpG-poor genomic regions by qMSP, a well-established method with a high accuracy, due to the difficulty in designing primers. Alternatively, *Msp*JI, a recently developed methylation-sensitive restriction enzyme, recognizes ^{m}CNNR (N=A, T, G or C; R=G or C) sequences and cleaves DNA when the C is methylated (24,25). Since the recognition sequence is applicable to the majority of CpG sites and cytosines in non-CpG sites are not methylated in somatic cells, the positive cleavage by *Msp*JI is used to determine methylation status of most CpG sites. Using qPTMR, the methylation levels of all the 19 candidate regions with few CpG sites were quantified. This new method is predicted to have various applications.

The methylation status of the genomic region around cg06436185 was unlikely to affect transcription of a known nearby gene (*PRKAG2*). However, its high methylation level in GCs, namely large fractions of cancer cells with methylation in cancer tissue, was associated with the presence of LNM and also with a poorer prognosis of the GC patients. One possible reason is that the region is located in a promoter region of unknown genes, including microRNA genes, or in enhancer regions whose methylation is critical for the regulation of gene expression levels. Another possible reason is that the methylation of the region is caused by an abnormality of unknown methylation regulation and that this abnormality is critical for tumor metsastasis or malignancy. In this case, other genomic regions are likely to be methylated in GCs with LNM or a poorer prognosis.

In conclusion, we identified one genomic region with a methylation status in primary GCs that was associated with the presence of LNM and a poorer prognosis of GC patients.

Acknowledgements

We thank Dr Michihiro Ishida, Dr Yukie Yoda and Dr Masahiro Maeda for their assistance during sample preparation. This study was supported by the Third-term Comprehensive Cancer Control Strategy from the Ministry of Health, Labour and Welfare, Japan, by the JSPS A3 Foresight Program and by the National Cancer Center Research and Development Fund. Y.S. is a recipient of Research Resident Fellowships from the Foundation for Promotion of Cancer Research.

References

1. Matsuda A and Matsuda T: Time trends in stomach cancer mortality (1950-2008) in Japan, the USA and Europe based on the WHO mortality database. Jpn J Clin Oncol 41: 932-933, 2011.
2. Kamangar F, Dores GM and Anderson WF: Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J Clin Oncol 24: 2137-2150, 2006.
3. Yokota T, Ishiyama S, Saito T, *et al*: Lymph node metastasis as a significant prognostic factor in gastric cancer: a multiple logistic regression analysis. Scand J Gastroenterol 39: 380-384, 2004.
4. Chen JH, Wu CW, Lo SS, *et al*: Lymph node metastasis as a single predictor in patients with Borrmann type I gastric cancer. Hepatogastroenterology 54: 981-984, 2007.
5. Saito H, Fukumoto Y, Osaki T, *et al*: Prognostic significance of level and number of lymph node metastases in patients with gastric cancer. Ann Surg Oncol 14: 1688-1693, 2007.
6. Sano T and Aiko T: New Japanese classifications and treatment guidelines for gastric cancer: revision concepts and major revised points. Gastric Cancer 14: 97-100, 2011.
7. Bonenkamp JJ, Songun I, Hermans J, *et al*: Randomised comparison of morbidity after D1 and D2 dissection for gastric cancer in 996 Dutch patients. Lancet 345: 745-748, 1995.
8. Cuschieri A, Fayers P, Fielding J, *et al*: Postoperative morbidity and mortality after D1 and D2 resections for gastric cancer: preliminary results of the MRC randomised controlled surgical trial. The Surgical Cooperative Group. Lancet 347: 995-999, 1996.
9. Degiuli M, Sasako M, Ponti A, Soldati T, Danese F and Calvo F: Morbidity and mortality after D2 gastrectomy for gastric cancer: results of the Italian Gastric Cancer Study Group prospective multicenter surgical study. J Clin Oncol 16: 1490-1493, 1998.
10. Seevaratnam R, Cardoso R, McGregor C, *et al*: How useful is preoperative imaging for tumor, node, metastasis (TNM) staging of gastric cancer? A meta-analysis. Gastric Cancer: Aug 12, 2011 (E-pub ahead of print).
11. Ganpathi IS, So JB and Ho KY: Endoscopic ultrasonography for gastric cancer: does it influence treatment? Surg Endosc 20: 559-562, 2006.
12. Yoshioka T, Yamaguchi K, Kubota K, *et al*: Evaluation of ^{18}F-FDG PET in patients with advanced, metastatic, or recurrent gastric cancer. J Nucl Med 44: 690-699, 2003.
13. Ha TK, Choi YY, Song SY and Kwon SJ: F18-fluorodeoxyglucose-positron emission tomography and computed tomography is not accurate in preoperative staging of gastric cancer. J Korean Surg Soc 81: 104-110, 2011.
14. Natsugoe S, Mueller J, Stein HJ, Feith M, Hofler H and Siewert JR: Micrometastasis and tumor cell microinvolvement of lymph nodes from esophageal squamous cell carcinoma: frequency, associated tumor characteristics, and impact on prognosis. Cancer 83: 858-866, 1998.
15. Kojima N, Yonemura Y, Bando E, *et al*: Optimal extent of lymph node dissection for T1 gastric cancer, with special reference to the distribution of micrometastasis, and accuracy of preoperative diagnosis for wall invasion. Hepatogastroenterology 55: 1112-1117, 2008.
16. Motoyama K, Inoue H, Mimori K, *et al*: Clinicopathological and prognostic significance of PDCD4 and microRNA-21 in human gastric cancer. Int J Oncol 36: 1089-1095, 2010.
17. Tanaka M, Kitajima Y, Edakuni G, Sato S and Miyazaki K: Abnormal expression of E-cadherin and beta-catenin may be a molecular marker of submucosal invasion and lymph node metastasis in early gastric cancer. Br J Surg 89: 236-244, 2002.
18. Arigami T, Natsugoe S, Uenosono Y, *et al*: CCR7 and CXCR4 expression predicts lymph node status including micrometastasis in gastric cancer. Int J Oncol 35: 19-24, 2009.
19. Shen Z, Ye Y, Dong L, *et al*: Kindlin-2: a novel adhesion protein related to tumor invasion, lymph node metastasis, and patient outcome in gastric cancer. Am J Surg 203: 222-229, 2012.
20. Bibikova M, Barnes B, Tsan C, *et al*: High density DNA methylation array with single CpG site resolution. Genomics 98: 288-295, 2011.
21. Oka D, Yamashita S, Tomioka T, *et al*: The presence of aberrant DNA methylation in noncancerous esophageal mucosae in association with smoking history: a target for risk diagnosis and prevention of esophageal cancers. Cancer 115: 3412-3426, 2009.
22. Niwa T, Tsukamoto T, Toyoda T, *et al*: Inflammatory processes triggered by *Helicobacter pylori* infection cause aberrant DNA methylation in gastric epithelial cells. Cancer Res 70: 1430-1440, 2010.
23. Niwa T, Yamashita S, Tsukamoto T, *et al*: Whole-genome analyses of loss of heterozygosity and methylation analysis of four tumor-suppressor genes in N-methyl-N'-nitro-N-nitrosoguanidine-induced rat stomach carcinomas. Cancer Sci 96: 409-413, 2005.
24. Zheng Y, Cohen-Karni D, Xu D, *et al*: A unique family of Mrr-like modification-dependent restriction endonucleases. Nucleic Acids Res 38: 5527-5534, 2010.
25. Cohen-Karni D, Xu D, Apone L, *et al*: The MspJI family of modification-dependent restriction endonucleases for epigenetic studies. Proc Natl Acad Sci USA 108: 11040-11045, 2011.
26. Yokota T, Ishiyama S, Saito T, *et al*: Is tumor size a prognostic indicator for gastric carcinoma? Anticancer Res 22: 3673-3677, 2002.
27. Simon R, Radmacher MD, Dobbin K and McShane LM: Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. J Natl Cancer Inst 95: 14-18, 2003.

OPEN ACCESS Freely available online

PLOS ONE

DNA Methylation Profiles at Precancerous Stages Associated with Recurrence of Lung Adenocarcinoma

Takashi Sato[1,2], Eri Arai[1]*, Takashi Kohno[3], Koji Tsuta[4], Shun-ichi Watanabe[5], Kenzo Soejima[2], Tomoko Betsuyaku[2], Yae Kanai[1]

1 Division of Molecular Pathology, National Cancer Center Research Institute, Tokyo, Japan, **2** Division of Pulmonary Medicine, Department of Medicine, Keio University School of Medicine, Tokyo, Japan, **3** Division of Genome Biology, National Cancer Center Research Institute, Tokyo, Japan, **4** Department of Pathology and Clinical Laboratories, Pathology Division, National Cancer Center Hospital, Tokyo, Japan, **5** Department of Thoracic Oncology, Thoracic Surgery Division, National Cancer Center Hospital, Tokyo, Japan

Abstract

The aim of this study was to clarify the significance of DNA methylation alterations at precancerous stages of lung adenocarcinoma. Using single-CpG resolution Infinium array, genome-wide DNA methylation analysis was performed in 36 samples of normal lung tissue obtained from patients without any primary lung tumor, 145 samples of non-cancerous lung tissue (N) obtained from patients with lung adenocarcinomas, and 145 samples of tumorous tissue (T). Stepwise progression of DNA methylation alterations from normal lung tissue to non-cancerous lung tissue obtained from patients with lung adenocarcinomas, and then tumorous tissue samples, was observed at 3,270 CpG sites, suggesting that non-cancerous lung tissue obtained from patients with lung adenocarcinomas was at precancerous stages with DNA methylation alterations. At CpG sites of 2,083 genes, DNA methylation status in samples of non-cancerous lung tissue obtained from patients with lung adenocarcinomas was significantly correlated with recurrence after establishment of lung adenocarcinomas. Among such recurrence-related genes, 28 genes are normally unmethylated (average β-values based on Infinium assay in normal lung tissue samples was less than 0.2) and their DNA hypermethylation at precancerous stages was strengthened during progression to lung adenocarcinomas ($\Delta\beta_{T-N}>0.1$). Among these 28 genes, we focused on 6 for which implications in transcription regulation, apoptosis or cell adhesion had been reported. DNA hypermethylation of the *ADCY5, EVX1, GFRA1, PDE9A,* and *TBX20* genes resulted in reduced mRNA expression in tumorous tissue samples. 5-Aza-2′-deoxycytidine treatment of lung cancer cell lines restored the mRNA expression levels of these 5 genes. Reduced mRNA expression in tumorous tissue samples was significantly correlated with tumor aggressiveness. These data suggest that DNA methylation alterations at precancerous stages determine tumor aggressiveness and outcome through silencing of specific genes.

Citation: Sato T, Arai E, Kohno T, Tsuta K, Watanabe S-i, et al. (2013) DNA Methylation Profiles at Precancerous Stages Associated with Recurrence of Lung Adenocarcinoma. PLoS ONE 8(3): e59444. doi:10.1371/journal.pone.0059444

Editor: Bernard W Futscher, The University of Arizona, United States of America

Received September 25, 2012; **Accepted** February 14, 2013; **Published** March 27, 2013

Funding: This study was supported by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NiBio, 10–42, http://www.nibio.go.jp/part/promote/fundamental/doc/index.html) and partially supported by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labor and Welfare of Japan (2, http://www.mhlw.go.jp/bunya/kenkyuujigyou/hojokin-koubo14/09.html), National Cancer Center Research and Development Fund (23-A-11, http://www.ncc.go.jp/jp/about/rinri/kaihatsu/files/h24_ncc_research_list.pdf), and Grants-in-Aid for Scientific Research (B, 233900 90, http://www.jsps.go.jp/j-grantsinaid/) and for Young Scientists (B, 23790456, http://www.jsps.go.jp/j-grantsinaid/) from the Japan Society for the Promotion of Science (JSPS). National Cancer Center Biobank is supported by the National Cancer Center Research and Development Fund (23-A-1, http://www.ncc.go.jp/jp/about/rinri/kaihatsu/files/h24_ncc_research_list.pdf), Japan. T. Sato is an awardee of a research resident fellowship from the Foundation for Promotion of Cancer Research in Japan (http://www.fpcr.or.jp/index.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: earai@ncc.go.jp

Introduction

Lung adenocarcinoma (LADC) is increasingly recognized as a clinicopathologically and molecularly heterogeneous disease: frequent mutations of the *EGFR, KRAS, BRAF, TP53, ERBB2, PIK3CA* and *MET* genes and *EML4-ALK* fusions have been reported in LADCs [1,2]. In addition, recent whole-exome sequencing has revealed frequent mutation of the *CSMD3* gene [3]. However, the molecular background responsible for the clinicopathological diversity of LADCs is not yet fully understood.

As well as genetic abnormalities, epigenetic changes in human cancers have also been described [4–7]. In LADCs, silencing of the *RASSF1A, CDKN2A, RARβ, MGMT, APC, DAPK, FHIT* and *CDH13* genes due to DNA hypermethylation around their promoter regions has been frequently observed [8]. In addition, DNA methylation alterations are known to occur even at the early and precancerous stages of carcinogenesis in various organs [5–7,9]. For example, we have reported that DNA hypermethylation at the D17S5 locus, where the *HIC-1* tumor suppressor gene has been identified, is evident even in non-cancerous lung tissue obtained from patients with non-small cell lung cancers, and is correlated with smoking history [10]. Other researchers have also reported DNA hypermethylation of specific tumor-related genes at precancerous stages associated with cigarette smoking [8,11]. However, it has been unclear whether DNA methylation status is simply altered at precancerous stages or whether DNA methylation alterations at these stages actually result in gene expression alterations in established LADCs. Moreover, in organs other than

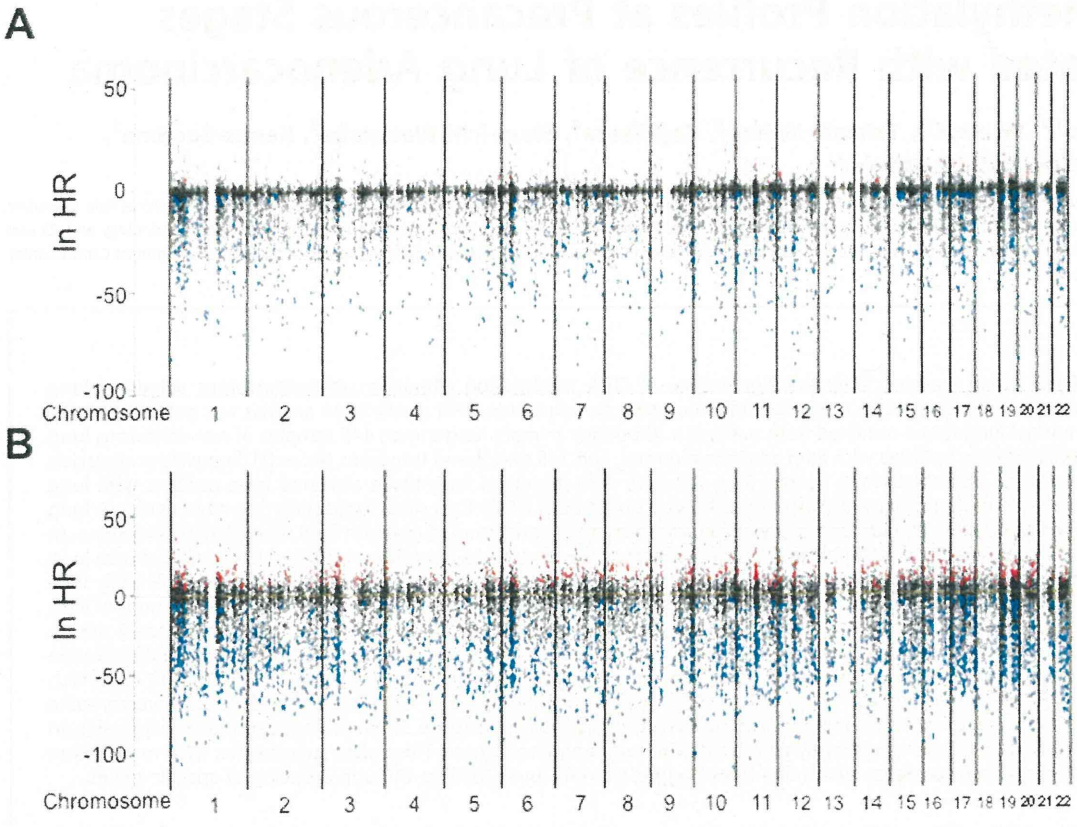

Figure 1. Hazard ratio (HR) obtained from the Cox regression model. Correlation between DNA methylation status (average β-values) and recurrence was examined in 145 samples of tumorous tissue (**A**) and 145 samples of the corresponding non-cancerous lung tissue (**B**) obtained from patients with lung adenocarcinomas who had undergone complete resection and had not received any adjuvant therapy after surgery. All of the examined 26,455 probes of the Infinium array are shown along the chromosomes. Red color means that higher β-values of the probes were observed in recurrence-positive patients than in recurrence-negative patients ($P<0.001$). Blue color means that lower β-values of the probes were observed in recurrence-positive patients than in recurrence-negative patients ($P<0.001$).
doi:10.1371/journal.pone.0059444.g001

the lung, it has been suggested that DNA methylation profiles at precancerous stages may determine tumor aggressiveness and outcome [12–14]. However, the clinicopathological impact of DNA methylation alterations at precancerous stages during lung carcinogenesis has not been clarified.

Recently, genome-wide DNA methylation analysis using the single-CpG resolution Infinium array has made it possible to interrogate 27,000 highly informative CpG sites, i.e. an average of two CpG sites within the proximal promoter regions of the transcription start sites of each of 14,475 consensus coding sequences in the National Center for Biotechnology Information Database, especially 3 to 20 CpG sites for more than 200 cancer-related and imprinted genes [15]. Although a few studies of lung cancers employing the Infinium assay have been reported [16,17], they did not focus on precancerous stages. In order to clarify the significance of DNA methylation alterations at precancerous stages of lung carcinogenesis, we performed the Infinium assay in association with mRNA expression and clinicopathological analyses of 36 samples of normal lung tissue (C) obtained from patients without any primary lung tumors, 145 samples of non-cancerous lung tissue (N) from patients with LADCs, and 145 corresponding samples of tissue from the tumors (T) themselves. Although the molecular classification of LADCs based on the results of the Infinium assay will be published elsewhere, we focused on specific genes methylated at precancerous stages in the present study.

Materials and Methods

Patients and Tissue Samples

The 145 paired samples of N and the corresponding T were obtained from patients with primary LADCs who underwent lung resection at the National Cancer Center Hospital, Japan, between December 1997 and March 2008. These patients had undergone complete resection and had not received any preoperative treatment or adjuvant therapy after surgery. Eighty-one patients were males and 64 were females with a median age of 61 years (range, 30–81 yr). Histological diagnosis and grading were based on the 2004 World Health Organization classification [18]. Recurrence was diagnosed by clinicians on the basis of physical examination and imaging modalities such as computed tomography, magnetic resonance imaging, scintigraphy or positron-emission tomography, and sometimes confirmed histopathologically by biopsy.

For comparison, 36 C samples were obtained from materials that had been surgically resected from patients without any primary lung tumor. Twenty-two of these patients were males and 14 were females, with a median age of 63 years (range, 27–83 yr).

Table 1. The 28 probes for which the average β-value in N samples (β_N) was higher in recurrence-positive patients than in recurrence-negative patients, for which the average β-value in C samples (β_C) was less than 0.2, and for which the average β-value in T samples minus that in corresponding N samples ($\Delta\beta_{T-N}$) was more than 0.1.

Target ID[a]	Chromosome	Position[b]	Gene symbol	P[c]	Adjusted P[d]
cg00516481	21	44,073,202	*PDE9A*	9.996×10^{-4}	1.193×10^{-2}
cg01295203	8	70,984,199	*PRDM14*	8.156×10^{-5}	3.236×10^{-3}
cg02008154	7	35,293,537	*TBX20*	7.188×10^{-4}	9.917×10^{-3}
cg02909790	6	26,271,587	*HIST1H3G*	2.896×10^{-6}	5.787×10^{-4}
cg03538436	12	117,799,370	*NOS1*	7.227×10^{-4}	9.944×10^{-3}
cg03963198	5	1,882,871	*IRX4*	5.133×10^{-4}	8.186×10^{-3}
cg06005396	19	590,541	*HCN2*	9.485×10^{-4}	1.158×10^{-2}
cg06269753	8	72,755,871	*MSC*	8.627×10^{-7}	3.426×10^{-4}
cg07651242	7	45,614,720	*ADCY1*	7.403×10^{-4}	1.010×10^{-2}
cg11612345	6	168,842,491	*SMOC2*	4.264×10^{-6}	6.753×10^{-4}
cg12087643	10	118,033,370	*GFRA1*	3.144×10^{-9}	2.489×10^{-5}
cg12265829	14	24,804,022	*ADCY4*	7.265×10^{-4}	9.984×10^{-3}
cg13262687	4	147,559,579	*POU4F2*	6.246×10^{-5}	2.795×10^{-3}
cg13449778	1	179,712,298	*FAM163A*	8.195×10^{-7}	3.426×10^{-4}
cg13878010	3	123,167,276	*ADCY5*	2.339×10^{-6}	5.220×10^{-4}
cg16254309	7	145,814,152	*CNTNAP2*	2.917×10^{-4}	6.240×10^{-3}
cg16387606	1	149,804,293	*HIST2H4A*	7.441×10^{-4}	1.011×10^{-2}
cg16604516	3	13,590,419	*FBLN2*	7.181×10^{-4}	9.916×10^{-3}
cg16652259	2	172,949,501	*DLX1*	4.175×10^{-4}	7.386×10^{-3}
cg17191178	3	157,824,217	*SHOX2*	8.872×10^{-5}	3.353×10^{-3}
cg18454685	17	48,639,239	*CACNA1G*	4.326×10^{-4}	7.487×10^{-3}
cg20286200	6	133,562,267	*EYA4*	1.541×10^{-9}	2.038×10^{-5}
cg21087137	12	75,728,469	*GLIPR1L1*	1.965×10^{-6}	4.821×10^{-4}
cg22461835	8	26,723,365	*ADRA1A*	7.297×10^{-4}	9.998×10^{-3}
cg23418591	20	57,090,317	*LOC149773*	4.075×10^{-4}	7.287×10^{-3}
cg25302419	5	11,904,015	*CTNND2*	5.149×10^{-4}	8.194×10^{-3}
cg25764191	10	105,037,215	*INA*	3.834×10^{-4}	7.026×10^{-3}
cg27626299	7	27,282,431	*EVX1*	2.164×10^{-4}	5.423×10^{-3}

[a]Probe ID for the Infinium HumanMethylation27 Bead Array (Illumina).
[b]National Center for Biotechnology Information database (Genome Build 37).
[c]Non-adjusted *P*-values and.
[d]Benjamini-Hochberg-adjusted *P*-values for the Cox regression model used for evaluation of correlation with recurrence.
doi:10.1371/journal.pone.0059444.t001

Thirty-five had undergone lung resection for metastatic lesions of primary cancers of the colon, rectum, kidney, urinary bladder, thyroid, breast, pancreas, ampulla of Vater and salivary gland, osteosarcoma, synovial sarcoma, leiomyosarcoma, rhabdmyosarcoma, liposarcoma, dermatofibrosarcoma, and myxofibrosarcoma. The remaining one patient had undergone chest wall resection for lipoma with removal of adjacent lung tissue. Histological observation confirmed that all of the C samples showed no remarkable histological abnormality and did not contain any contaminating tumor cells that had metastasized from organs other than the lung.

Tissue specimens were provided by the National Cancer Center Biobank, Japan. This study was approved by the Ethics Committee of the National Cancer Center, Tokyo, Japan, and was performed in accordance with the Declaration of Helsinki 1975. All patients included in this study provided written informed consent.

Cell Lines

The characteristics of the four lung cancer cell lines used in this study are summarized in Table S1.

Infinium Assay

Genomic DNA was extracted using a QIAamp DNA Mini kit (Qiagen, Valencia, CA, USA) and phenol-chloroform extraction followed by dialysis [19] from all tissue samples and cell lines, respectively. Five-hundred-nanogram aliquots of DNA were subjected to bisulfite conversion using an EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA). DNA methylation status at 27,578 CpG loci was examined at single-CpG resolution using the Infinium HumanMethylation27 Bead Array (Illumina, San Diego, CA, USA). After hybridization, the specifically hybridized DNA was fluorescence-labeled by a single-base extension reaction and detected using a BeadScan reader (Illumina) in accordance with the manufacturer's protocols. The data were then assembled using GenomeStudio methylation

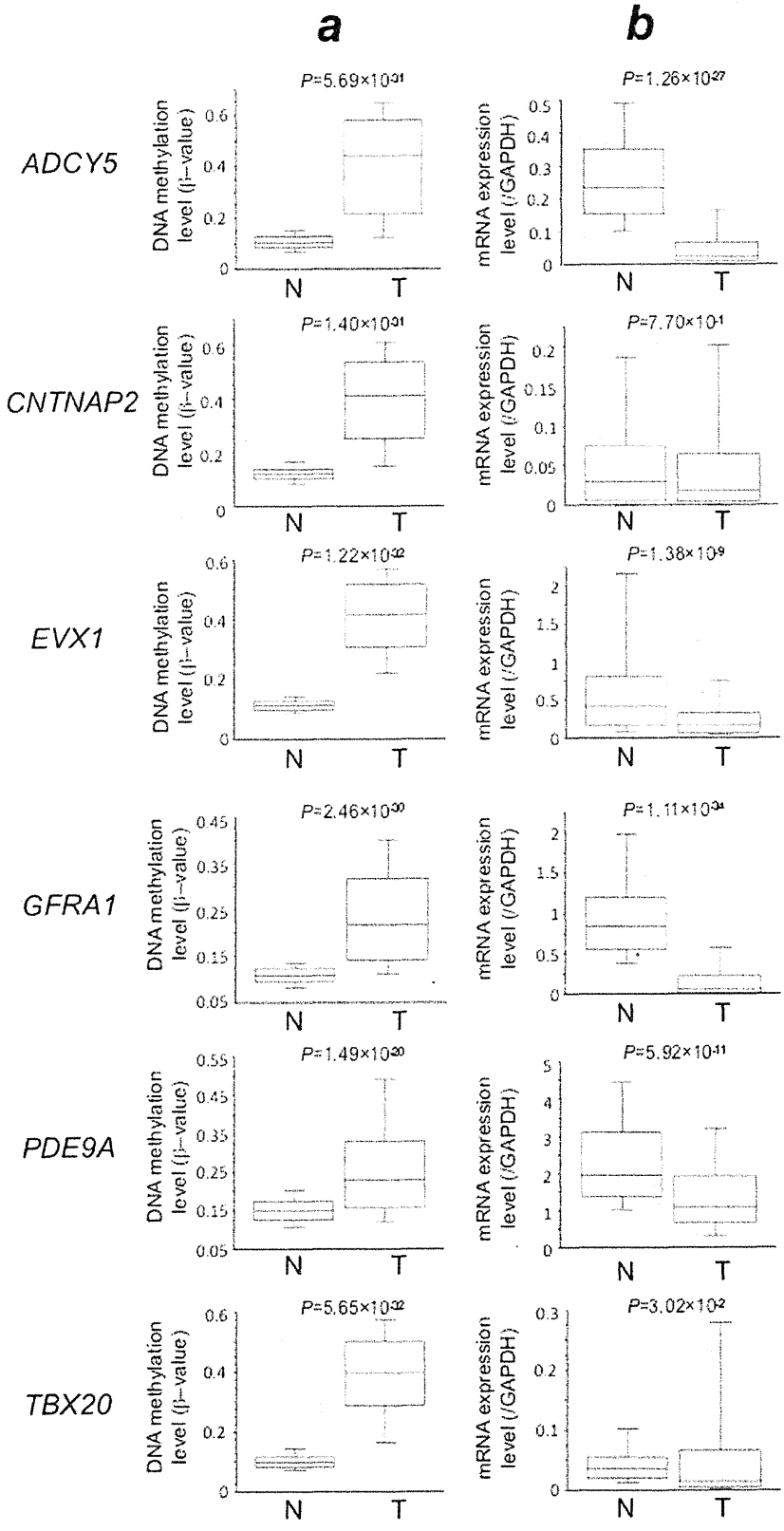
a
b
ADCY5
CNTNAP2
EVX1
GFRA1
PDE9A
TBX20
DNA methylation level (β-value)
mRNA expression level (/GAPDH)
P=5.69×10-31
P=1.26×10-27
P=1.40×10-31
P=7.70×10-1
P=1.22×10-32
P=1.38×10-9
P=2.46×10-30
P=1.11×10-34
P=1.49×10-30
P=5.92×10-11
P=5.65×10-32
P=3.02×10-2
N
T

Figure 2. Correlation between DNA methylation levels and mRNA expression levels. DNA methylation levels (average β-values) (**A**) and mRNA expression levels (**B**) for the *ADCY5, CNTNAP, EVX1, GFRA1, PDE9A* and *TBX20* genes in samples of non-cancerous lung tissue (N) from patients with lung adenocarcinomas and samples of the corresponding tumorous tissue (T) were determined by Infinium assay and quantitative real-time reverse transcription-PCR analysis, respectively. DNA methylation levels for all six genes were significantly higher in T samples than in N samples, and levels of expression of mRNAs for the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes were significantly lower in T samples than in N samples, although the reduction in the expression of the *CNTNAP2* gene did not reach statistical significance. These results suggested that DNA hypermethylation of the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes may result in reduced mRNA expression in tissue samples from the same cohort.
doi:10.1371/journal.pone.0059444.g002

software (Illumina). At each CpG site, the ratio of the fluorescence signal was measured using a methylated probe relative to the sum of the methylated and unmethylated probes, i.e. the so-called β-value, which ranges from 0.00 to 1.00, reflecting the methylation level of an individual CpG site.

Quantitative Real-time Reverse Transcription (RT)-PCR Analysis

Total RNA was extracted from 132 N and 151 T samples for which additional tissue specimens were available and cell lines using TRIzol reagent (Life Technologies, Carlsbad, CA, USA) in accordance with the manufacturer's instructions. cDNA was synthesized from total RNA with random primers using Super-Script III Reverse Transcriptase (Life Technologies) and pre-amplified using TaqMan PreAmp Master Mix (Life Technologies).

To evaluate mRNA expression levels, fluorescence-labeled locked nucleic acid hydrolysis probes were selected from the Universal Probe Library collection (Roche Applied Science, Mannheim, Germany) and specific PCR primers yielding intron-spanning amplicons were designed using ProbeFinder assay design software (https://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id = UP030000). The probe ID and primer sequences are summarized in Table S2. Quantitative real-time PCR was performed using TaqMan Universal Master Mix II (Life Technologies) and the relative standard curve method in the BioMark HD System (Fluidigm, South San Francisco, CA, USA). Ct values were normalized to that of *GAPDH* in the same sample. All assays were performed in triplicate.

5-aza-2′-deoxycytidine (5-aza-dC) Treatment

A549, PC9, VMRC-LCD and EBC-1 cells were seeded at a density of 9×10^5 cells per 15-cm dish on day 0 and then allowed to attach for a 24-h period. Then, 5-aza-dC (Sigma-Aldrich, St. Louis, MO, USA) was added to a final concentration of 1 μM. Cells were passaged at a subculture ratio of 1:2 on day 3. At 24 h after replating, 5-aza-dC was added again to the same final concentration. Since toxicity had been obvious during preliminary experiments, the final concentration of 5-aza-dC was reduced to 0.1 μM for EBC-1 cells. Genomic DNA and total RNA were extracted from all cells on days 3 and 6.

Statistics

In the Infinium assay, all CpG sites on chromosomes X and Y were excluded, to avoid any gender-specific methylation bias. The call proportions (*P*-values for detection of signals above the background <0.01) for 31 probes (shown in Table S3) in all of the tissue samples examined were less than 90%. Since such a low proportion may be attributable to polymorphism at the probe CpG sites, these 31 probes were excluded from the present assay, leaving a final total of 26,455 autosomal CpG sites.

Infinium probes showing ordered differences from 36 C to 145 N, and then to the 145 T samples themselves, were examined by the cumulative logit model adjusted by sex, age and experimental batch ($P<1\times10^{-14}$). Correlations between β-values in N and T samples and recurrence were assessed by the Cox regression model adjusted by sex, age and experimental batch ($P<0.001$). Benjamini-Hochberg correction was performed to adjust for multiple testing. Differences of β-values and mRNA expression levels between N and T samples were examined by Mann-Whitney U test. Correlations between mRNA expression levels and clinicopathological parameters were assayed by analysis of variance between groups (ANOVA) and Welch's T-test: after adjusted Bonferroni correction to adjust for multiple testing, corrected *P* values of <0.05 were considered to be significant. All statistical analyses were performed using programming language R.

Results

DNA Methylation Profiles Associated with Recurrence are Established at Precancerous Stages

DNA methylation levels of CpG sites of the *CNTNAP2, EVX1, GFRA1, PDE9A* and *TBX20* genes based on the Infinium assay were clearly verified using the quantitative pyrosequencing method (Figure S1), indicating the reliability of the Infinium assay. The cumulative logit model ($P<1\times10^{-14}$) revealed ordered progression of DNA methylation alterations from C to N, and then to T samples, on 3,270 probes; DNA methylation alterations occurred even in N samples compared to C samples, and such DNA methylation alterations were inherited by, or strengthened in, T samples, indicating that Ns were at precancerous stages with DNA methylation alterations. Among the 3,270 probes, the number showing average β-values in T samples minus average β-values in C samples ($\Delta\beta_{T\text{-}C}$) of >0.1 and <−0.1 were 1,209 and 1,056, respectively. Thus, when we defined differentially methylated probes as probes showing a $\Delta\beta_{T\text{-}C}$ value of >0.1 or <−0.1, the false positivity rate by the cumulative logit model was 4.3%.

Correlations between DNA methylation status and recurrence were examined using the Cox regression model ($P<0.001$) for 145 patients. In T samples, DNA methylation status on 944 probes for the 916 genes was significantly correlated with recurrence: on 87 probes (red dots in Figure 1A), higher β-values were observed in recurrence-positive patients than in recurrence-negative patients, whereas lower β-values on 857 probes (blue dots in Figure 1A) were observed in recurrence-positive patients. Surprisingly, even in N samples, the DNA methylation status on 2,215 probes for the 2,083 genes was significantly correlated with recurrence: on 425 probes (red dots in Figure 1B), higher β-values were observed in recurrence-positive patients than in recurrence-negative patients, whereas lower β-values on 1,790 probes (blue dots in Figure 1B) were observed in recurrence-positive patients.

In order to identify recurrence-related genes that are normally unmethylated and for which DNA hypermethylation at precancerous stages is strengthened in the established LADCs, among the 425 probes (red dots in Figure 1B), we initially focused on 28 probes for which the average β-values in C samples (β_C) were less than 0.2 and the average β-values in T samples minus that in the corresponding N samples ($\Delta\beta_{T\text{-}N}$) were more than 0.1 (Table 1).