difference analysis and its application to cancer research. *Ann NY Acad Sci* 983, 131–141.
3. Ushijima, T. (2005) Detection and interpretation of altered methylation patterns in cancer cells. *Nat Rev Cancer* 5, 223–231.
4. Takai, D., Yagi, Y., Wakazono, K., et al. (2001) Silencing of *HTR1B* and reduced expression of *EDN1* in human lung cancers, revealed by methylation-sensitive representational difference analysis. *Oncogene* 20, 7505–7513.
5. Kaneda, A., Kaminishi, M., Nakanishi, Y., et al. (2002) Reduced expression of the *insulin-induced protein 1* and *p41 Arp2/3 complex* genes in human gastric cancers. *Int J Cancer* 100, 57–62.
6. Miyamoto, K., Asada, K., Fukutomi, T., et al. (2003) Methylation-associated silencing of heparan sulfate D-glucosaminyl 3-*O*-sulfotransferase-2 (3-OST-2) in human breast, colon, lung and pancreatic cancers. *Oncogene* 22, 274–280.
7. Miyamoto, K., Fukutomi, T., Akashi-Tanaka, S., et al. (2005) Identification of 20 genes aberrantly methylated in human breast cancers. *Int J Cancer* 116, 407–414.
8. Hagihara, A., Miyamoto, K., Furuta, J., et al. (2004) Identification of 27 5′ CpG islands aberrantly methylated and 13 genes silenced in human pancreatic cancers. *Oncogene* 23, 8705–8710.
9. Abe, M., Ohira, M., Kaneda, A., et al. (2005) CpG island methylator phenotype is a strong determinant of poor prognosis in neuroblastomas. *Cancer Res* 65, 828–834.
10. Furuta, J., Nobeyama, Y., Umebayashi, Y., et al. (2006) Silencing of peroxiredoxin 2 and aberrant methylation of 33 CpG islands in putative promoter regions in human malignant melanomas. *Cancer Res* 66, 6080–6086.
11. Kaneda, A., Wakazono, K., Tsukamoto, T., et al. (2004) *Lysyl oxidase* is a tumor suppressor gene inactivated by methylation and loss of heterozygosity in human gastric cancers. *Cancer Res* 64, 6410–6415.
12. Abe, M., Westermann, F., Nakagawara, A., et al. (2007) Marked and independent prognostic significance of the CpG island methylator phenotype in neuroblastomas. *Cancer Lett* 247, 253–258.
13. Lisitsyn, N., Lisitsyn, N., Wigler, M. (1993) Cloning the differences between two complex genomes. *Science* 259, 946–951.
14. Henke, W., Herdel, K., Jung, K., et al. (1997) Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res* 25, 3957–3958.

Original Article

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

Daiji Oka, MD[1,2]; Satoshi Yamashita, PhD[1]; Tadashi Tomioka, BTech[1]; Yukihiro Nakanishi, MD, PhD[3]; Hoichi Kato, MD, PhD[4]; Michio Kaminishi, MD, PhD[2]; and Toshikazu Ushijima, MD, PhD[1]

BACKGROUND: Esophageal squamous cell carcinomas (ESCCs) tend to have multiple primary lesions, and it is believed that they arise from background mucosae with accumulation of genetic/epigenetic alterations. In this study, the objective was to elucidate the effects of smoking and drinking on the accumulation of epigenetic alterations in background mucosae. **METHODS:** Genes that are silenced in human ESCCs were searched for by treating 3 ESCC cell lines with the demethylating agent, 5-aza-2′-deoxycytidine and performing oligonucleotide microarrays. Methylation levels were analyzed by quantitative methylation-specific polymerase chain reaction analysis of 60 ESCCs and their corresponding background mucosae. **RESULTS:** Forty-seven genes were identified as methylation-silenced in at least 1 of the 3 ESCC cell lines, and 14 of those genes (claudin 6 [*CLDN6*]; G protein-coupled receptor 158 [*GPR158*]; homeobox A9 [*HOXA9*]; metallothionein 1M [*MT1M*]; neurofilament, heavy polypeptide 200 kDa [*NEFH*]; plakophilin 1 [*PKP1*]; protein phosphatase 1, regulatory [inhibitor] subunit 14A [*PPP1R14A*]; pyrin domain and caspase recruitment domain containing [*PYCARD*]; R-spondin family, member 4 [*RSPO4*]; testis-specific protein, Y-encoded-like 5 [*TSPYL5*]; ubiquitin carboxyl-terminal esterase L1 [*UCHL1*]; zinc-finger protein 42 homolog [*ZFP42*]; zinc-finger protein interacting with K protein 1 homolog [*ZIK1*]; and zinc-finger and SCAN domain containing 18 [*ZSCAN18*]) were used as markers. In the background mucosae, methylation levels of 5 genes (*HOXA9*, *MT1M*, *NEFH*, *RSPO4*, and *UCHL1*) had significant correlations with smoking duration ($\rho = .268$; $P = .044$; $\rho = .405$; $P = .002$; $\rho = .285$; $P = .032$; $\rho = .300$; $P = .024$; and $\rho = .437$; $P = .001$, respectively). In contrast, an inverse correlation between *PYCARD* methylation levels and alcohol intake was observed ($\rho = -.334$, $P = .025$) among individuals with the inactive aldehyde dehydrogenase 2 (*ALDH2*) genotype. **CONCLUSIONS:** The current results suggested that ESCCs developed from an epigenetic field for cancerization, which was induced by exposure to carcinogenic factors, such as tobacco smoking. The epigenetic field defect will be a novel target for risk diagnosis and prevention of ESCCs. **Cancer 2009;115:3412–26.**

KEY WORDS: epigenetics, DNA methylation, esophageal cancer, tobacco smoking, alcohol drinking.

Corresponding author: Toshikazu Ushijima, MD, PhD, Carcinogenesis Division, National Cancer Center Research Institute, 1-1 Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan; Fax: (011) 81-3-5565-1753; tushijim@ncc.go.jp

[1]Carcinogenesis Division, National Cancer Center Research Institute, Tokyo, Japan; [2]Department of Gastrointestinal Surgery, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [3]Pathology Division, National Cancer Center Research Institute, Tokyo, Japan; [4]Esophageal Surgery Division, National Cancer Center Hospital, Tokyo, Japan

Received: August 26, 2008; **Revised:** December 9, 2008; **Accepted:** January 7, 2009

Published online: May 26, 2009 © 2009 American Cancer Society

DOI: 10.1002/cncr.24394, www.interscience.wiley.com

Esophageal cancer is 1 of the most lethal cancers and imposes reduced quality of life even in patients who receive curative treatment.[1] Squamous cell carcinoma (SCC) and adenocarcinoma are 2 major histologic types of esophageal cancer, and SCC is the predominant histologic type in Asian countries.[2] Most patients with esophageal SCC (ESCC) have a history of chronic smoking and/or heavy drinking, and these are established risk factors for ESCC.[1] With regard to the interactions between smoking and drinking, controversial reports are available; however, to our knowledge, to date, the combined risk for ESCC has not been clarified.[3-5]

Patients with ESCC often have multiple primary lesions,[6] and the frequency of multiple occurrence reaches up to 30%.[7,8] In addition, dysplastic lesions frequently are observed in the background mucosae surrounding ESCCs.[9] The incidence of multiple occurrences of ESCC and/or dysplastic lesions reportedly is high in heavy smokers and heavy drinkers.[7,10,11] It is believed that both tobacco smoking and alcohol drinking can induce genetic/epigenetic alterations in esophageal mucosal cells and that genetic/epigenetic alterations already have accumulated in the normal-appearing esophageal mucosae of individuals who have a long history of smoking and drinking, forming a "field for cancerization."[12]

DNA methylation of a CpG island (CGI) in a promoter region causes silencing of its downstream gene and is known as a major epigenetic mechanism for inactivation of tumor-suppressor genes.[13] In ESCCs, various tumor-suppressor genes, such as adenomatous polyposis coli (*APC*), cadherin 1 (*CDH1*), cyclin-dependent kinase inhibitor 2A (*CDKN2A*), retinoic acid receptor β (*RARβ*), Ras association domain family member 1-isoform A (*RASSF1A*), and ubiquitin carboxyl-terminal esterase L1 (*UCHL1*), reportedly are methylated.[14,15] In contrast to the deep involvement of aberrant DNA methylation in ESCCs, its inducers in the esophagus have been poorly clarified, except for induction methylation of the fragile histidine triad gene *FHIT* by nicotine, a tobacco component, in human esophageal epithelial cell lines.[16] Tobacco smoking reportedly induced methylation of several genes in bronchial and oral epithelium,[17–20] but no information was available in esophageal epithelial cells.

In the current study, our objective was to clarify the effects of tobacco smoking and alcohol drinking on the induction of DNA methylation in esophageal mucosae. The mucosae are not clonal, and the degree of methylation is correlated with the power of methylation induction. Because different genes have different susceptibility to methylation induction,[21] first, we screened genes that were silenced in ESCCs by treating ESCC cell lines with a demethylating agent and analyzing the genes that were up-regulated. Then, we quantified the methylation levels of the "marker" genes and 4 tumor-suppressor genes in ESCCs and in noncancerous esophageal mucosae by using quantitative methylation-specific polymerase chain reaction (PCR) (MSP) analysis.

MATERIALS AND METHODS

Cell Lines and 5-Aza-2′-deoxycitidine Treatment

The ESCC cell lines KYSE30, KYSE220, and KYSE270, which were derived from well differentiated, moderately differentiated, and well differentiated ESCCs, respectively,[22] were purchased from Health Science Research Resources Bank (Osaka, Japan). For 5-aza-2′-deoxycitidine (5-Aza-dC) treatment, 2×10^5 KYSE30 cells per 10-cm dish, 5×10^5 KYSE220 cells per 10-cm dish, and 4×10^5 KYSE270 cells per 10-cm dish were seeded on Day 0 and exposed to freshly prepared 5-Aza-dC (Sigma, St. Louis, MO) for 24 hours on Days 1 and 3. The concentration of 5-Aza-dC was 1 μM for KYSE30 and KYSE220 cells and 0.3 μM for KYSE270. After each treatment, the cells were placed in fresh medium and harvested on Day 4. Total RNA was extracted using Isogen (Nippon Gene, Tokyo, Japan) and purified using an RNeasy Mini kit (Qiagen, Valencia, Calif). DNA methyltransferase 1 depletion was confirmed by Western blot analysis of the whole cell extract.[23]

Oligonucleotide Microarray Analysis and Database Search

Oligonucleotide microarray analysis was performed using the GeneChip Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, Calif) with 47,400 probe sets from 38,500 genes and GeneChip Operating Software. The signal intensities were normalized so that the average

of all genes on a GeneChip was 500. Database searches were performed at a GenBank website, and we searched for CGIs by using a modified criteria of Takai and Jones[24]: 1) CpG score ≥0.60, 2) guanine (G) and cytosine (C) content ≥50%, and 3) length ≥500 base pairs.

Patients and Tissue Samples

Sixty primary ESCC specimens and their paired noncancerous background mucosae were collected from patients who underwent esophagectomy and who were diagnosed histologically with invasive SCC at the National Cancer Center Hospital, Tokyo, Japan (51 men and 9 women; average age, 63.3 years [range, 41-83 years]). Informed consent was obtained from all patients. Background mucosae were resected from areas stained by iodine that were considered histologically normal.[25] All samples were stored in RNAlater (Applied Biosystems, Foster City, Calif) at −80°C until the extraction of genomic DNA. Genomic DNA was extracted by using the phenol-chloroform method.

History of tobacco smoking and alcohol drinking was obtained from 57 patients and 55 patients, respectively, in interviews with the patients by medical staff. Mean daily alcohol intake was calculated, converting a cup of sake (180 mL), a cup of shochu (180 mL), a single finger of whisky (30 mL), a bottle of wine (750 mL), and a bottle of beer (633 mL) into 27 g, 45 g, 12 g, 105 g, and 25 g of alcohol, respectively.

Aldehyde Dehydrogenase Genotyping

Aldehyde dehydrogenase 2 (*ALDH2*) genotyping of each patient was performed on genomic DNA extracted from background mucosae by PCR-restriction fragment length polymorphism.[26] Exon 12 of the *ALDH2* gene was amplified with forward primer (5′-CAAATTACAGGGTCAACTG CT-3′) and reverse primer (5′-CCACACTCACAGTTTTC TCTT-3′). The PCR products were digested with *Eam*1104I (*Ksp*632I) (Fermentas International Inc., Burlington, Ontario, Canada) and electrophoresed on ethidium bromide-stained 2% NuSieve agarose gels.

Bisulfite Modification and Semiquantitative Methylation-specific PCR

DNA was digested by *Bam*HI, and 1 μg of the digested DNA was denatured in 0.3 N NaOH at 37°C for 15 minutes. The samples underwent 15 cycles of 30-second denaturation at 95°C and a 15-minute incubation at 50°C in 3.1 N sodium bisulfite (pH 5.0) and 0.5 mM hydroquinone. The samples were desalted with Zymo-Spin IC Columns (Zymo Research, Orange, Calif), desulfonated in 0.3 N NaOH, and dissolved in 40 μL to 160 μL of Tris-ethylene diamine tetraacetic acid buffer.

MSP was performed with a primer set specific to the methylated (M) or unmethylated (U) sequence using 12.5 ng of the sodium bisulfite-treated DNA and an iCycler Thermal Cycler (Bio-Rad Laboratories, Hercules, Calif). The primers used are listed in Table 1. DNA methylated with *Sss*I methylase (New England Biolabs, Beverly, Mass) and DNA amplified twice by a GenomiPhi DNA amplification kit (GE Healthcare Bio-Sciences, Buckinghamshire, England) were used as fully methylated and unmethylated control DNA, respectively. For semiquantitative MSP, "10%-methylated DNA" was prepared by mixing 10% of methylated DNA and 90% of fully unmethylated control DNA after bisulfite modification.

Quantitative Methylation-specific PCR and Calculation of Methylation Levels

Real-time MSP was performed using 50 ng of the sodium bisulfite-treated DNA, SYBR Green I (BioWhittaker Molecular Applications, Rockland, Md), and an iCycler Thermal Cycler (Bio-Rad Laboratories). Methylated and unmethylated standard DNA was prepared by cloning PCR products from fully methylated and unmethylated control DNA, respectively, into the pGEM-T Easy vector (Promega, Madison, Wis) or by purifying PCR products using the Wizard SV Gel and PCR clean-up system (Promega). The numbers of methylated and unmethylated molecules in a sample were determined by comparing its amplification with that of methylated and unmethylated standard DNA, respectively, that contained known numbers of molecules (10 to 10^6 molecules).

On the basis of the numbers of methylated and unmethylated molecules for a genomic region in a sample, methylation levels were calculated as the fraction of methylated molecules in the total number of DNA molecules (the number of methylated molecules + the number of

Table 1. Primers for Qualitative or Quantitative Methylation-specific Polymerase Chain Reaction

Gene Symbol	Status	Primer Sequence: Forward	Primer Sequence: Reverse	Annealing Temperature, °C: Qualitative	Annealing Temperature, °C: Quantitative	No. of Cycles*
CLDN6	M	ATAAGTTTGGGATTCGTAC	ATCTTAAAAAAACGATAACG	52	54	35
	U	TTGATAAGTTTGGGATTTGTAT	CAAAAATCTTAAAAAAACAATAACA	52	54	35
GPR158	M	GTAATTTTTAGTCGGTTTTTC	GAATAACTAAACTACCGTCG	56	52	35
	U	TGTAATTTTTAGTTGGTTTTTTGT	CCAAATAACTAAACTACCATCA	56	58	35
HOXA9	M	TCGGATTATTAATAGCGTGC	ATCACCTAATAAATTACCGACG	60	58	33
	U	TAATAGTGTGTGGAGTGATTTATGT	CAATCACCTAATAAATTACCAACA	60	60	33
MT1M	M	GTATAGTTTTTTTCGCGTAAATTC	AACCCAACATAAATACCGAACG	59	55	32
	U	TTATTTGGTGTATAGTTTTTTTTGT	TAAACCCAACATAAATACCAAACA	56	52	38
NEFH	M	TTAAGGGGTTGGATTCGGTC	CGAAACGAAACGAAAAACACTACG	61	58	35
	U	GTTAAGGGGTTGGATTTGGTT	CCAAAACAAAACAAAAAACACTACA	61	61	35
PKP1	M	TTTTGTTTTTAAGAGCGTTGGTTTC	ACCCACTCCACCGAACCG	58	62	35
	U	TTTTGTTTTTAAGAGTGTTGGTTTT	CACTCCACCAAACCACACA	58	61	34
PPP1R14A	M	ATTTCGGTTCGGGAGTTTTC	TCGACTTAAAACACGCAATCG	58	60	35
	U	GGAGTTTTTGATGTAGGGATTT	TCAACTTAAAACACACAATCATACA	58	57	35
PYCARD	M	CGGGGAATCGCGGAGGTTTC	AATAAAACCCGAAAAAAAACCG	55	57	35
	U	GGTTTGGGGAATTGTGGAGGTTTT	ATCACACCCTCCAACTAACCTACA	55	60	33
RSPO4	M	CGTTAGGGTAGTGTTTCGGTTTC	TACTATAAACGCGCCGAACG	57	60	35
	U	TTTTTTTTGTTAGGGTAGTGTTTT	ATAAACACACCAAACACATCCA	57	58	35
TSPYL5	M	GGGTCGTTTTTTGCGTAGTC	GTCACGAACGTACAACTATACCG	62	60	35
	U	GGTTGTTTTTTGTGTAGTTGTAGT	CATCACAAACATACAACTATACCA	62	62	35
UCHL1	M	TCGTATTTATTTGGTCGCGATC	CTATAAAACGCCGACCAAACG	62	64	35
	U	GGTTTGTATTTATTTGGTTGTGATT	CAACTATAAAACACCAACCAAACA	60	61	33
ZFP42	M	GCGTCGTTTAGGTGTTAGGC	AAAACGTAACCGACCCCG	62	57	35
	U	GTGTTTTAGGGTGGGGTTGGTTAT	AAACCCACCCTCCAAACTAACACA	62	63	35
ZIK1	M	GTTTGAGGTGACGTTGGGC	GACCCTTTTCTCAACGCGA	64	62	35
	U	TTTGAGGTGATGTTGGGTG	AACAACCCTTTTCTCAACACA	60	59	35
ZSCAN18	M	TTTTTTGTTCGTTTCGGTGC	GATAACGACCGACAAACTACG	59	62	35
	U	TGTTATGGTTTTTTTGTTTGTTTT	CTACACACTAAACCTCACCACA	59	60	35
CDH1	M	TAGGTTTTAGTGAGTTATCGGC	AAACGAAACTAACGACCCG		59	
	U	ATTTTAGGTTAGAGGGTTATTGTG	ATAAACCCCAAAAACACCA		59	
CDKN2A	M	TTGGTAGTTAGGAAGGTTGTATCGC	TCCCTACTCCCAACCGCG		62	
	U	GGTAGTTAGGAAGGTTGTATTGT	TCCCTACTCCCAACCACA		61	
MLH1	M	CGTTAAGTATTTTTTTCGTTTTGC	TCCGCTCTTCCTATTAATTCG		59	
	U	AGTGTTAAGTATTTTTTTTGTTTTGT	CTATCCACTCTTCCTATTAATTCA		56	
RASSF1A	M	GTCGTCGTTGTGGTCGTTC	AACCCGAAAACGAAACTAAACG		62	
	U	TGTTGTTGTTGTGGTTGTTT	AAAACAAAACTAAACACACTCTCA		62	

CLDN6 indicates claudin 6; M, specific to methylated DNA; A, adenine; T, thymine; G, guanine; C, cytosine; U, specific to unmethylated DNA; *GPR158*, G protein-coupled receptor 158; *HOXA9*, homeobox A9; *MT1M*, metallothionein 1M; *NEFH*, neurofilament, heavy polypeptide 200 kDa; *PKP1*, plakophilin 1; *PPP1R14A*, protein phosphatase 1, regulatory (inhibitor) subunit 14A; *PYCARD*, pyrin domain and caspase recruitment domain containing; *RSPO4*, R-spondin family, member 4; *TSPYL5*, testis-specific protein, Y-encoded–like 5; *UCHL1*, ubiquitin carboxyl-terminal esterase L1; *ZFP42*, zinc-finger protein 42 homolog; *ZIK1*, zinc-finger protein interacting with K protein 1 homolog; *ZSCAN18*, zinc-finger and SCAN domain containing 18; *CDH1*, cadherin 1; *CDKN2A*, cyclin-dependent kinase inhibitor 2A; *MLH1*, MutL homolog 1; *RASSF1A*, Ras association domain family member 1, isoform A.
*The number of polymerase chain reaction (PCR) cycles for qualitative methylation-specific PCR.

unmethylated molecules). For each gene, the deviation value of a methylation level in the background mucosa of a case was calculated.

Statistical Analysis

Correlations between methylation levels in background mucosae and risk factors of patients were analyzed by Spearman rank-order correlation coefficients (ρ). Methylation levels in groups with different clinicopathologic characters were compared using the Kruskal-Wallis test and the Mann-Whitney *U* test. Correlations between methylation frequencies in ESCCs and methylation levels in background mucosae (mean deviation values) were analyzed using Pearson correlation coefficients. The statistical calculations were conducted with SPSS software (13.0J; SPSS Inc., Chicago, Ill).

RESULTS

Genes Up-regulated by 5-Aza-deoxycitidine Treatment and Their Methylation Analysis

Three ESCC cell lines (KYSE30, KYSE220, and KYSE270) were treated with 5-Aza-dC, and changes in gene expression were analyzed by oligonucleotide microarrays. We searched for genes that 1) were up-regulated above a threshold (8-fold, 16-fold, or 32-fold), 2) had signal intensities of $\leq$100 before the treatment and >100 after the treatment, 3) were not located on chromosome X, and 4) had CGIs 5′ upstream of their putative transcription start sites. The higher threshold we adopted, the fewer genes were up-regulated. Because the objective of the screening was to isolate marker genes for exposure to tobacco smoking and alcohol drinking, we adopted a cutoff value of 16-fold so that a manageable number of candidate genes (72 candidate genes in total) would be obtained.

The methylation status of the CGIs at the putative transcription start sites of the 72 genes was analyzed by MSP in the KYSE30, KYSE220, and KYSE270 cell lines. Forty-seven genes were confirmed as completely methylated at least in 1 of the 3 cell lines and were considered to be methylation silenced. Then, their methylation status was analyzed in 6 primary ESCCs and their background mucosae by using semiquantitative MSP. Thirty-nine of those genes were methylated in at least 1 primary ESCC, but 15 genes were methylated too heavily in the background mucosae ($\geq$10% in all 6 samples). Therefore, the remaining 24 genes were considered the most informative.

Methylation Quantification in ESCCs and Background Mucosae

Among the 24 genes, primers for quantitative MSP were designed successfully for 14 genes: claudin 6 (*CLDN6*); G protein-coupled receptor 158 (*GPR158*); homeobox A9 (*HOXA9*); metallothionein 1M (*MT1M*); neurofilament, heavy polypeptide 200 kDa (*NEFH*); plakophilin 1 (*PKP1*); protein phosphatase 1, regulatory (inhibitor) subunit 14A (*PPP1R14A*); pyrin domain and caspase recruitment domain containing (*PYCARD*); R-spondin family, member 4 (*RSPO4*); testis-specific protein, Y-encoded–like 5 (*TSPYL5*); *UCHL1*; zinc-finger protein 42 homolog (*ZFP42*); zinc-finger protein interacting with K protein 1 homolog (*ZIK1*); and zinc-finger and SCAN domain containing 18 (*ZSCAN18*). Methylation levels of these 14 genes and of 4 tumor-suppressor genes (*CDH1*, *CDKN2A*, mutL homolog 1 [*MLH1*], and *RASSF1A*), 3 of which reportedly are silenced in ESCCs,[27–29] were analyzed in 60 ESCCs and their paired noncancerous background mucosae. Similar to results from an examination of gastric cancers and their background mucosae,[30] distributions of methylation levels revealed different patterns between ESCCs and their background mucosae (Fig. 1).

In the background mucosae, different genes had different methylation distribution from the viewpoints of the fraction of methylation-positive samples and their absolute methylation levels. The first group of genes (*CLDN6*, *CDKN2A*, *MLH1*, and *RASSF1A*) had no methylation. The second group of genes (*PYCARD*, *RSPO4*, *TSPYL5*, *ZIK1*, and *ZSCAN18*) was methylated in a small number of samples, and the levels were low ($\leq$3%). The third group of genes (*UCHL1* only) was methylated also in only a small number of samples, but the methylation level was high in some samples, reaching up to 20%. The fourth group of genes (*GPR158*, *HOXA9*, *MT1M*, *NEFH*, *PKP1*, *PPP1R14A*, *ZFP42*, and *CDH1*) was methylated in a large number of samples, and the methylation levels revealed unimodal distribution with various highest values ranging from 6.8% (*MT1M*) to 25.9% (*ZFP42*).

In the 60 ESCCs, 12 marker genes (*CLDN6*, *GPR158*, *HOXA9*, *MT1M*, *NEFH*, *PPP1R14A*, *RSPO4*, *TSPYL5*, *UCHL1*, *ZIK1*, *ZFP42*, and *ZSCAN18*) and 2 tumor-suppressor genes (*CDH1* and *RASSF1A*) were methylated in 11 to 49 ESCCs and in 3 to 4 ESCCs, respectively, with a cutoff threshold of 6%.[30,31] Two marker genes (*PKP1* and *PYCARD*) and 2 tumor-suppressor genes (*CDKN2A* and *MLH1*) were not methylated. The distribution of methylation levels in methylation-positive ESCCs was much broader than the levels in background mucosae. Also, 11 of 12 marker genes (*CLDN6*, *GPR158*, *HOXA9*, *MT1M*, *NEFH*, *PPP1R14A*, *RSPO4*, *TSPYL5*, *UCHL1*, *ZIK1*, and *ZSCAN18*) and 2 tumor-suppressor genes (*CDH1* and *RASSF1A*) had large numbers of methylation-negative samples at the same time.

Correlations Between Methylation Levels in the Background Mucosae and Exposure Levels to ESCC Risk Factors

Next, we examined correlations between methylation levels in the background mucosae and risk factors for ESCCs; age, smoking duration, and mean daily alcohol intake (Table 2). From the initial 14 genes, *CLDN6*, which did

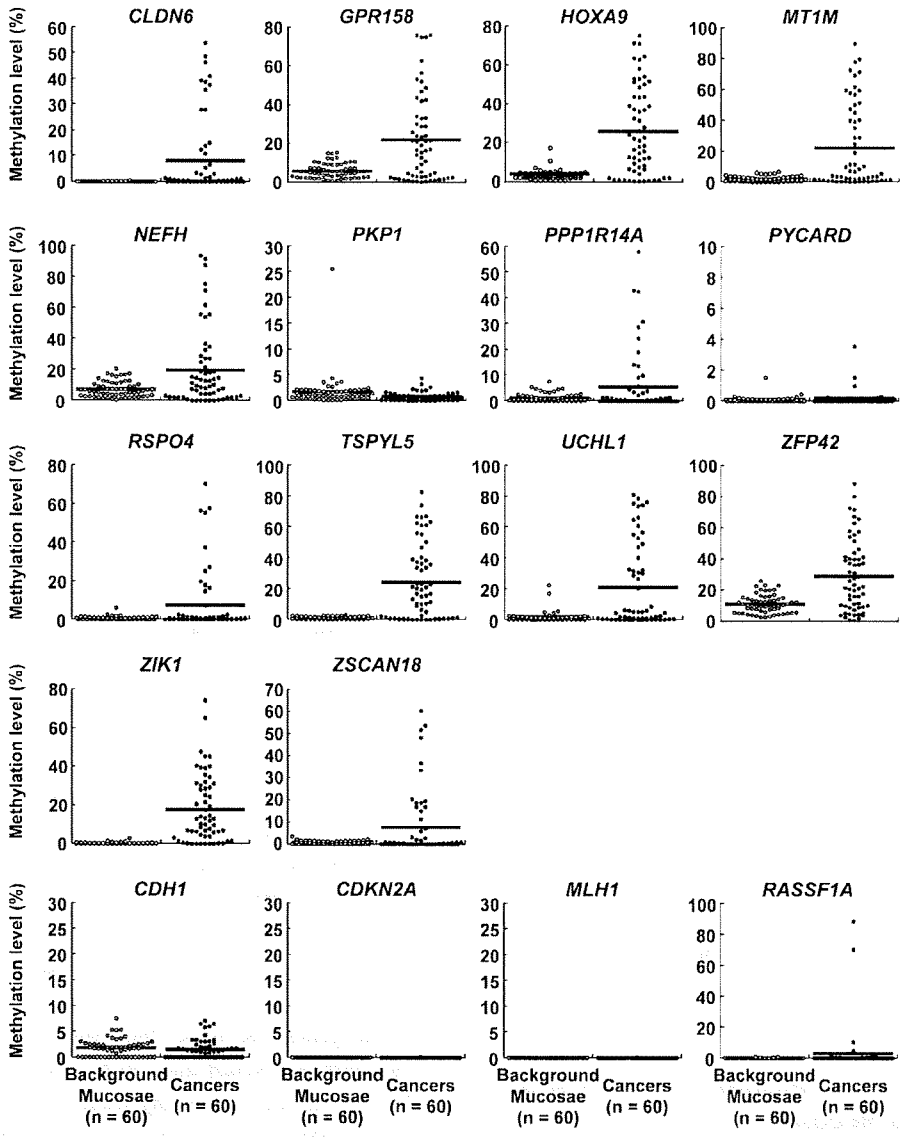

FIGURE 1. Methylation levels of the 14 selected genes and 4 tumor suppressor genes in esophageal squamous cell carcinomas (ESCCs) and their background mucosae are shown. Distinct distributions of methylation levels were observed in the background mucosae and in ESCCs. Mean methylation levels of individual groups are indicated by horizontal lines. *CLDN6* indicates claudin 6; *GPR158*, G protein-coupled receptor 158; *HOXA9*, homeobox A9; *MT1M*, metallothionein 1M; *NEFH*, neurofilament, heavy polypeptide 200 kDa; *PKP1*, plakophilin 1; *PPP1R14A*, protein phosphatase 1, regulatory (inhibitor) subunit 14A; *PYCARD*, pyrin domain (PYD) and caspase recruitment domain (CARD) containing; *RSPO4*, R-spondin family, member 4; *TSPYL5*, testis-specific protein, Y-encoded-like 5; *UCHL1*, ubiquitin carboxyl-terminal esterase L1; *ZFP42*, zinc-finger protein 42 homolog; *ZIK1*, zinc-finger protein interacting with K protein 1 homolog; *ZSCAN18*, zinc-finger and SCAN domain containing 18; *CDH1*, cadherin 1; *CDKN2A*, cyclin-dependent kinase inhibitor 2A; *MLH1*, MutL homolog 1; *RASSF1A*, Ras association domain family member 1, isoform A.

not have methylation in the background mucosae, was excluded. With age, a significant correlation was observed only for *TSPYL5* ($\rho = .273$; $P = .035$). It is interesting to note that, with smoking duration (n = 57), significant correlations were observed for 5 genes; *HOXA9* ($\rho = .268$; $P = .044$), *MT1M* ($\rho = .405$; $P = .002$), *NEFH* ($\rho = .285$; $P = .032$), *RSPO4* ($\rho = .300$; $P = .024$), and *UCHL1* ($\rho = .437$; $P = .001$). Scatter diagrams are shown in Figure 2, and the mean methylation levels in 3 tertile groups are shown in Figure 3.

Table 2. Correlations (Shown as *P* Values) Between Tobacco Use or Alcohol Use and Methylation Levels in Background Mucosae of the 13 Selected Genes

Variable	*GPR158*	*HOXA9*	*MT1M*	*NEFH*	*PKP1*	*PPP1R-14A*	*PYCARD*	*RSPO4*	*TSPYL5*	*UCHL1*	*ZFP42*	*ZIK1*	*ZSCAN18*
Age													
ρ	.177	.218	.035	.225	.077	.246	.216	.100	.273	−.079	.200	.061	.049
P	.176	.094	.789	.084	.560	.058	.097	.449	.035*	.551	.126	.642	.707
Smoking duration													
ρ	.155	.268	.405	.285	.158	.130	−.118	.300	.215	.437	.181	.027	.108
P	.248	.044*	.002*	.032*	.242	.335	.383	.024*	.109	.001*	.177	.841	.424
Alcohol consumption													
ρ	−.154	−.078	.071	−.139	−.072	−.168	−.249	−.089	−.214	.224	−.249	−.024	−.251
P	.263	.571	.605	.312	.600	.219	.067	.517	.116	.100	.067	.861	.064
Inactive ALDH2†													
ρ	.021	−.036	.228	−.039	−.034	.030	−.334	-.058	−.111	.259	−.142	.009	−.202
P	.890	.812	.132	.798	.822	.843	.025*	.705	.469	.085	.353	.952	.184

GPR158 indicates; G protein-coupled receptor 158; *HOXA9*, homeobox A9; *MT1M*, metallothionein 1M; *NEFH*, neurofilament, heavy polypeptide 200 kDa; *PKP1*, plakophilin 1; *PPP1R14A*, protein phosphatase 1, regulatory (inhibitor) subunit 14A; *PYCARD*, pyrin domain and caspase recruitment domain containing; *RSPO4*, R-spondin family, member 4; *TSPYL5*, testis-specific protein, Y-encoded–like 5; *UCHL1*, ubiquitin carboxyl-terminal esterase L1; *ZFP42*, zinc-finger protein 42 homolog; *ZIK1*, zinc-finger protein interacting with K protein 1 homolog; *ZSCAN18*, zinc-finger and SCAN domain containing 18; ALDH2, aldehyde dehydrogenase 2.

*$P < .05$.

† The $ALDH2^1/ALDH2^2$ heterozygote (n=45) and the $ALDH2^2/ALDH2^2$ homozygote (n=0).

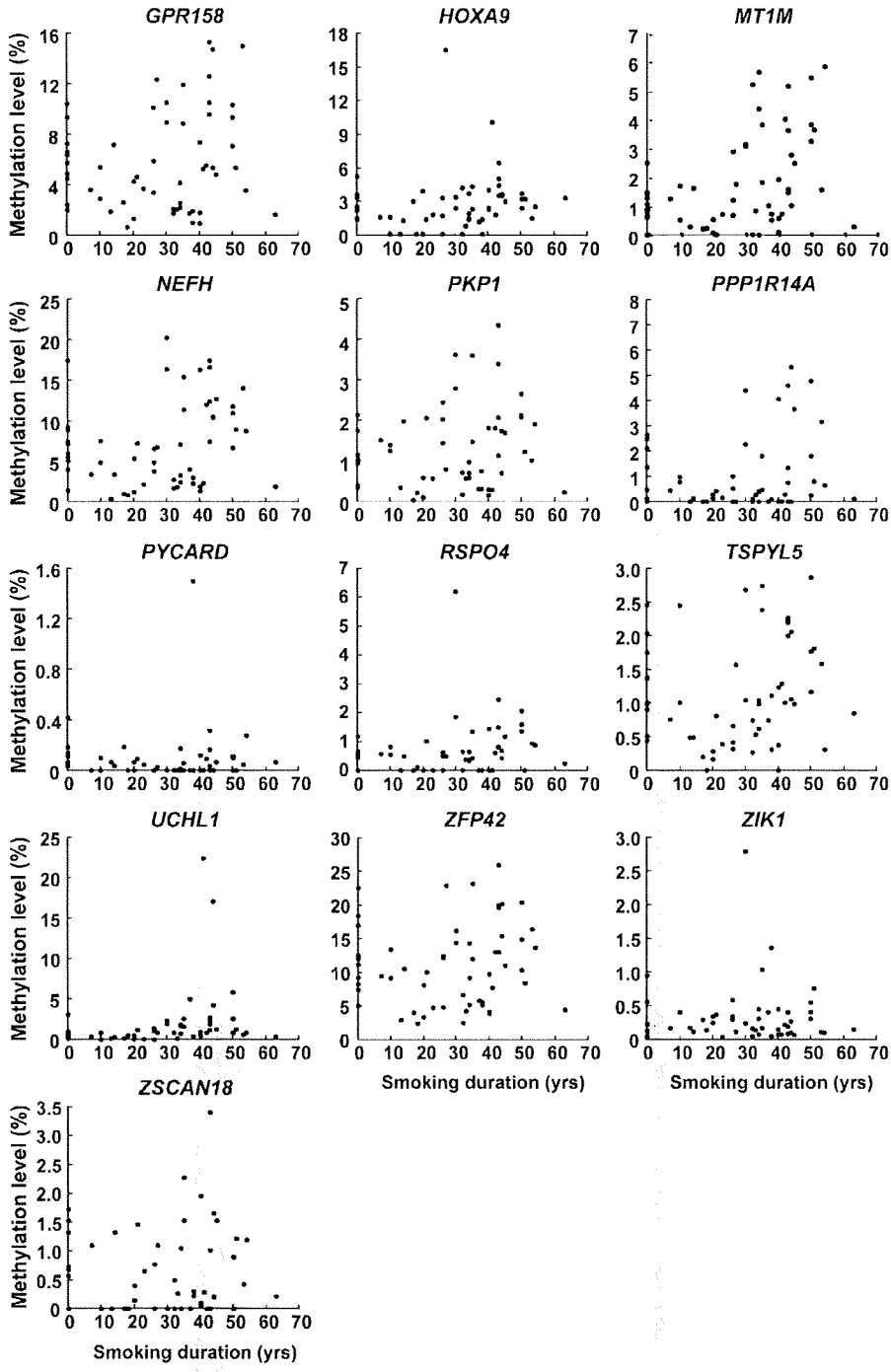

FIGURE 2. Correlations between smoking duration and methylation levels of the 13 selected genes in the background mucosae are illustrated. A positive correlation was observed for homeobox A9 (*HOXA9*) ($P = .044$), metallothionein 1M (*MT1M*) ($P = .002$), neurofilament, heavy polypeptide 200 kDa (*NEFH*) ($P = .032$), R-spondin family, member 4 (*RSPO4*) ($P = .024$), and ubiquitin carboxyl-terminal esterase L1 (*UCHL1*) ($P = .001$) methylation levels. *GPR158* indicates G protein-coupled receptor 158; *PKP1*, plakophilin 1; *PPP1R14A*, protein phosphatase 1, regulatory (inhibitor) subunit 14A; *PYCARD*, pyrin domain (PYD) and caspase recruitment domain (CARD) containing; *TSPYL5*, testis-specific protein, Y-encoded-like 5; *ZFP42*, zinc-finger protein 42 homolog; *ZIK1*, zinc-finger protein interacting with K protein 1 homolog; *ZSCAN18*, zinc-finger and SCAN domain containing 18.

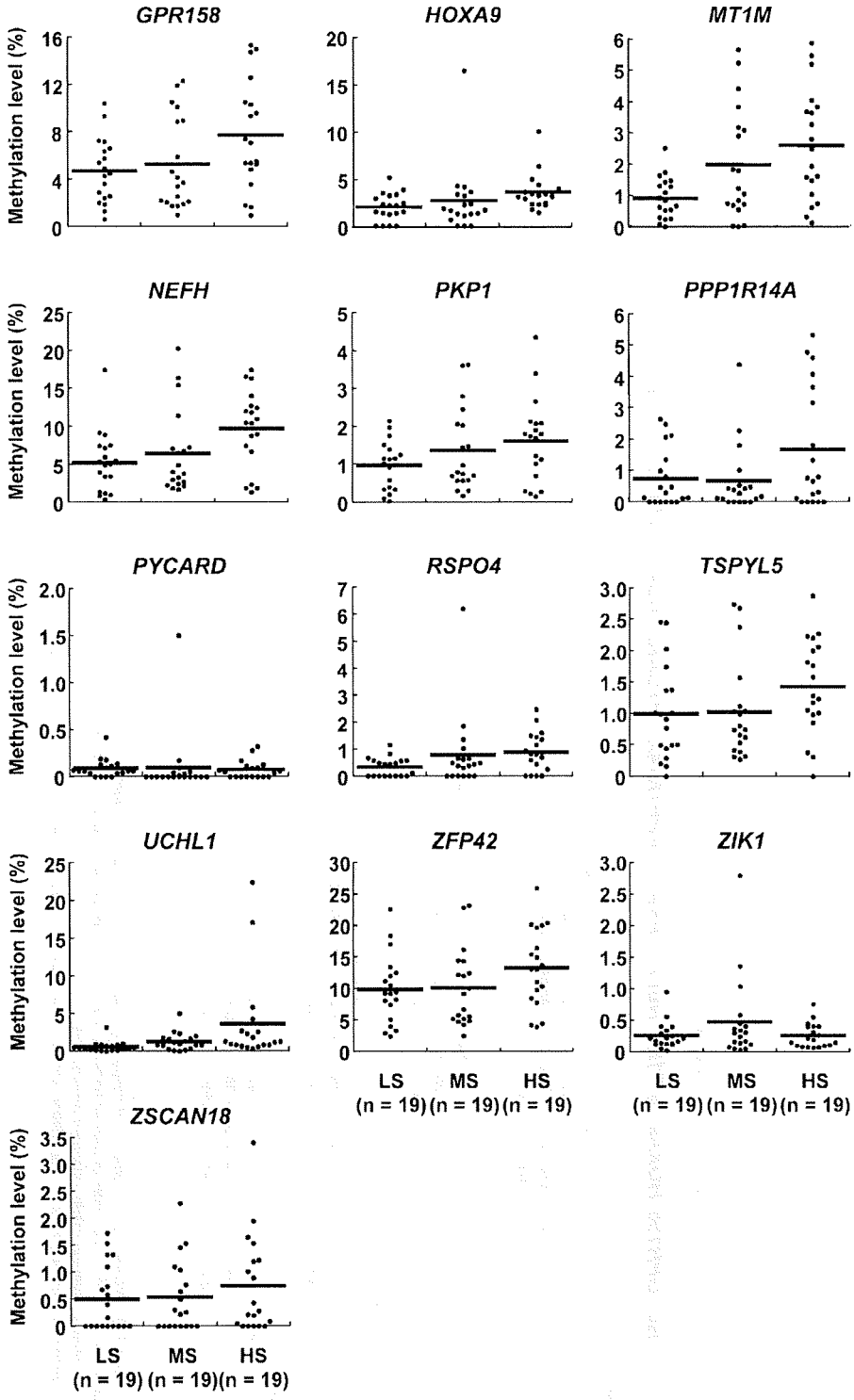

FIGURE 3. Methylation levels of the 13 selected genes in tertiles of smoking duration are shown light smokers (LS) (no or short smoking duration; <21 years); middle smokers (MS) (smoking duration <40 years but ≥21 years); and heavy smokers (HS) (smoking duration ≥40 years). Significantly increasing trends of methylation levels were observed for homeobox A9 (*HOXA9*), metallothionein 1M (*MT1M*), neurofilament heavy polypeptide 200 kDa (*NEFH*), R-spondin family member 4 (*RSPO4*), and ubiquitin carboxyl-terminal esterase L1 (*UCHL1*). *GPR158* indicates G protein-coupled receptor 158; *PKP1*, plakophilin 1; *PPP1R14A*, protein phosphatase 1, regulatory (inhibitor) subunit 14A; *PYCARD*, pyrin domain (PYD) and caspase recruitment domain (CARD) containing; *TSPYL5*, testis-specific protein, Y-encoded-like 5; *ZFP42*, zinc-finger protein 42 homolog; *ZIK1*, zinc-finger protein interacting with K protein 1 homolog; *ZSCAN18*, zinc-finger and SCAN domain containing 18.

Table 3. Associations Between Clinicopathologic Characters and Methylation Levels in the Background Mucosae of the 13 Selected Genes*

Variable	No.	*GPR158*	*HOXA9*	*MT1M*	*NEFH*	*PKP1*	*PPP1R14A-*	*PYCARD*	*RSPO4*	*TSPYL5*	*UCHL1*	*ZFP42*	*ZIK1*	*ZSCAN18*
Tumor differentiation														
Poor	25	4.79±3.42	2.45±1.92	1.80±1.99	6.21±5.16	1.06±0.89	1.10±1.87	0.12±0.30	0.62±1.24	0.99±0.77	2.22±5.36	9.8±5.5	0.28±0.54	0.53±0.58
Moderate	20	6.44±4.40	2.49±1.66	1.65±1.47	7.99±5.61	1.55±1.11	1.10±1.44	0.08±0.12	0.68±0.70	1.18±0.76	1.49±1.38	10.7±6.0	0.36±0.32	0.63±0.93
Well	13	6.72±3.90	3.63±4.24	2.34±1.72	7.46±4.56	3.32±6.72	1.07±1.42	0.03±0.03	0.62±0.58	1.39±0.82	1.45±1.51	13.2±6.4	0.40±0.25	0.65±0.75
P		.224	.328	.340	.410	.160	.441	.605	.568	.339	.262	.362	.006†	.948
Depth of tumor														
T1/T2	12	4.40±2.82	2.00±2.81	1.56±1.90	4.81±4.05	1.21±0.87	1.13±2.26	0.02±0.03	0.34±0.38	0.96±0.75	2.44±6.30	9.5±6.0	0.27±0.18	0.32±0.56
T3/T4	46	6.16±4.11	2.97±2.43	1.95±1.73	7.70±5.30	1.88±3.70	1.08±1.42	0.10±0.23	0.72±1.02	1.19±0.79	1.63±2.66	11.2±5.9	0.35±0.46	0.66±0.77
P		.216	.024†	.372	.076	.715	.277	.046†	.201	.296	.246	.328	.977	.081
Lymph node metastasis														
Negative	8	6.14±3.04	2.23±1.02	2.40±1.42	8.00±5.67	1.72±1.08	1.39±1.90	0.05±0.05	0.73±0.64	1.06±0.77	1.24±1.00	11.9±5.9	0.51±0.46	0.80±0.93
Positive	52	5.79±4.00	2.77±2.67	1.79±1.78	7.02±5.06	1.73±3.49	1.14±1.64	0.09±0.22	0.64±0.97	1.16±0.77	1.88±3.84	10.9±6.0	0.30±0.40	0.55±0.70
P		.500	.446	.158	.704	.254	.429	.946	.480	.863	.761	.557	.131	.522
Multiplicity of tumor														
Solitary	48	5.99±3.84	2.59±1.79	1.97±1.82	7.38±5.04	1.84±3.60	1.26±1.71	0.06±0.09	0.68±0.96	1.18±0.75	1.89±3.95	11.3±5.7	0.34±0.43	0.52±0.62
Multiple	10	4.86±4.37	2.74±2.65	1.38±1.35	5.77±5.85	1.22±1.33	0.31±0.44	0.21±0.46	0.45±0.82	0.98±0.94	1.33±1.61	9.0±6.9	0.31±0.33	0.95±1.13
P		.237	.845	.446	.206	.249	.126	.345	.111	.229	.571	.138	.571	.202

GPR158 indicates; G protein-coupled receptor 158; *HOXA9*, homeobox A9; *MT1M*, metallothionein 1M; *NEFH*, neurofilament, heavy polypeptide 200 kDa; *PKP1*, plakophilin 1; *PPP1R14A*, protein phosphatase 1, regulatory (inhibitor) subunit 14A; *PYCARD*, pyrin domain and caspase recruitment domain containing; *RSPO4*, R-spondin family, member 4; *TSPYL5*, testis-specific protein, Y-encoded–like 5; *UCHL1*, ubiquitin carboxyl-terminal esterase L1; *ZFP42*, zinc-finger protein 42 homolog; *ZIK1*, zinc-finger protein interacting with K protein 1 homolog; *ZSCAN18*, zinc-finger and SCAN domain containing 18.

* Methylation levels are described as average values±standard deviation (%).

† Significant associations were observed only for *ZIK1* methylation and tumor differentiation, *HOXA9* methylation and depth of tumor, and *PYCARD* methylation and depth of tumor.

With alcohol intake (n = 55), in contrast, no correlation was observed in the patients overall. When 55 patients were classified into those who had the active ALDH2 type (*ALDH2*^{1}/*ALDH2*^{1} homozygote; 10 patients) and those who had the inactive ALDH2 type (*ALDH2*^{1}/*ALDH2*^{2} heterozygote; 45 patients), a significant inverse correlation was present for *PYCARD* ($\rho = -0.334$; $P = .025$) among patients who had the inactive ALDH2 type. No positive correlation was observed with any genes in either group.

We also examined associations between methylation levels in the background mucosae and clinicopathologic characters, including tumor differentiation, depth of tumor, positive lymph node metastasis, and multiplicity of tumors (Table 3). Associations were observed only in 3 analyses, *ZIK1* methylation and tumor differentiation, *HOXA9* methylation and depth of tumor, and *PYCARD* methylation and depth of tumor, and the other 49 analyses were negative.

Relation Between Methylation in the Background Mucosae and in ESCCs

Methylation in the background mucosae reflects methylation events in numerous stem/progenitor cells, and its degree can be assessed by methylation levels. To incorporate methylation levels of multiple genes analyzed for a sample into 1 value, we calculated deviation values for the genes, and their average was used. In contrast, methylation in cancer tissue, if it occurred in cancer precursor cells, theoretically is present in all cancer cells in a sample. To assess the degree of methylation in such cancer precursor cells, we obtained the number of methylated genes in an ESCC. Eleven genes were used for the analysis, because *CLDN6* was not methylated at all in the background mucosae, and *PKP1* and *PYCARD* were not methylated in any of the ESCCs (methylation levels, $\leq$6%). No significant correlation between the methylation levels in the background mucosae and the methylation frequencies in the ESCCs (n = 60) was observed ($r = .212$; $P = .104$) (Fig. 4).

DISCUSSION

In this study, we demonstrated that duration of tobacco smoking is correlated significantly with DNA methylation levels of promoter CGIs of *HOXA9*, *MT1M*, *NEFH*, *RSPO4*, and *UCHL1* in esophageal mucosae. This strongly indicates that chronic tobacco smoking induces aberrant DNA methylation of multiple genes in esophageal mucosae and that a predisposed field for ESCCs is formed (epigenetic field defect or epigenetic field for cancerization). Ishii et al demonstrated in a qualitative analysis of 14 genes that methylation was more frequent in the background mucosae from patients with ESCC than in mucosae from healthy volunteers,[32] indicating that the presence of aberrant methylation in esophageal mucosae is associated with ESCC development. The degree of aberrant methylation in gastric mucosae is correlated with gastric cancer risk,[21,33] and the presence of aberrant methylation in noncancerous tissues also is associated with the risk of liver cancer,[34] colon cancer,[35] breast cancer,[36] and renal cancer.[37] Therefore, it is highly possible that the degree of aberrant methylation is correlated with the risk of ESCC.

The current study clearly indicated a correlation between the quantity of aberrant DNA methylation and smoking duration, although an association between methylation incidence and smoking (or alcohol intake) was not observed in a previous study.[32] This "discrepancy" most likely occurred because our quantitative analysis was able to detect differences even among methylation-positive individuals and also because we screened and selected genes with methylation levels that were correlated with tobacco smoking. Different genes have different susceptibility to methylation induction by specific methylation-inducing agents,[21,38] partly because genes with low transcription have high susceptibility to methylation induction.[39] In bronchial epithelia, an association between tobacco smoking and the methylation of some genes has been reported by qualitative studies.[17–19] A more extensive search for genes that are methylated in association with smoking duration may lead to the isolation of more marker genes.

The mechanism(s) with which tobacco smoking induces aberrant DNA methylation is important. Generally, as an inducer of aberrant DNA methylation, chronic inflammation is considered important.[39] In gastric mucosae, *Helicobacter pylori* infection induces aberrant methylation, possibly through the induction of chronic inflammation.[21] In colonic mucosae from patients with ulcerative colitis, it is known that aberrant methylation is

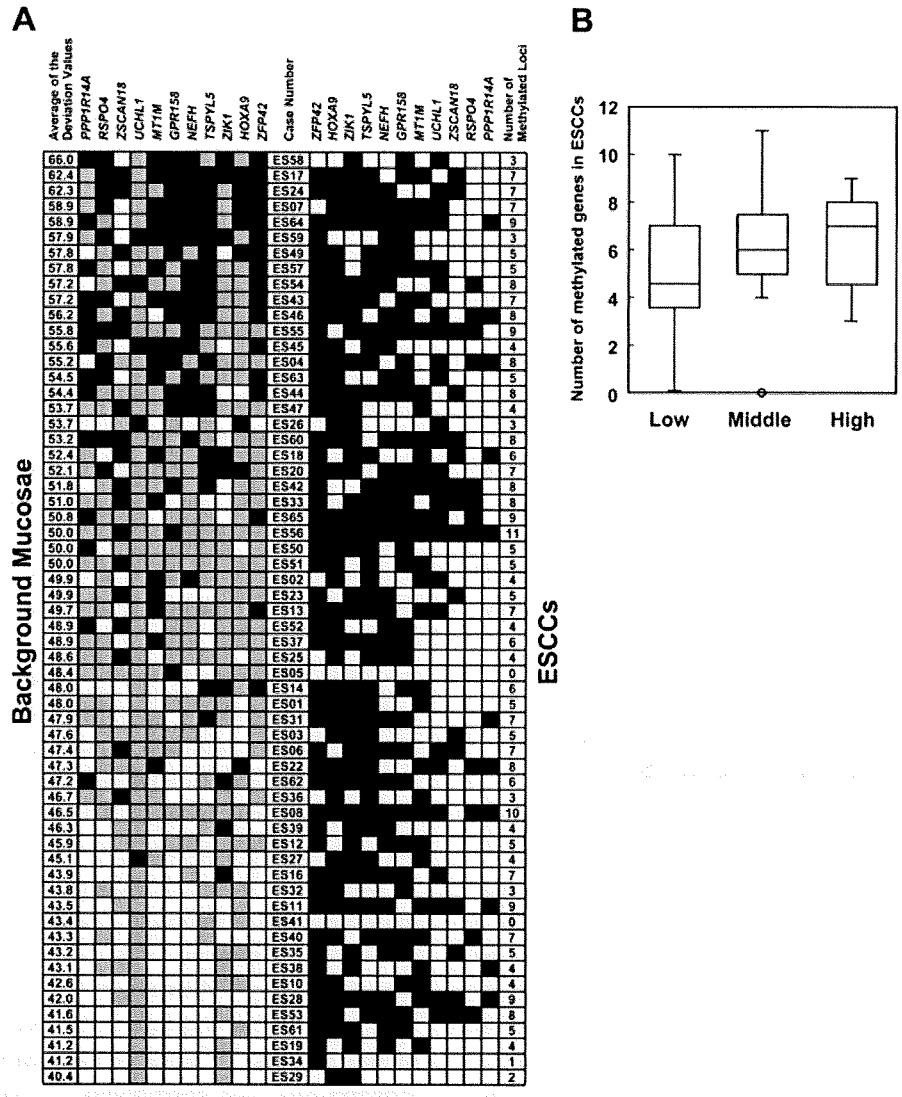

FIGURE 4. Correlations between methylation levels in the background mucosae and frequent methylation in esophageal squamous cell carcinomas (ESCCs) are illustrated. (A) Color coded representation of deviation values in the background mucosae and presence of methylation in ESCCs. The analysis was performed for 11 marker genes with methylation in the background mucosae and methylation-positive ESCCs, and the 11 genes were aligned in the order of frequency of methylation in ESCCs. The 60 samples were aligned in the order of high average of the deviation values. No trend toward a decreasing frequency of methylation in ESCCs was observed. *PPP1R14A* indicates protein phosphatase 1, regulatory (inhibitor) subunit 14A; *RSPO4*, R-spondin family, member 4; *ZSCAN18*, zinc-finger and SCAN domain containing 18; *UCHL1*, ubiquitin carboxyl-terminal esterase L1; *MT1M*, metallothionein 1M; *GPR158*, G protein-coupled receptor 158; *NEFH*, neurofilament, heavy polypeptide 200 kDa; *TSPYL5*, testis-specific protein, Y-encoded-like 5; *ZIK1*, zinc-finger protein interacting with K protein 1 homolog; *ZFP42*, zinc-finger protein 42 homolog; *HOXA9*, homeobox A9. (B) Box graph representation of the number of methylated genes in tertiles of the methylation levels in the background mucosae. The tertiles were obtained by the average of the deviation values in A. No difference in the number of methylated genes was observed.

present.[40] There is little information whether or not tobacco smoking induces chronic inflammation in esophageal mucosae. In bronchial epithelia, it is known that tobacco smoking induces chronic inflammation,[41] and tobacco ingredients that were swallowed induced similar inflammation in esophageal mucosae. It was reported recently that bladder cancers in smokers tended to have methylation of the runt-related transcription factor 3 gene *RUNX3*,[42] and systemic effects of tobacco ingredients also may be possible.

Alcohol drinking, in contrast with smoking, did not induce methylation. Rather, *PYCARD* methylation levels in the background esophageal mucosae decreased significantly with alcohol consumption in patients with who had the inactive ALDH2 genotype. It has been reported that chronic alcohol consumption induces global hypomethylation in rats,[43,44] and it has been suggested that alcohol has an epigenetic action different from tobacco smoking, although both are risk factors for ESCC. In addition, it is suggested that the methylation profile associated with tobacco smoking is different from that associated with alcohol drinking. Once a methylation profile specific to tobacco smoking or alcohol drinking is established, it can be used as a methylation fingerprint to assess past exposure to these factors for clinicopathologic analysis and epidemiology. The use of DNA methylation profiles as fingerprints of exposure to carcinogenic factors is expected to become an important field.[45,46]

To assess the extent of tobacco smoking and alcohol drinking, we used duration (not the amount of intake) and mean intake (not duration), respectively. This was because previous reports suggested that ESCC risk depends mainly on smoking duration rather than mean tobacco intake and on mean alcohol intake rather than drinking duration.[3,5]

Methylation levels in cancers were useful for estimating the roles of genes in ESCC development. *PKP1* and *PYCARD* methylation levels in ESCCs were <6%, and it was unlikely that their methylation took place before monoclonal growth of cancer cells. Similarly, *CDH1* methylation levels in ESCC were <7.1% in all ESCCs, suggesting that *CDH1* methylation was unlikely to be involved in early stages of ESCC development. This finding was in accordance with a previous report that *CDH1* methylation was involved in metastatic progression.[47] The *CDKN2A* and *MLH1* tumor-suppressor genes were not methylated in ESCCs or in their background mucosae. In contrast, the *RASSF1A* tumor-suppressor gene had methylation levels >30% in 2 ESCCs and zero in most ESCCs. This suggested that *RASSF1A* silencing may be involved in the early stages of ESCC development, but the incidence was low. Aberrant methylation of *CDH1*, *CDKN2A*, and *RASSF1A* in ESCCs was reported in 14 of 20 ESCCs, 17 of 34 ESCCs, and 25 of 48 ESCCs, respectively, by qualitative MSP.[27–29] It is known that qualitative MSP tends to overestimate methylation frequencies,[48] and the incidences reported here were considered reasonable. Some genes, such as *NEFH* and *ZFP42* had methylation levels of almost 100% in some ESCCs. Because cancer tissues contained not only cancer cells but also stromal cells, these genes should have been methylated in both cancer cells and stromal cells. Recent studies demonstrated that cancer stromal cells have distinct epigenetic changes,[49] and some of these genes may be involved in such changes.

There was no correlation between methylation levels in the background mucosae and methylation frequencies in ESCCs, as in our previous study on gastric cancers and their background mucosae.[30] This suggests that methylation levels in the background mucosae do not necessarily reflect methylation levels in cancer precursor cells. In conclusion, we have demonstrated that chronic tobacco smoking is associated with the accumulation of aberrant methylation of multiple genes in esophageal mucosae.

Conflict of Interest Disclosures

Supported by Grants-in-aid for the Third-term Comprehensive Cancer Control Strategy and for Pioneering Basic Research from the Ministry of Health, Labor and Welfare, Japan.

Dr. Oka is the recipient of a Research Resident Fellowships from the Foundation for Promotion of Cancer Research.

References

1. Enzinger PC, Mayer RJ. Esophageal cancer. *N Engl J Med.* 2003;349:2241-2252.
2. Vizcaino AP, Moreno V, Lambert R, Parkin DM. Time trends incidence of both major histologic types of esophageal carcinomas in selected countries, 1973-1995. *Int J Cancer.* 2002;99:860-868.
3. Castellsague X, Munoz N, De Stefani E, et al. Independent and joint effects of tobacco smoking and alcohol drinking on the risk of esophageal cancer in men and women. *Int J Cancer.* 1999;82:657-664.
4. Lee CH, Wu DC, Lee JM, et al. Carcinogenetic impact of alcohol intake on squamous cell carcinoma risk of the oesophagus in relation to tobacco smoking. *Eur J Cancer.* 2007; 43:1188-1199.
5. Launoy G, Milan CH, Faivre J, Pienkowski P, Milan CI, Gignoux M. Alcohol, tobacco and oesophageal cancer: effects of the duration of consumption, mean intake and current and former consumption. *Br J Cancer.* 1997;75: 1389-1396.

6. Kuwano H, Ohno S, Matsuda H, Mori M, Sugimachi K. Serial histologic evaluation of multiple primary squamous cell carcinomas of the esophagus. *Cancer.* 1988;61:1635-1638.
7. Kanamoto A, Yamaguchi H, Nakanishi Y, Tachimori Y, Kato H, Watanabe H. Clinicopathological study of multiple superficial oesophageal carcinoma. *Br J Surg.* 2000;87:1712-1715.
8. Pesko P, Rakic S, Milicevic M, Bulajic P, Gerzic Z. Prevalence and clinicopathologic features of multiple squamous cell carcinoma of the esophagus. *Cancer.* 1994;73:2687-2690.
9. Saeki H, Kimura Y, Ito S, Miyazaki M, Ohga T. Biologic and clinical significance of squamous epithelial dysplasia of the esophagus. *Surgery.* 2002;131(1 suppl):S22-S27.
10. Nakanishi Y, Ochiai A, Yoshimura K, et al. The clinicopathologic significance of small areas unstained by Lugol's iodine in the mucosa surrounding resected esophageal carcinoma: an analysis of 147 cases. *Cancer.* 1998;82:1454-1459.
11. Miyazaki M, Ohno S, Futatsugi M, Saeki H, Ohga T, Watanabe M. The relation of alcohol consumption and cigarette smoking to the multiple occurrence of esophageal dysplasia and squamous cell carcinoma. *Surgery.* 2002; 131(1 suppl):S7-S13.
12. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer.* 1953;6:963-968.
13. Jones PA, Baylin SB. The epigenomics of cancer. *Cell.* 2007;128:683-692.
14. Sato F, Meltzer SJ. CpG island hypermethylation in progression of esophageal and gastric cancer. *Cancer.* 2006; 106:483-493.
15. Mandelker DL, Yamashita K, Tokumaru Y, et al. PGP9.5 promoter methylation is an independent prognostic factor for esophageal squamous cell carcinoma. *Cancer Res.* 2005;65:4963-4968.
16. Soma T, Kaganoi J, Kawabe A, Kondo K, Imamura M, Shimada Y. Nicotine induces the fragile histidine triad methylation in human esophageal squamous epithelial cells. *Int J Cancer.* 2006;119:1023-1027.
17. Belinsky SA, Palmisano WA, Gilliland FD, et al. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. *Cancer Res.* 2002; 62:2370-2377.
18. Soria JC, Rodriguez M, Liu DD, Lee JJ, Hong WK, Mao L. Aberrant promoter methylation of multiple genes in bronchial brush samples from former cigarette smokers. *Cancer Res.* 2002;62:351-355.
19. Zochbauer-Muller S, Lam S, Toyooka S, et al. Aberrant methylation of multiple genes in the upper aerodigestive tract epithelium of heavy smokers. *Int J Cancer.* 2003;107:612-616.
20. Chang HW, Ling GS, Wei WI, Yuen AP. Smoking and drinking can induce p15 methylation in the upper aerodigestive tract of healthy individuals and patients with head and neck squamous cell carcinoma. *Cancer.* 2004;101:125-132.
21. Maekita T, Nakazawa K, Mihara 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;12:989-995.
22. Shimada Y, Imamura M, Wagata T, Yamaguchi N, Tobe T. Characterization of 21 newly established esophageal cancer cell lines. *Cancer.* 1992;69:277-284.
23. Moriguchi K, Yamashita S, Tsujino Y, Tatematsu M, Ushijima T. Larger numbers of silenced genes in cancer cell lines with increased de novo methylation of scattered CpG sites. *Cancer Lett.* 2007;249:178-187.
24. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A.* 2002;99:3740-3745.
25. Sugimachi K, Ohno S, Matsuda H, Mori M, Kuwano H. Lugol-combined endoscopic detection of minute malignant lesions of the thoracic esophagus. *Ann Surg.* 1988;208:179-183.
26. Takeshita T, Morimoto K, Mao X, Hashimoto T, Furuyama J. Characterization of the 3 genotypes of low Km aldehyde dehydrogenase in a Japanese population. *Hum Genet.* 1994;94:217-223.
27. Xing EP, Nie Y, Wang LD, Yang GY, Yang CS. Aberrant methylation of p16INK4a and deletion of p15INK4b are frequent events in human esophageal cancer in Linxian, China. *Carcinogenesis.* 1999;20:77-84.
28. Si HX, Tsao SW, Lam KY, Srivastava G, Liu Y, Wong YC, et al. E-cadherin expression is commonly downregulated by CpG island hypermethylation in esophageal carcinoma cells. *Cancer Lett.* 2001;173:71-78.
29. Kuroki T, Trapasso F, Yendamuri S, et al. Promoter hypermethylation of RASSF1A in esophageal squamous cell carcinoma. *Clin Cancer Res.* 2003;9:1441-1445.
30. Enomoto S, Maekita T, Tsukamoto T, et al. Lack of association between CpG island methylator phenotype in human gastric cancers and methylation in their background non-cancerous gastric mucosae. *Cancer Sci.* 2007;98:1853-1861.
31. Ogino S, Cantor M, Kawasaki T, et al. CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. *Gut.* 2006;55:1000-1006.
32. Ishii T, Murakami J, Notohara K, et al. Oesophageal squamous cell carcinoma may develop within a background of accumulating DNA methylation in normal and dysplastic mucosa. *Gut.* 2007;56:13-19.
33. Nakajima T, Maekita T, Oda I, et al. Higher methylation levels in gastric mucosae significantly correlate with higher risk of gastric cancers. *Cancer Epidemiol Biomarkers Prev.* 2006;15:2317-2321.
34. Kondo Y, Kanai Y, Sakamoto M, Mizokami M, Ueda R, Hirohashi S. Genetic instability and aberrant DNA

methylation in chronic hepatitis and cirrhosis—A comprehensive study of loss of heterozygosity and microsatellite instability at 39 loci and DNA hypermethylation on 8 CpG islands in microdissected specimens from patients with hepatocellular carcinoma. *Hepatology.* 2000;32:970-979.

35. Shen L, Kondo Y, Rosner GL, et al. MGMT promoter methylation and field defect in sporadic colorectal cancer. *J Natl Cancer Inst.* 2005;97:1330-1338.
36. Yan PS, Venkataramu C, Ibrahim A, et al. Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. *Clin Cancer Res.* 2006;12:6626-6636.
37. Arai E, Kanai Y, Ushijima S, Fujimoto H, Mukai K, Hirohashi S. Regional DNA hypermethylation and DNA methyltransferase (DNMT) 1 protein overexpression in both renal tumors and corresponding nontumorous renal tissues. *Int J Cancer.* 2006;119:288-296.
38. Nakajima T, Yamashita S, Maekita T, Niwa T, Nakazawa K, Ushijima T. The presence of a methylation fingerprint of *Helicobacter pylori* infection in human gastric mucosae. *Int J Cancer.* 2009;124:905-910.
39. Ushijima T, Okochi-Takada E. Aberrant methylations in cancer cells: where do they come from? *Cancer Sci.* 2005;96:206-211.
40. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in ulcerative colitis. *Cancer Res.* 2001;61:3573-3577.
41. van der Vaart H, Postma DS, Timens W, ten Hacken NH. Acute effects of cigarette smoke on inflammation and oxidative stress: a review. *Thorax.* 2004;59:713-721.
42. Wolff EM, Liang G, Cortez CC, et al. RUNX3 methylation reveals that bladder tumors are older in patients with a history of smoking. *Cancer Res.* 2008;68:6208-6214.
43. Garro AJ, McBeth DL, Lima V, Lieber CS. Ethanol consumption inhibits fetal DNA methylation in mice: implications for the fetal alcohol syndrome. *Alcohol Clin Exp Res.* 1991;15:395-398.
44. Choi SW, Stickel F, Baik HW, Kim YI, Seitz HK, Mason JB. Chronic alcohol consumption induces genomic but not p53-specific DNA hypomethylation in rat colon. *J Nutr.* 1999;129:1945-1950.
45. Herceg Z. Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. *Mutagenesis.* 2007;22:91-103.
46. Ushijima T, Nakajima T, Maekita T. DNA methylation as a marker for the past and future. *J Gastroenterol.* 2006; 41:401-407.
47. Graff JR, Gabrielson E, Fujii H, Baylin SB, Herman JG. Methylation patterns of the E-cadherin 5′ CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. *J Biol Chem.* 2000;275:2727-2732.
48. Shames DS, Minna JD, Gazdar AF. Methods for detecting DNA methylation in tumors: from bench to bedside. *Cancer Lett.* 2007;251:187-198.
49. Hu M, Yao J, Cai L, et al. Distinct epigenetic changes in the stromal cells of breast cancers. *Nat Genet.* 2005;37: 899-905.

Cancer Science
The official journal of the Japanese Cancer Association JCA

Review Article

Genome-wide DNA methylation profiles in precancerous conditions and cancers

Yae Kanai[1]

Pathology Division, National Cancer Center Research Institute, Tokyo, Japan

(Received August 15, 2009/Revised September 25, 2009/Accepted September 27, 2009/Online publication November 4, 2009)

Alterations of DNA methylation, which result in chromosomal instability and silencing of tumor-related genes, are among the most consistent epigenetic changes observed in human cancers. Analysis of tissue specimens has revealed that DNA methylation alterations participate in multistage carcinogenesis, even from the early and precancerous stages, especially in association with chronic inflammation and/or persistent viral infection, such as chronic hepatitis or liver cirrhosis resulting from infection with hepatitis B or C virus. DNA methylation alterations can account for the histological heterogeneity and clinicopathological diversity of human cancers. Overexpression of DNA methyltransferase 1 is not a secondary result of increased cell proliferative activity, but is significantly correlated with accumulation of DNA hypermethylation in CpG islands of tumor-related genes. Alteration of DNA methyltransferase 3b splicing may result in chromosomal instability through DNA hypomethylation in pericentromeric satellite regions. Genome-wide analysis of DNA methylation status has revealed that the DNA methylation profile at the precancerous stage is basically inherited by the corresponding cancers developing in individual patients. DNA methylation status is not simply altered at the precancerous stage; rather, DNA methylation alterations at the precancerous stage may confer vulnerability to further genetic and epigenetic alterations, generate more malignant cancers, and thus determine patient outcome. Therefore, genome-wide DNA methylation profiling may provide optimal indicators for carcinogenetic risk estimation and prognostication, and thus provide an avenue for cancer prevention and therapy on an individual basis. (*Cancer Sci* 2010; 101: 36–45)

DNA methylation, a covalent chemical modification resulting in addition of a methyl group at the carbon five position of the cytosine ring in CpG dinucleotides, is one of the most consistent epigenetic changes observed in human cancers.[1] DNMTs transfer methyl groups from *S*-adenosylmethionine to cytosines.[2] The preference of DNMT1, a major and well-known DNMT, for hemimethylated over unmethylated substrates *in vitro*,[3] and its targeting of replication foci by binding to PCNA,[4,5] are believed to allow copying of the DNA methylation pattern on the parental strand to the newly synthesized daughter DNA strand. Thus, DNMT1 has been recognized as a ''maintenance'' DNMT,[6] whereas DNMT3a and DNMT3b show *de novo* DNA methylation activity.[7] DNA methylation normally promotes a highly condensed heterochromatin structure associated with deacetylation of histones H3 and H4, loss of histone H3, lysine 4 (H3K4) methylation, and gain of H3K9 and H3K27 methylation.[8] When methyl-CpG-binding proteins, such as MeCP2[9,10] and MBD2,[11] bind to methylated CpG dinucleotide, their transcriptional repression domain recruits a co-repressor complex containing histone deacetylases. However, histone methyltransferases, such as G9A[12] and SUV39H1,[13] are required to recruit DNMTs. DNA methylation is a stable modification inherited throughout consecutive cell divisions, being essential for the normal development and function of adult organs, particularly for X-chromosome inactivation, genome imprinting, silencing of transposons and other parasitic elements, and proper expression of genes.[14]

Reduction of DNMT1 activity in genetically engineered animals alters the number of tumors or the timing of tumor development, suggesting a causal relationship between DNA methylation alterations and tumorigenesis.[15,16] In 1995, when the *RB* and *VHL* genes were the only tumor suppressor genes known to be silenced by DNA methylation, we showed that the E-cadherin tumor suppressor gene is silenced by DNA methylation around the promoter region.[17] The list of tumor-related genes whose expression levels are altered due to DNA hypo- or hypermethylation is increasing.[18–22] Transcriptionally repressive chromatin modifications within the promoters of tumor-related genes silenced by DNA methylation are known to resemble the chromatin modifications of these genes in normal embryonic stem cells, for example, polycomb complex binding and H3K27 methylation.[23] These genes also have an active marker, H3K4 methylation, in normal stem cells, and this bivalent state is converted to a primary active or repressive chromatin conformation after differentiation cues have been received.[23] During carcinogenesis, such modifications may render the genes vulnerable to errors, resulting in aberrant DNA methylation.[24] DNA hypomethylation induces chromosomal instability through decondensation of heterochromatin and enhancement of chromosomal recombination during carcinogenesis.[25] Translational epigenetics have come of age,[26,27] and empirical analysis of DNA methylation status in clinical tissue samples in connection with the clinicopathological diversity of human cancers is assuming increasing importance for the diagnosis, prevention, and therapy of cancers.[28,29]

Alterations of DNA methylation during multistage carcinogenesis

Alterations of DNA methylation at the precancerous stage. DNA methylation alterations play a key role in the early steps of human carcinogenesis. In the 1990s, although LOH on chromosome 16 was frequently detected by classical Southern blotting in HCCs that were poorly differentiated, large in size, and associated with metastasis,[30] only a few of the molecular events occurring in the earlier stage of hepatocarcinogenesis were known. Since DNA methylation alterations may be correlated with chromosomal instability, we examined the DNA methylation status on chromosome 16 using Southern blotting

[1]To whom correspondence should be addressed.
E-mail: ykanai@ncc.go.jp

doi: 10.1111/j.1349-7006.2009.01383.x

with a DNA methylation-sensitive restriction enzyme. DNA methylation alterations at multiple loci on chromosome 16, compared to normal liver tissue samples, were frequently revealed even in samples of non-cancerous liver tissue showing chronic hepatitis or liver cirrhosis,[31,32] which are widely considered to be precancerous conditions,[33] indicating that DNA methylation alterations are a very early event during multistage hepatocarcinogenesis. This was one of the earliest reports of DNA methylation alterations at the precancerous stage.[31]

DNA hypermethylation around the promoter region of the E-cadherin tumor suppressor gene (16q22.1), which encodes a Ca^{2+}-dependent cell-cell adhesion molecule,[34] has been detected even in samples of non-cancerous liver tissue showing chronic hepatitis or cirrhosis.[35] Heterogeneous E-cadherin expression in such non-cancerous liver tissue, which is associated with small focal areas of hepatocytes showing only slight E-cadherin immunoreactivity, might be due, at least partly, to DNA hypermethylation.[35] Reduction of E-cadherin expression due to DNA methylation around the promoter region may participate even in the very early stage of hepatocarcinogenesis through loss of intercellular adhesiveness and destruction of tissue morphology.

Studies of LOH by PCR using microsatellite markers have been reported, using specimens microdissected from precancerous lesions in several organ types. Whether aberrant DNA methylation precedes chromosomal instability during hepatocarcinogenesis was re-examined using microdissected specimens obtained from lobules, pseudo lobules or regenerative nodules in non-cancerous liver tissue from patients with HCCs by bisulfite modification. Although no degree of DNA methylation of any of the examined C-type CpG islands, which are generally methylated in a cancer-specific but not age-dependent manner, was ever detected in normal liver tissue from patients without HCCs, DNA hypermethylation of such islands was frequently found, even in microdissected specimens of non-cancerous liver tissue showing no remarkable histological changes obtained from patients with HCCs in which LOH was never detected.[36] Thus it was confirmed that aberrant DNA methylation is an earlier event preceding chromosomal instability during hepatocarcinogenesis.

As another example of inflammation-associated carcinogenesis, ductal carcinomas of the pancreas frequently develop after chronic damage due to pancreatitis. At least a proportion of peripheral pancreatic ductal epithelia with an inflammatory background may be at the precancerous stage. When the DNA methylation status of the *p14*, *p15*, *p16*, *p73*, *APC*, *hMLH1*, *MGMT*, *BRCA1*, *GSTP1*, *TIMP-3*, *E-cadherin*, and *DAPK-1* genes was examined, the average number of methylated tumor-related genes and the incidence of DNA methylation of at least one gene were increased in peripheral pancreatic ductal epithelia with an inflammatory background and in another precancerous lesion, PanIN, in comparison with normal peripheral pancreatic duct epithelia.[37]

UCs of the urinary bladder, renal pelvis, and ureter are clinically remarkable because of their multicentricity and tendency to recur (Fig. 1a).[38] A possible mechanism for such multiplicity is the "field effect." Even non-cancerous urothelia showing no remarkable histological changes obtained from patients with UCs can be considered precancerous, because they may have been exposed to carcinogens in the urine. When the DNA methylation status of multiple C-type CpG islands was examined, the average number of methylated C-type CpG islands was increased in non-cancerous urothelia showing no remarkable histological changes obtained from patients with UCs, in comparison with normal urothelia obtained from patients without UCs.[39]

Cigarette smoking is another background factor associated with alterations of DNA methylation during multistage carcinogenesis. DNA hypermethylation at the D17S5 locus, where the *HIC (hypermethylated-in-cancer)-1* tumor suppressor gene was identified, is observed even in non-cancerous lung tissue, which may contain progenitor cells for cancers, obtained from patients with non-small-cell lung cancers. The incidence of DNA hypermethylation in non-cancerous lung tissue obtained from patients with non-small-cell lung cancers is significantly correlated with both smoking history and the extent of pulmonary anthracosis, as an index of the cumulative effects of smoking.[40] Thus, DNA methylation alterations are frequently found even at the precancerous stage in various organs, especially in association with chronic inflammation[41,42] and/or persistent infection with viruses[43–45] or other pathogenic microorganisms, and with cigarette smoking.

DNA methyltransferase 1 overexpression and regional DNA hypermethylation. With respect to the molecular backgrounds of DNA methylation alterations,[46] it has been reported that levels of DNMT1 mRNA expression are significantly higher in samples of non-cancerous liver tissue showing chronic hepatitis or cirrhosis than in normal liver tissue, and are even higher in HCCs.[47,48] The incidence of DNMT1 overexpression in HCCs is significantly correlated with poorer tumor differentiation and portal vein involvement.[49] Moreover, the recurrence-free and overall survival rates of patients with HCCs showing DNMT1 overexpression are significantly lower than those of patients with HCCs that do not.[49]

As mentioned above, at least a proportion of peripheral pancreatic ductal epithelia with an inflammatory background may be at the precancerous stage. The incidence of DNMT1 protein expression increases with progression from peripheral pancreatic ductal epithelia with an inflammatory background, to PanIN, to well-differentiated ductal carcinoma, and finally to poorly differentiated ductal carcinoma of the pancreas, in comparison with normal peripheral pancreatic duct epithelia.[50] DNMT1 overexpression in ductal carcinomas of the pancreas is significantly correlated with the extent of invasion to the surrounding tissue, an advanced stage, and poorer patient outcome.[50] The average number of methylated tumor-related genes in microdissected specimens of peripheral pancreatic ductal epithelia with an inflammatory background, PanIN, and ductal carcinoma was significantly correlated with the level of DNMT1 protein expression examined immunohistochemically in precisely microdissected areas.[37]

Expression levels of DNMT1 mRNA and protein are significantly correlated with poorer differentiation and the CIMP, a cancer phenotype characterized by accumulation of DNA methylation of C-type CpG islands,[51,52] in stomach cancers,[53] but no such association has been observed for the expression of DNMT2, DNMT3a, or DNMT3b.[54] Epstein–Barr virus infection in stomach cancers is significantly associated with marked accumulation of DNA methylation of C-type CpG islands and overexpression of DNMT1 protein.[53] *Helicobacter pylori* infection, another etiologic factor for stomach carcinogenesis, has also been reported to strongly promote regional DNA hypermethylation[55] but is not correlated with DNMT1 expression levels.[53]

It is debatable whether increased DNMT1 expression is due to an increase in the proportion of dividing cells or to an acute increase of DNMT1 expression per individual cancer cell. Immunohistochemical examinations have clearly revealed that the incidence of nuclear DNMT1 immunoreactivity is already higher in non-cancerous urothelia showing no remarkable histological changes obtained from patients with UCs, which may already be exposed to carcinogens in the urine but in which the PCNA labeling index had not yet increased, compared to that in normal urothelia from patients without UCs, indicating that

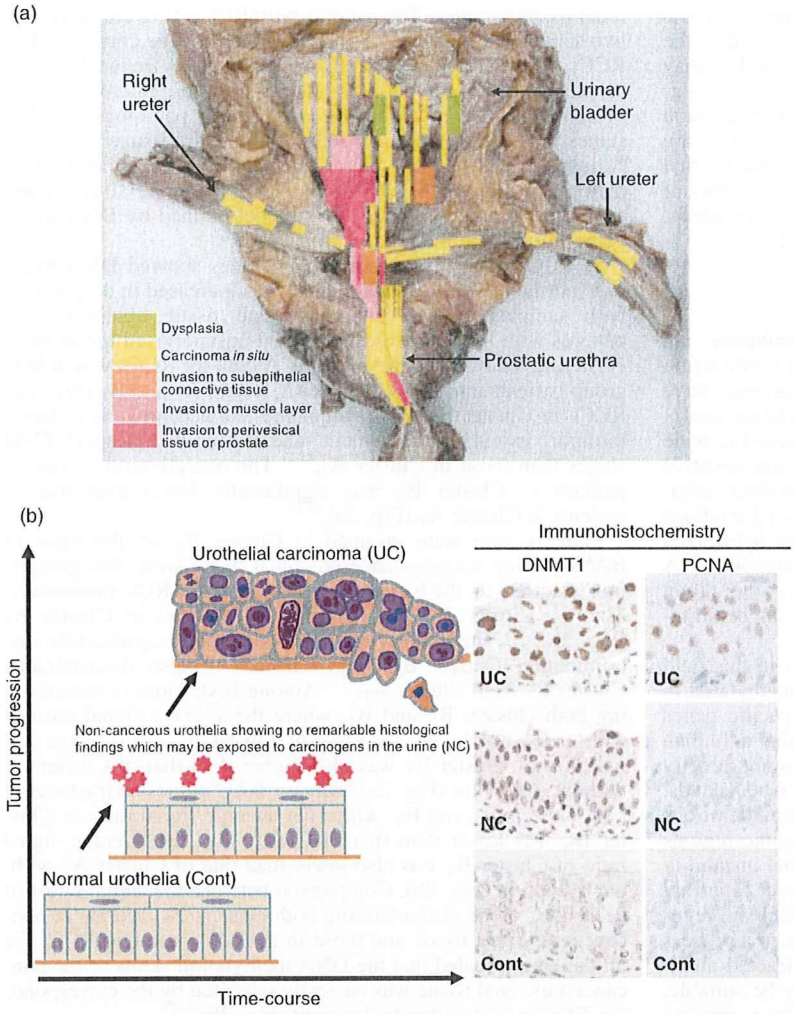

Fig. 1. Overexpression of DNA methyltransferase (DNMT) 1 protein during multistage urothelial carcinogenesis. (a) Specimen obtained by radical cystectomy for multiple urothelial carcinomas (UCs) of the urinary bladder, bilateral ureters, and prostatic urethra. UCs are clinically remarkable because of their multicentricity and tendency to recur: synchronously or metachronously multifocal UCs often develop in individual patients.[38] A possible mechanism for such multiplicity is the "field effect." Even non-cancerous urothelia showing no remarkable histological changes obtained from patients with UCs can be considered precancerous, because they may be exposed to carcinogens in the urine. (b) Immunohistochemical examination for DNMT1 and proliferating cell nuclear antigen (PCNA) in tissue specimens. The incidence of nuclear DNMT1 immunoreactivity had already increased in non-cancerous urothelia showing no remarkable histological changes obtained from patients with UCs (NC), where the PCNA labeling index had not yet increased, compared to that in normal urothelia obtained from patients without UCs (Cont), indicating that DNMT1 overexpression preceded any increase of cell proliferative activity.[56] The intensity of nuclear DNMT1 immunoreactivity was further increased in UCs.[56]

DNMT1 overexpression preceded increased cell proliferative activity (Fig. 1b).[56] The incidence of nuclear DNMT1 immunoreactivity showed a further and progressive increase in dysplastic urothelia, and during transition to UCs (Fig. 1b).[56] Among all examined microdissected specimens of non-cancerous urothelia showing no remarkable histological changes from patients with UCs, or dysplastic urothelia and UCs, accumulation of DNA methylation of C-type CpG islands was significantly correlated with the level of DNMT1 protein expression.[39]

Thus DNMT1 overexpression participates not only in the precancerous stage but also in the malignant progression of various cancers, and has a prognostic impact on patients. DNMT1 overexpression is frequently associated with CIMP of cancers. Although the maintenance activities of DNMT1 are related to its *in vitro* preference for hemimethylated substrates, excessive amounts of DNMT1 in comparison to PCNA may participate in *de novo* methylation of CpG islands. The molecular mechanisms that target DNMT1 to unmethylated substrates in cancers need to be clarified.

Splicing alteration of DNMT3b and DNA hypomethylation in pericentromeric satellite regions. DNA hypomethylation in pericentromeric satellite regions is known to result in centromeric decondensation and enhanced chromosome recombination. In HCCs[57] and UCs,[58] DNA hypomethylation of these regions is correlated with copy number alterations on chromosomes 1 and 9, respectively, where satellite regions are rich. DNMT3b is required for DNA methylation of pericentromeric satellite regions in early mouse embryos, and germline mutations of the *DNMT3b* gene have been reported in patients with immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome, a rare recessive autosomal disorder characterized by DNA hypomethylation of pericentromeric satellite regions.[59] The major splice variant of DNMT3b in normal liver tissue samples is DNMT3b3, which possesses the conserved catalytic domains.[60] DNMT activity of human DNMT3b3 has been confirmed *in vitro*.[61] In contrast, DNMT3b4 lacks the conserved catalytic domains, although it retains the *N*-terminal domain required for targeting to heterochromatin sites. Samples of normal liver tissue show only a trace level of DNMT3b4 expression.[60] The levels of DNMT3b4 mRNA expression and the ratio of DNMT3b4 mRNA to DNMT3b3 in samples of non-cancerous liver tissue obtained from patients with HCCs, and in HCCs themselves, are significantly correlated with the degree of DNA hypomethylation in pericentromeric satellite regions.[60] DNA demethylation on satellite 2 has been observed in DNMT3b4-transfected human epithelial 293 cells.[60] As DNMT3b4 lacking DNMT activity competes with DNMT3b3 for targeting to pericentromeric satellite regions, DNMT3b4 overexpression may lead to chromosomal instability through induction of DNA hypomethylation in such regions.

doi: 10.1111/j.1349-7006.2009.01383.x

Furthermore, the growth rate of DNMT3b4 transfectants is approximately double that of mock-transfectants soon after the introduction of DNMT3b4, when chromosomal instability may not yet have accumulated.[62] Genes implicated in interferon signaling including signal transducer and activator of transcription (STAT) 1, which acts as an effector of interferon signaling, are upregulated in DNMT3b4 transfectants,[62] suggesting that DNMT3b may act to maintain the DNA methylation status of not only pericentromeric satellite regions but also specific genes, probably in cooperation with DNMT1, in cancer cells.

Genome-wide DNA methylation profiling

DNA methylation profiles in precancerous conditions are inherited by cancers. The above findings that DNA methylation alterations are associated with multistage carcinogenesis have prompted us to carry out genome-wide DNA methylation analysis of tissue specimens. Recently, analysis on a genomic-wide scale has become possible using DNA methylation-sensitive restriction enzyme-based or anti-methyl-cytosine antibody affinity techniques that enrich methylated and unmethylated fractions of genomic DNA.[63,64] These fractions can then be hybridized to DNA microarrays or sequenced. Ultra-high-throughput DNA sequencing technologies are being introduced for the direct sequencing of enriched, methylated fragments or for bisulfite-converted genomic sequencing.[65]

We have used BAMCA.[66–69] Many researchers in this field use the promoter arrays to identify genes that are methylated in cancer cells. However, the promoter regions of specific genes are not the only target of DNA methylation alterations in human cancers. DNA methylation status in genomic regions not directly participating in gene silencing, such as the edges of CpG islands, may be altered at the precancerous stage before the alterations of the promoter regions themselves occur.[70] Genomic regions in which DNA hypomethylation affects chromosomal instability may not be contained in promoter arrays. Moreover, aberrant DNA methylation of large chromosome regions, which are regulated in a coordinated manner in human cancers due to a process of long-range epigenetic silencing, has recently attracted attention.[71] Therefore, we used a BAC array that may be suitable, not for focusing on specific promoter regions, but for overviewing the DNA methylation status of individual large regions among all chromosomes.

When BAMCA methods were applied to samples of non-cancerous renal tissue obtained from patients with clear cell RCCs, many BAC clones showed DNA hypo- or hypermethylation in comparison to normal renal tissue samples from patients without any primary renal tumors.[72] RCCs are usually well demarcated and covered by a fibrous capsule, and hardly ever contain fibrous stroma between cancer cells (Fig. 2a). We were therefore able to obtain cancer cells of high purity from surgical specimens, avoiding contamination with either non-cancerous epithelial cells or stromal cells (Fig. 2a). Therefore, the DNA methylation alterations observed in samples of non-cancerous renal tissue from patients with RCCs cannot be attributable to contamination during sampling. Moreover, DNA methylation alterations in non-cancerous renal tissue did not depend on the distance from the RCC itself to the site from which the non-cancerous renal tissue samples were taken. Because of the lack of any remarkable histological changes or any association with chronic inflammation and persistent infection with viruses or other pathogenic microorganisms, precancerous conditions in the kidney have rarely been described. However, from the viewpoint of DNA methylation, we can consider that non-cancerous renal tissue from patients with RCCs is already at the precancerous stage, showing genome-wide DNA methylation alterations.

We then carried out two-dimensional unsupervised hierarchical clustering analysis based on BAMCA data for non-cancerous renal tissue samples. The patients with RCCs were clustered into two subclasses, clusters A_N and B_N (Fig. 2a). The corresponding RCCs of patients in Cluster B_N showed more frequent macroscopically evident renal vein tumor thrombi, microscopically evident vascular involvement, and higher pathological TNM stages than those in Cluster A_N.[72] The overall survival rate of patients in Cluster B_N was significantly lower than that of patients in Cluster A_N (Fig. 2a).[72] Tumor aggressiveness and even patient outcome might thus be determined by DNA methylation profiles at the precancerous stage.

In RCCs themselves, more BAC clones showed DNA hypo- or hypermethylation, and its degree was increased in comparison with samples of non-cancerous renal tissue obtained from patients with RCCs. Two-dimensional unsupervised hierarchical clustering analysis based on BAMCA data for RCCs was able to group patients into two subclasses, Clusters A_T and B_T (Fig. 2a). RCCs in Cluster B_T more frequently showed renal vein tumor thrombi, vascular involvement, and higher pathological TNM stages than those in Cluster A_T.[72] The overall survival rate of patients in Cluster B_T was significantly lower than that of patients in Cluster A_T (Fig. 2a).[72]

Patients who were grouped in Cluster B_N on the basis of BAMCA data for non-cancerous renal tissue were also grouped in Cluster B_T on the basis of BAMCA data for RCC themselves. That is, Cluster B_N was completely included in Cluster B_T (Fig. 2b).[72] The majority of the BAC clones significantly discriminating Cluster B_N from Cluster A_N also discriminated Cluster B_T from Cluster A_T.[72] Among BAC clones characterizing both clusters B_N and B_T, where the average signal ratio of Cluster B_N was higher than that of Cluster A_N, the average signal ratio of Cluster B_T was also higher than that of Cluster A_T without exception (Fig. 2b). Among BAC clones characterizing both clusters B_N and B_T, where the average signal ratio of Cluster B_N was lower than that of Cluster A_N, the average signal ratio of Cluster B_T was also lower than that of Cluster A_T without exception (Fig. 2b). Comparison between the signal ratios of each BAC clone characterizing both clusters B_N and B_T in non-cancerous renal tissue and those in the corresponding RCCs for all patients revealed that the DNA methylation status of the non-cancerous renal tissue was basically inherited by the corresponding RCC in each individual patient (Fig. 2b).[72]

In non-cancerous renal tissue showing no remarkable histological changes and consisting mainly of renal tubules with specialized functions, no progenitor cell is able to gain a growth advantage, and clonal expansion is unable to occur. Therefore, the distinct DNA methylation profile of Cluster B_N, which is clinicopathologically valid, cannot be established through the selection of one of a number of random DNA methylation profiles in non-cancerous renal tissue in patients with clear cell RCCs, and instead may be established through distinct target mechanisms. As the DNA methylation profiles in Cluster B_T are shared by phenotypically similar patients, who all suffer from clinicopathologically aggressive tumors and show a poor outcome, DNA methylation alterations in at least a proportion of the BAC regions characterizing Cluster B_T cannot be passenger changes. It is clear that cancer itself can induce alterations in DNA methylation. However, DNA methylation alterations of BAC regions characterizing Cluster B_T may significantly participate in carcinogenesis, as the DNA methylation profile in Cluster B_N was established at a very early and precancerous stage of carcinogenesis and inherited during progression of the cancers themselves as Cluster B_T. At least a proportion of DNA methylation alterations at the precancerous stage may be "epigenetic gatekeepers"[21] and which allow time for further epigenetic and genetic alterations including genetic gatekeeper mutations (Fig. 3).

In fact, when the DNA methylation status of C-type CpG islands was examined,[73] the average number of methylated