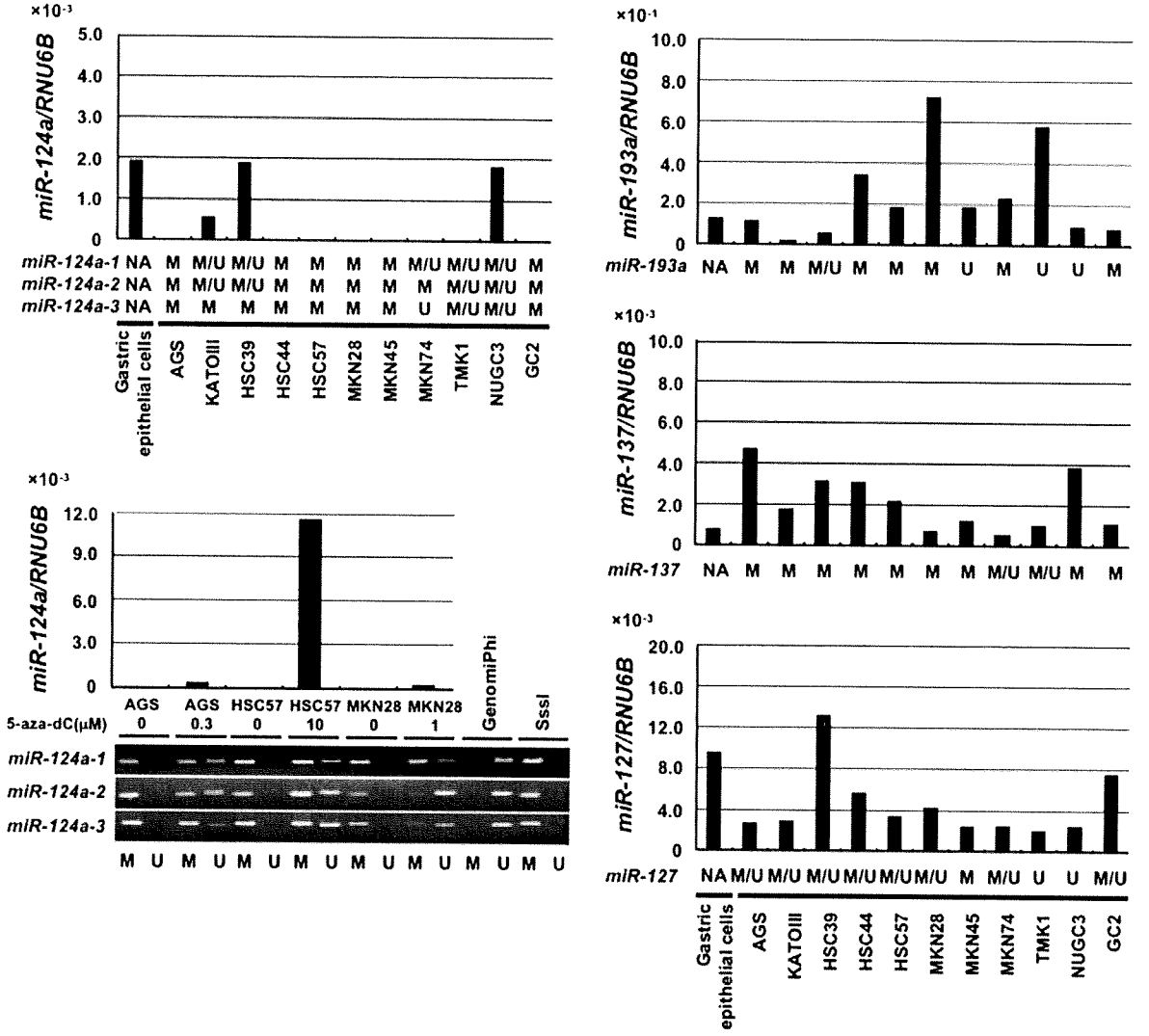

FIGURE 2 – Expression levels of miRNAs in gastric epithelial cells and gastric cancer cell lines. Expression levels were analyzed by quantitative RT-PCR, and normalized to *RNU6B* expression. Gastric epithelial cells were obtained by gland isolation technique from noncancerous tissues of eight gastric cancer patients, and average expression levels of the eight patient samples are shown. Results of MSP were duplicated from Figure 1 for convenience. M, M/U and U represent the presence of only methylated DNA, both methylated and unmethylated DNA, and only unmethylated DNA, respectively. NA, not applicable. Only *miR-124a* showed consistent repression in cell lines without unmethylated DNA molecules. After treatment by 5-aza-dC, *miR-124a* was re-expressed abundantly in HSC57, in association with demethylation, and in AGS and MKN28.

TaqMan® MicroRNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA). Real-time PCR was performed using the ABI Prism 7300 Fast Real-Time PCR System (Applied Biosystems). Expression levels of target miRNAs were normalized to that of a small nuclear RNA *RNU6B* transcript.

Statistical analysis

A difference in mean methylation levels was analyzed by the *t*-test Welch method, and differences in methylation incidence in gastric cancer tissues were analyzed by the chi-square test. Correlation between the age and methylation levels of miRNA genes, and correlation between methylation levels of each gene were analyzed using Spearman's rank correlation coefficient. All the analyses were performed using SPSS (SPSS, Inc., Chicago, IL), and the results were considered significant when a p value less than 0.05 was obtained by two-sided tests.

Results

Identification of miRNAs silenced in gastric cancer cell lines

Six genes of four miRNAs (*miR-124a, miR-137, miR-193a* and *miR-127*) were reported to have a tumor-suppressive function and be controlled by DNA methylation in colon, bladder and oral cancers.[22–24] We first analyzed methylation statuses of their putative promoter regions in 11 gastric cancer cell lines and two normal gastric mucosae of healthy individuals without *H. pylori*. It was found that *miR-124a-1, miR-124a-2, miR-124a-3* and *miR-137* were unmethylated in the normal gastric mucosae, but were completely methylated (no unmethylated DNA molecules detected)

frequently in the cell lines (Figs. 1*b*–1*d*). *miR-193a* was partially methylated in one of the two normal gastric mucosae, and completely methylated frequently in gastric cancer cell lines. In contrast, *miR-127* was completely methylated in the normal gastric mucosae, but unmethylated in the gastric cancer cell lines.

We then examined the effect of methylation of the putative promoter regions on miRNA expression (*miR-124a* for *miR-124a-1*, *miR-124a-2* and *miR-124a-3* genes; *miR-137*; *miR-193a*; *miR-127*) in the 11 gastric cancer cell lines and gastric epithelial cells obtained by the gland isolation technique (Fig. 2). *miR-124a* was consistently unexpressed in six cell lines with simultaneous methylation of its three isoforms (*miR-124a-1*, *miR-124a-2* and *miR-124a-3*), but was expressed in the gastric epithelial cells. In contrast, *miR-137*, *miR-193a*, and *miR-127* were expressed even in cell lines with complete methylation. This showed that these three miRNA genes were not silenced by their "promoter" methylation, and indicated that, in contrast, *miR-124a* was silenced by promoter methylation of its three isoforms.

Methylation-silencing of *miR-124a* was further confirmed by analyzing its re-expression in association with its promoter demethylation after treatment with a demethylating agent, 5-aza-dC, in three cell lines (AGS, HSC57 and MKN28). Re-expression and appearance of unmethylated DNA molecules were observed in all the three cell lines, HSC57 being prominent. This further orted that *miR-124a* was methylation-silenced.

The presence of miR-124a *methylation in primary gastric cancers*

Since methylation-silencing was identified only for *miR-124a*, we analyzed methylation levels of its three genes (*miR-124a-1*, *miR-124a-2* and *miR-124a-3*), along with a representative protein-coding gene (*LOX*), in 28 primary gastric cancer tissues (13 intestinal and 15 diffuse types) by qMSP. The fact that densely methylated DNA molecules were being measured was confirmed by bisulfite sequencing (Supp. Info. Fig. 1). *miR-124a-1* showed a distribution of methylation levels similar to *LOX*, some having no methylation and the others having various levels of methylation (Fig. 3*a*). This was consistent with our previous finding that cancer samples could be essentially classified into two groups (cancers with and without methylation), and that the various degrees of methylation levels in methylation-positive cancer samples were mainly due to contamination of normal cells.[18] On the other hand, *miR-124a-2* and *miR-124a-3* showed a unimodal distribution of methylation levels, suggesting that they are susceptible to methylation induction in cancer tissues. Using a cut-off value of 6%, as in previous reports,[31,32] *miR-124a-1*, *miR-124a-2* and *miR-124a-3* were methylated in 11, 23 and 26 of the 28 samples, respectively. Between the two histological types, the incidences of methylation were the same for *miR-124a-1*, *miR-124a-2* and *miR-124a-3* (p = 0.95, 0.84 and 0.67) (Supp. Info. Fig. 2a).

We further analyzed an association between methylation and expression of *miR-124a* in an additional 19 gastric cancer samples. Using a cut-off value of 6%, eight samples had methylation of all the three *miR-124a* genes, and the other 11 samples had methylation of only one or two genes and retained at least one unmethylated gene. *miR-124a* was barely expressed in all the eight samples with methylation of the three genes whereas it was expressed in 5 of 11 cancer samples with at least one unmethylated gene (Fig. 3*b*).

Accumulation of methylation in H. pylori *positive gastric mucosae, and its association with gastric cancer risk*

Methylation levels of *miR-124a-1*, *miR-124a-2* and *miR-124a-3*, again along with *LOX*, were analyzed by qMSP in gastric mucosae of 56 healthy volunteers (28 volunteers with *H. pylori* and 28 without) and noncancerous gastric mucosae of 45 gastric cancer patients (29 patients with *H. pylori* and 16 without) (Fig. 3*b*). Among the healthy volunteers, the mean methylation levels of *miR-124a-1*, *miR-124a-2*, *miR-124a-3* and *LOX* in the *H. pylori*-positive individuals were 13.1-, 7.8-, 8.9- and 46.7-fold, respectively, as high as those in *H. pylori*-negative individuals. This showed that *H. pylori* infection was associated with aberrant methylation of not only protein-coding genes but also miRNA genes.

Next, methylation levels in gastric mucosae of healthy volunteers were compared with those of noncancerous gastric mucosae of gastric cancer patients. Since potent methylation induction by *H. pylori* can mask a difference in *H. pylori*-positive individuals, the comparison was made among *H. pylori*-negative individuals only (28 healthy volunteers and 16 gastric cancer patients) (Table II; Fig. 3*c*). The mean methylation levels of the three miRNA genes and *LOX* were much higher in noncancerous gastric mucosae of gastric cancer patients than those of gastric mucosae of healthy volunteers (15.5-, 7.2-, 13.3- and 24.7-fold, respectively). Between the two histological types, the mean methylation levels were not different (Supp. Info. Fig. 2b).

Correlations among methylation levels of miRNA genes and *LOX* were examined by calculating correlation coefficients. Correlations among the three miRNA genes were very strong, but correlations between a miRNA gene and *LOX* were weak or absent (Table III; Supp. Info. Fig. 3).

No effect of age and sex on methylation levels on miRNA genes

Methylation of various CGIs is reported to be correlated with age.[33,34] Also, males have twice as high an incidence of gastric cancers as females.[1] In *H. pylori*-negative healthy volunteers, methylation levels of *miR-124a-1*, *miR-124a-2 and miR-124a-3* were not correlated with age (Spearman correlation test: r = 0.19, 0.01 and 0.29; p = 0.35, 0.94 and 0.15), and not associated with sex (p = 0.05, 0.68 and 0.19). Also, in *H. pylori*-positive healthy volunteers, methylation levels were not correlated with age (r = 0.13, 0.18 and −0.1; p = 0.51, 0.35 and 0.51), and not associated with sex (p = 0.70, 0.20 and 0.67).

Discussion

The present study showed that significantly higher methylation levels of three miRNA genes (*miR-124a-1*, *miR-124a-2* and *miR-124a-3*) were present in gastric mucosae of *H. pylori*-positive healthy volunteers, indicating that *H. pylori* infection can induce DNA methylation of miRNA genes, in addition to protein-coding genes. Moreover, it was also shown that methylation levels of the miRNA genes in noncancerous gastric mucosae of gastric cancer patients were higher than those in gastric mucosae of healthy volunteers among *H. pylori* negative individuals, indicating that miRNA silencing is involved in the formation of a field defect for gastric cancers. To our knowledge, the presence of miRNA silencing in a field for cancerization was shown here for the first time.

Recent studies demonstrated that expression of some miRNAs is regulated by epigenetic mechanisms.[24,35] From six miRNA genes that were reported to be silenced by promoter methylation and to have tumor-suppressor functions, we were able to confirm that three genes of *miR-124a* were methylation-silenced in gastric cancer cell lines. The other three genes, *miR-137*, *miR-193a* and *miR-127*, were expressed even in cell lines with complete methylation, and were unlikely to be silenced by promoter methylation in gastric cancers. Since methylation of putative promoter regions consistently represses transcription of their downstream genes,[29,30] the presence of the expression of the three genes in gastric cancer cell lines with complete methylation of their "promoter" CGI indicated that the three genes had additional or alternative promoters.

Lujambio *et al.*[22] discovered that *miR-124a* was silenced by promoter methylation after screening 320 miRNA genes. They also found that *miR-124a* down-regulates CDK6, a demonstrated oncogene involved in cell cycle progression and differentiation, and induces hypophosphorylation of RB.[22] Therefore, it is possible that *miR-124a* silencing is also involved in gastric carcinogenesis, and the presence of its silencing in noncancerous tissues

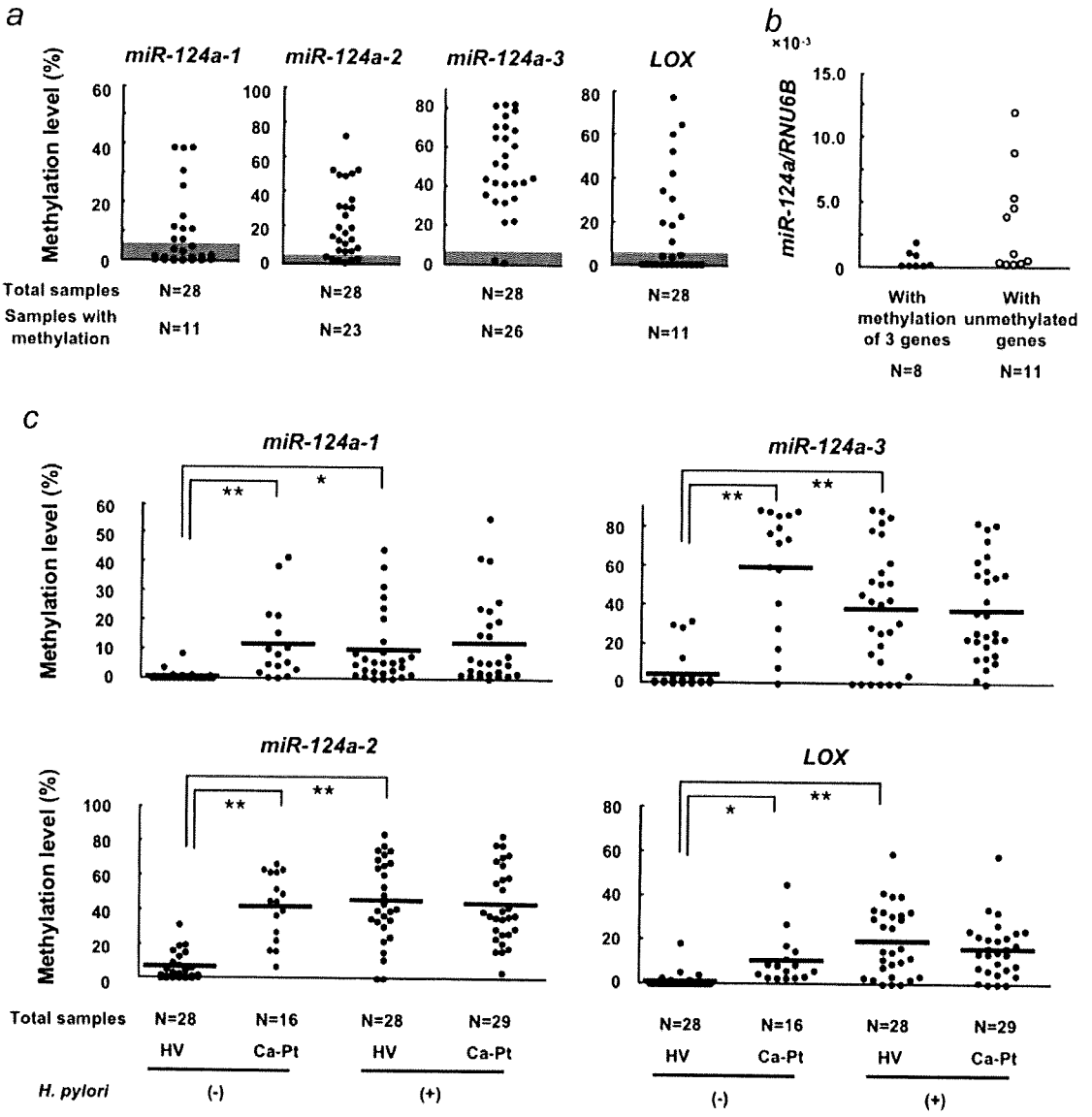

FIGURE 3 – Methylation levels of *miR-124a-1*, *miR-124a-2*, *miR-124a-3* and *LOX* in gastric mucosae of healthy volunteers, noncancerous mucosae of gastric cancer patients, and cancer tissues. (*a*) Distribution of methylation levels in gastric cancers. Gray areas show samples with methylation levels below the cut-off value of 6%. (*b*) Expression of *miR-124a* in eight cancer samples with methylation (three *miR-124a* genes, methylation positive) and 11 cancer samples with at least one unmethylated gene (*miR-124a-1* or *miR124a-3*). Five of the 11 cancers with unmethylated genes had high *miR-124a* expression levels. (*c*) Distribution of methylation levels in gastric mucosae of healthy volunteers (HV) and noncancerous mucosae of gastric cancer patients (Ca-Pt). A horizontal line represents a mean methylation level for each group. Among the healthy volunteers, *H. pylori*-positive individuals had 7.8–46.7-fold as high methylation levels as *H. pylori*-negative individuals (*$p < 0.005$; **$p < 0.001$). Among the *H. pylori*-negative individuals, noncancerous gastric mucosae of gastric cancer patients had 7.2–24.7-fold as high methylation levels as gastric mucosae of healthy volunteers (*$p < 0.005$; **$p < 0.001$).

TABLE II – MEAN METHYLATION LEVELS AND STANDARD DEVIATIONS OF THE FOUR GENES IN GASTRIC MUCOSAE OF HEALTHY VOLUNTEERS AND GASTRIC CANCER PATIENTS

		N	*miR-124a-1*	*miR-124a-2*	*miR-124a-3*	*LOX*
H.pylori(−)	(1) Healthy volunteers	28	0.76 ± 1.70	5.75 ± 7.73	4.45 ± 9.29	0.43 ± 1.22
	(2) Gastric cancer patients	16	11.82 ± 12.94	41.66 ± 19.33	59.42 ± 30.71	10.64 ± 11.33
H.pylori(+)	(3) Healthy volunteers	28	9.96 ± 12.28	44.79 ± 23.96	39.66 ± 30.08	20.10 ± 15.72
	(4) Gastric cancer patients	29	12.28 ± 14.31	46.33 ± 28.36	37.48 ± 25.13	15.93 ± 12.58
p value	(1) *vs.* (3)		<0.001	$<10^{-8}$	$<10^{-5}$	$<10^{-6}$
	(1) *vs.* (2)		0.004	$<10^{-5}$	$<10^{-6}$	0.003
	(3) *vs.* (4)		0.51	0.77	0.77	0.28

TABLE III – CORRELATION AMONG METHYLATION LEVEL OF *miR-124A-1, miR-124A-2, miR-124A-3* AND *LOX*

	miR-124a-1		*miR-124a-2*		*miR-124a-3*		*LOX*	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
miR-124a-1	–	–	0.70	$<10^{-15}$	0.77	$<10^{-20}$	0.03	0.76
miR-124a-2	0.70	10^{-15}	–	–	0.72	$<10^{-16}$	0.37	$<10^{-3}$
miR-124a-3	0.77	$<10^{-20}$	0.72	$<10^{-16}$	–	–	0.20	0.04

r, correlation coefficient.

could be directly associated with predisposition to developing gastric cancers.

As repeatedly shown by epidemiological studies, the majority of *H. pylori*-negative individuals with a gastric cancer are considered to have past exposure to *H. pylori*.[36] Methylation levels of protein-coding genes, including *LOX*, in the gastric mucosae of individuals with past infection (gastric cancer patients without *H. pylori*) were lower than those of individuals with current infection (both healthy volunteers and gastric cancer patients) in our previous study.[11] Actually, incidences of aberrant methylation and methylation levels of *CDH1* are reported to decrease after the eradication of *H. pylori*,[19,37] showing that DNA methylation in gastric mucosae decreases when *H. pylori* infection discontinues. Interestingly, methylation levels of the three miRNA genes in the gastric mucosae of individuals with past infection by *H. pylori* (gastric cancer patients without *H. pylori*) did not decrease compared with those of individuals with current infection (healthy volunteers and gastric cancer patients). Since aberrant methylation induced in stem cells is expected to persist even after *H. pylori* infection discontinues, DNA methylation of these miRNA genes might be relatively more easily induced in gastric stem cells than those of protein-coding genes.

In conclusion, our data indicated that DNA methylation of certain miRNA genes was associated with *H. pylori* infection, in addition to protein-coding genes, and involved in the formation of field defect for gastric cancers.

Acknowledgements

The authors thank Drs. Eriko Okochi-Takada and Satoshi Yamashita for their advice.

References

1. Ushijima T, Sasako M. Focus on gastric cancer. Cancer Cell 2004;5:121–5.
2. Nakajima T, Oda I, Gotoda T, Hamanaka H, Eguchi T, Yokoi C, Saito D. Metachronous gastric cancers after endoscopic resection: how effective is annual endoscopic surveillance? Gastric Cancer 2006;9: 93–8.
3. Lee KJ, Inoue M, Otani T, Iwasaki M, Sasazuki S, Tsugane S. Gastric cancer screening and subsequent risk of gastric cancer: a large-scale population-based cohort study, with a 13-year follow-up in Japan. Int J Cancer 2006;118:2315–21.
4. Kang CH, Yu TJ, Hsieh HH, Yang JW, Shu K, Huang CC, Chiang PH, Shiue YL. The development of bladder tumors and contralateral upper urinary tract tumors after primary transitional cell carcinoma of the upper urinary tract. Cancer 2003;98:1620–6.
5. Poon RT, Fan ST, Ng IO, Lo CM, Liu CL, Wong J. Different risk factors and prognosis for early and late intrahepatic recurrence after resection of hepatocellular carcinoma. Cancer 2000;89:500–7.
6. Yokoyama A, Omori T, Yokoyama T, Sato Y, Kawakubo H, Maruyama K. Risk of metachronous squamous cell carcinoma in the upper aerodigestive tract of Japanese alcoholic men with esophageal squamous cell carcinoma: a long-term endoscopic follow-up study. Cancer Sci 2008;99:1164–71.
7. Hartman M, Czene K, Reilly M, Adolfsson J, Bergh J, Adami HO, Dickman PW, Hall P. Incidence and prognosis of synchronous and metachronous bilateral breast cancer. J Clin Oncol 2007;25:4210–6.
8. Colice GL, Rubins J, Unger M. Follow-up and surveillance of the lung cancer patient following curative-intent therapy. Chest 2003; 123:S272–83.
9. Rex DK, Kahi CJ, Levin B, Smith RA, Bond JH, Brooks D, Burt RW, Byers T, Fletcher RH, Hyman N, Johnson D, Kirk L, et al. Guidelines for colonoscopy surveillance after cancer resection: a consensus update by the American Cancer Society and the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology 2006;130:1865–71.
10. Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications. Cancer Res 2003;63:1727–30.
11. Maekita T, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M, Arii K, Kaneda A, Tsukamoto T, Tatematsu M, Tamura G, Saito D, 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.
12. Nakajima T, Maekita T, Oda I, Gotoda T, Yamamoto S, Umemura S, Ichinose M, Sugimura T, Ushijima T, Saito D. Higher methylation levels in gastric mucosae significantly correlate with higher risk of gastric cancers. Cancer Epidemiol Biomarkers Prev 2006;15:2317–21.
13. 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;32:970–9.
14. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, Houlihan PS, Krouse RS, Prasad AR, Einspahr JG, Buckmeier J, Alberts DS, et al. MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst 2005;97:1330–8.
15. Ishii T, Murakami J, Notohara K, Cullings HM, Sasamoto H, Kambara T, Shirakawa Y, Naomoto Y, Ouchida M, Shimizu K, Tanaka N, Jass JR, 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.
16. Yan PS, Venkataramu C, Ibrahim A, Liu JC, Shen RZ, Diaz NM, Centeno B, Weber F, Leu YW, Shapiro CL, Eng C, Yeatman TJ, et al. Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. Clin Cancer Res 2006;12:6626–36.
17. 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;119:288–96.
18. Enomoto S, Maekita T, Tsukamoto T, Nakajima T, Nakazawa K, Tatematsu M, Ichinose M, Ushijima T. Lack of association between CpG island methylator phenotype in human gastric cancers and methylation in their background non-cancerous gastric mucosae. Cancer Sci 2007;98:1853–61.
19. Perri F, Cotugno R, Piepoli A, Merla A, Quitadamo M, Gentile A, Pilotto A, Annese V, Andriulli A. Aberrant DNA methylation in non-neoplastic gastric mucosa of *H. Pylori* infected patients and effect of eradication. Am J Gastroenterol 2007;102:1361–71.
20. Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857–66.
21. Esquela-Kerscher A, Slack FJ. Oncomirs—microRNAs with a role in cancer. Nat Rev Cancer 2006;6:259–69.
22. Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setien F, Casado S, Suarez-Gauthier A, Sanchez-Cespedes M, Git A, Spiteri I, Das PP, et al. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. Cancer Res 2007;67:1424–9.
23. Kozaki K, Imoto I, Mogi S, Omura K, Inazawa J. Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. Cancer Res 2008;68:2094–105.
24. Saito Y, Liang G, Egger G, Friedman JM, Chuang JC, Coetzee GA, Jones PA. Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. Cancer Cell 2006;9:435–43.

25. 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: 178–87.
26. Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma—2nd English edition. Gastric Cancer 1998;1: 10–24.
27. Lauren P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. Acta Pathol Microbiol Scand 1965; 64:31–49.
28. 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.
29. Li B, Carey M, Workman JL. The role of chromatin during transcription. Cell 2007;128:707–19.
30. Lin JC, Jeong S, Liang G, Takai D, Fatemi M, Tsai YC, Egger G, Gal-Yam EN, Jones PA. Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island. Cancer Cell 2007;12:432–44.
31. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787–93.
32. Ogino S, Odze RD, Kawasaki T, Brahmandam M, Kirkner GJ, Laird PW, Loda M, Fuchs CS. Correlation of pathologic features with CpG island methylator phenotype (CIMP) by quantitative DNA methylation analysis in colorectal carcinoma. Am J Surg Pathol 2006;30: 1175–83.
33. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res 2001;61:3573–7.
34. Ushijima T, Okochi-Takada E. Aberrant methylations in cancer cells: where do they come from? Cancer Sci 2005;96:206–11.
35. Yang N, Coukos G, Zhang L. MicroRNA epigenetic alterations in human cancer: one step forward in diagnosis and treatment. Int J Cancer 2008;122:963–8.
36. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, Taniyama K, Sasaki N, Schlemper RJ. *Helicobacter pylori* infection and the development of gastric cancer. N Engl J Med 2001;345:784–9.
37. Miyazaki T, Murayama Y, Shinomura Y, Yamamoto T, Watabe K, Tsutsui S, Kiyohara T, Tamura S, Hayashi N. E-cadherin gene promoter hypermethylation in *H. pylori*-induced enlarged fold gastritis. Helicobacter 2007;12:523–31.

Int. J. Cancer: **124,** 905–910 (2009)

The presence of a methylation fingerprint of *Helicobacter pylori* infection in human gastric mucosae

Takeshi Nakajima[1,2], Satoshi Yamashita[1], Takao Maekita[3], Tohru Niwa[1], Kazuyuki Nakazawa[3] and Toshikazu Ushijima[1]*

[1]*Carcinogenesis Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan*
[2]*Endoscopy Division, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan*
[3]*Second Department of Internal Medicine, Wakayama Medical University, 811-1 Kimiidera, Wakayama-city, Wakayama 641-0012, Japan*

Aberrant DNA methylation is deeply involved in human cancers, but its inducers and targets are still mostly unclear. *Helicobacter pylori* infection was recently shown to induce aberrant methylation in gastric mucosae, and produce a predisposed field for cancerization. Here, we analyzed the presence of target genes in methylation induction by *H. pylori* and the mechanism for the gene specificity. Noncancerous gastric mucosae were collected from 4 groups of individuals (with and without a gastric cancer, and with and without current *H. pylori* infection; $N = 11$ for each group), and methylation of promoter CpG islands of 48 genes that can be methylated in gastric cancer cell lines was analyzed by methylation-specific PCR. In total, 26 genes were consistently methylated in individuals with current or past infection by *H. pylori*, whereas 7 genes were not methylated at all. In addition, 14 genes were randomly or intermediately methylated in individuals with gastric cancers and the remaining 1 gene was methylated in all the cases. The methylation-susceptible genes had significantly lower mRNA expression levels than the methylation-resistant genes. *H. pylori* infection did not induce mRNA and protein expression of DNA methyltransferases; *DNMT1*, *DNMT3A* or *DNMT3B*. Gene specificity was present in the induction of aberrant DNA methylation by *H. pylori* infection, and low mRNA expression, which could precede methylation, was one of the mechanisms for the gene specificity. These findings open up the possibility that a methylation fingerprint can be used as a novel marker for past exposure to a specific carcinogenic factor.

Key words: DNA methylation; epigenetic; fingerprint; *Helicobacter pylori*; molecular epidemiology

Aberrant DNA methylation is deeply involved in human cancer development and progression.[1] In some cancer types, such as gastric cancers, tumor-suppressor genes are more frequently inactivated by aberrant DNA methylation than by mutations.[2] Nevertheless, only limited information is available for inducers of aberrant DNA methylation, which include aging, viral infection and ulcerative colitis.[3,4] Also, almost no information is available for gene specificity in methylation induction by a specific factor. Using cancer tissues, it is very difficult to clarify an association between a specific inducer and methylation of a gene. Aberrant methylation of a gene can be present in cancer tissues because its methylation conferred a growth advantage although it was a rare and random event, or because its methylation was carried over from a precursor tissue to a cancer tissue since it was frequently induced in the precursor tissue. In contrast, using a noncancerous tissue, one can assess an effect of a methylation inducer by the fraction of cells with methylation in the polyclonal tissue.

Gastric mucosa infected by *Helicobacter pylori* is a useful model to examine the possible presence of gene specificity in methylation induction by a specific factor since *H. pylori* infection was recently shown to induce aberrant DNA methylation potently in gastric mucosae.[5] Moreover, the fraction of DNA molecules with aberrant methylation (methylation level) in gastric mucosae of individuals without current *H. pylori* infection was correlated with gastric cancer risk,[5,6] indicating that methylation in noncancerous tissues is related to gastric carcinogenesis. So far, 6 CpG islands in gene promoter regions methylated in gastric cancers[7] were analyzed, and all were methylated in gastric mucosae with current and past infection with *H. pylori*. However, it is unknown whether these 6 genes are preferentially methylated by *H. pylori* infection or *H. pylori* infection induces methylation of random genes.

In this study to analyze the presence of gene specificity for methylation induction, firstly we examined the methylation status of 48 promoter CpG islands in the noncancerous gastric mucosae of 4 groups of individuals (with and without a gastric cancer, and with and without current *H. pylori* infection). The 48 genes were selected as genes that can be methylation-silenced in gastric cancer cell lines[8] because the vast majority of CpG islands in gene promoter regions are not methylated at all in noncancerous tissues, and we had to newly select genes that have better chances to be methylated in noncancerous tissues. Secondly, we analyzed an association between susceptibility to methylation induction and mRNA expression levels in normal tissue without and with *H. pylori* infection.

Material and methods

Tissue samples and DNA/RNA extraction

For methylation analysis, (noncancerous) gastric mucosa samples were collected from 4 groups of individuals (with and without a gastric cancer, and with and without current *H. pylori* infection; $N = 11$ for each group, average age = 60.8 ± 13.8 years). For analysis of mRNA expression that determines gene specificity of methylation induction, we need to analyze the mRNA expression level in gastric mucosae free of methylation, which, once induced, will cause decreased gene transcription to avoid confusion between cause and consequence. Therefore, samples were collected from 11 healthy volunteers, who were considered to have less chance for methylation induction by *H. pylori* than elderly individuals (7 males and 4 females; 6 with *H. pylori* infection and 5 without; average age = 34.8 ± 3.1 years). Biopsy specimens were taken from one standard site of the stomach (antral regions in the lesser curvature) using sterilized biopsy forceps (Olympus, Tokyo, Japan). *H. pylori* infection status was analyzed by culture test (Eiken, Tokyo, Japan) and rapid urease test (Otsuka, Tokushima, Japan). All the materials were obtained with written informed consents, and the procedures were approved by the institutional review board. High molecular weight DNA was extracted by the standard phenol/chloroform method and total RNA was isolated using ISOGEN (Nippon Gene, Tokyo, Japan) and an RNeasy Mini kit (Qiagen, Valencia, CA).

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: The Ministry of Health, Labour and Welfare, Japan (for the Third-term Comprehensive Cancer Control Strategy, Pioneering Basic Research and Cancer Research).

***Correspondence to:** Carcinogenesis Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. Fax: +81 3 5565 1753. E-mail: tushijim@ncc.go.jp

Received 15 June 2008; Accepted after revision 11 September 2008
DOI 10.1002/ijc.24018
Published online 24 September 2008 in Wiley InterScience (www.interscience.wiley.com).

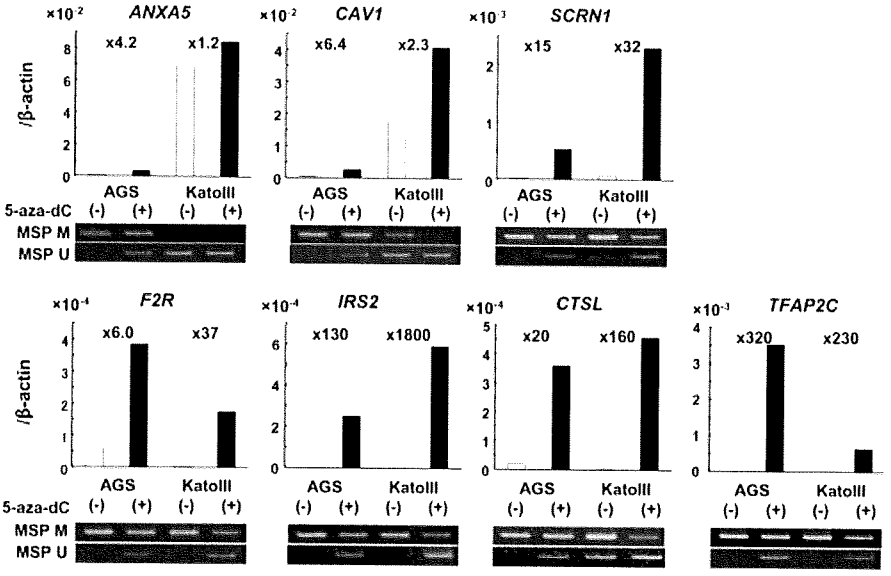

FIGURE 1 – Gene silencing due to methylation of the regions analyzed. mRNA expression and methylation were analyzed by real-time RT-PCR and MSP, respectively, in gastric cancer cell lines (AGS and KATO-III) before and after 5-aza-dC treatment. The fold increases after 5-aza-dC treatment is shown for each cell line. No or little mRNA expression in a cell line(s) without unmethylated DNA molecules and upregulation by the 5-aza-dC treatment was confirmed for the 7 genes randomly selected from the 48 genes.

Cell lines and 5-aza-dC treatment

Gastric cancer cell lines, AGS and KATO-III, were obtained from the Japanese Collection of Research Bioresources (Tokyo, Japan) and the American Type Culture Collection (Manassas, VA). For treatment with a demethylating agent, 5-aza-2′-deoxycitidine (5-aza-dC, Sigma, St. Louis, MO), cells were seeded on day 0, media containing 0.3 μM 5-aza-dC was freshly added on days 1 and 3, and cells were harvested on day 5. Genomic DNA and total RNA were isolated in the same way as the primary samples.

Bisulfite treatment and methylation-specific PCR

Bisulfite treatment was performed as previously described.[9] Briefly, DNA samples (1 μg each) digested by *Bam*HI were denatured in 0.3 N NaOH at 37°C for 15 min. The samples underwent 15 cycles of 30-sec denaturation at 95°C and 15-min incubation at 50°C in 3.1 N sodium bisulfite (pH 5.0) and 0.5 mM hydroquinone. The samples were desalted with the Wizard DNA Clean-Up system (Promega, Madison, WI), and desulfonated in 0.3 N NaOH. DNA was ethanol precipitated and dissolved in 40 μL of TE buffer.

Methylation-specific PCR (MSP) was performed with a primer set specific to the methylated or unmethylated sequence (M or U set), respectively,[8] using 2 μL of the sodium bisulfite-treated DNA. A region upstream of a putative transcriptional start site (200 bp or less) was analyzed, and CpG maps of all the genes are shown in the Supporting Information Figure 1. DNA methylated with *Sss*I methylase was used to determine a specific condition of PCR for the M set, and DNA amplified by a GenomiPhi DNA amplification kit (GE Healthcare Bio-Sciences) was used for the U set. A number of PCR cycles that would yield a minimal visible band was determined using these fully methylated DNA (for M primers) and fully unmethylated DNA (for U primers), and a further 4 cycles were added for actual analysis of test samples. Methylation levels were classified as none (−), low (+), high (++) according to the intensity of the band for methylated DNA molecules compared with that for unmethylated DNA, respectively.

Quantitative reverse transcription PCR

cDNA was synthesized from 1 μg of total RNA using a Superscript II kit (Life Technologies, Rockville, MD) with a random primer. Real-time PCR was performed using an iCycler Thermal Cycler (Bio-Rad Laboratories, Hercules, CA) with SYBR Green I (BioWhittaker Molecular Applications, Rockland, ME). The number of molecules of a specific gene in a sample was measured by comparing its amplification with that of standard samples, which contained 10^1–10^7 copies of the gene. The standard samples were produced by PCR amplification and purification using Zymo-Spin I™ Columns (Zymo Research, Orange, CA). The amount of the standard samples was measured by OD 260 nm and also by quantification of band intensities after electrophoresis. The mRNA quantity of each gene was normalized to that of β-actin. The primers and PCR conditions are shown in the Supporting Information Table 1. The difference of mRNA expression levels between 2 groups of genes was analyzed by the Welch *t*-test method (both sided).

Western blot analysis

Each 100 μg whole-cell lysate sample was subjected to SDS-PAGE (10% acrylamide gel) and blotted to PVDF membrane. DNMT1 and DNMT3A were detected using rabbit polyclonal antibody against human DNMT1 (NEB, Beverly, MA), human DNMT3A (Cell Signaling Technology, Danvers, MA), respectively at 1/1,000 dilution. DNMT3B was detected using goat polyclonal antibody against human DNMT3B (Santa Cruz Biotechnology, Santa Cruz, CA) at 1/500 dilution. Horse radish peroxidase-conjugated secondary antibody (antirabbit; Cell Signaling Technology, antigoat; Santa Cruz Biotechnology) was used at 1/5,000 dilution.

Results

Confirmation of gene silencing due to promoter CpG islands

The 48 genes consisted of 32 randomly and 16 arbitrarily selected genes from 421 genes that had been identified as methylation-silenced genes in a gastric cancer cell line using microarray analysis of cells treated with 5-aza-dC, and MSP analysis.[8] First,

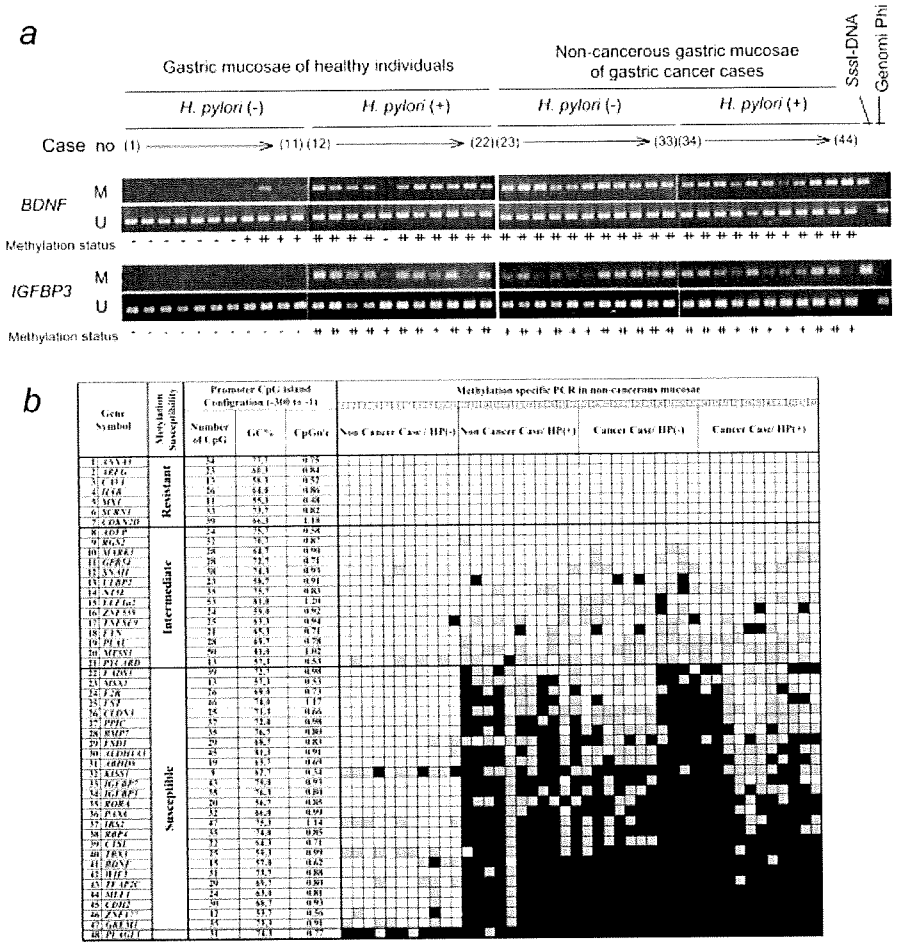

FIGURE 2 – Methylation profile of the 48 genes in noncancerous gastric mucosae. (*a*) Representative results of MSP. Samples 1–11, gastric mucosae of healthy individuals without *H. pylori* infection; 12–22, those with *H. pylori* infection; 23–33, noncancerous gastric mucosae of gastric cancer cases without *H. pylori* infection; and 34–44, those with *H. pylori* infection. Methylation levels were classified as none (−), low (+), high (++) according to the intensity of the band for methylated DNA molecules compared with that of fully methylated control DNA. (*b*) Overview of the results of all the 48 genes. The genes were aligned in the order of increasing numbers of individuals with methylation. Closed, hatched, and open boxes represent the methylation levels of high (++), low (+), and none (−), respectively. Rows 1–7, the 7 genes completely resistant to methylation induction in any cases; rows 8–21, genes methylated randomly or more frequently in individuals with cancers; and rows 22–47, genes susceptible to methylation induction by *H. pylori* infection. CpG island configuration (number of CpG sites, G+C content, and CpG score) in 300 bp upstream regions from transcription start sites is also shown. The presence of methylation-resistant and methylation-susceptible genes was clearly revealed. No clear difference in the CpG island configuration was observed between the 2 groups.

we analyzed mRNA expression of 7 of the 48 genes before and after 5-aza-dC treatment using real-time RT-PCR (Fig. 1). It was confirmed that no or little mRNA expression was present in cell lines without unmethylated DNA molecules and that mRNA expression was upregulated by the 5-aza-dC treatment.

Gene specificity in methylation induction by H. pylori infection in gastric mucosae

We then analyzed the methylation status of the promoter CpG islands of the 48 genes in the (noncancerous) gastric mucosae of 4 groups of individuals; those with and without *H. pylori* infection and with and without a gastric cancer. Since MSP can produce inconsistent results if inappropriately performed, we carefully selected a PCR cycle for each primer set so that false positive and negative results were not produced. We scored the methylation status as negative, weakly positive or positive by comparing the band density with that of a fully methylated control (representative results in Fig. 2*a*).

When all the genes were aligned in the order of number of samples with methylation (Fig. 2*b*), the 48 genes were divided into 3 groups: (*i*) 7 genes that were completely unmethylated in any of the 4 groups (genes 1–7 in Fig. 2*b*; methylation-resistant genes), (*ii*) 14 genes that were methylated randomly or more frequently in individuals with cancers (genes 9–21; intermediate genes); and (*iii*) 26 genes that were consistently methylated in the individuals with *H. pylori* infection or with a gastric cancer (genes 22–47; methylation-susceptible genes). The remaining 1 gene, *PLAGL1*, was methylated in all the individuals. This demonstrated that some genes are resistant to methylation induction by *H. pylori* infection while others are susceptible, namely the presence of gene specificity in methylation induction.

Lack of association between CpG island configuration and methylation susceptibility

The 48 genes analyzed here all had CpG islands in their promoter regions. However, based on recent reports,[10] there was a

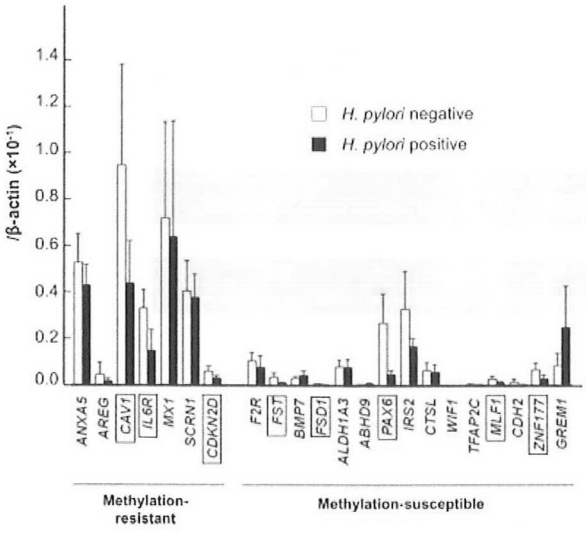

FIGURE 3 – The mRNA expression levels of genes resistant and susceptible to methylation induction. mRNA expression levels of 22 genes (7 resistant and 15 susceptible genes) in the noncancerous gastric mucosae of young healthy individuals with (closed columns) and without (open columns) *H. pylori* infection was analyzed by real-time RT-PCR. Error bar: standard deviation. The average mRNA expression level of methylation-resistant genes was much higher than that of methylation-susceptible genes among individuals without *H. pylori* infection (4.3×10^{-2} *vs.* 7.3×10^{-3}; $p = 0.0008$) and also among individuals with *H. pylori* infection (2.9×10^{-2} *vs.* 5.1×10^{-3}; $p = 0.0012$). The genes whose names are boxed showed a significant decrease in their mRNA expression levels by *H. pylori* infection ($p < 0.05$). Considering that all these 48 genes are those that can be methylated in gastric cancer cell lines, downregulation of mRNA expression could be involved in methylation induction.

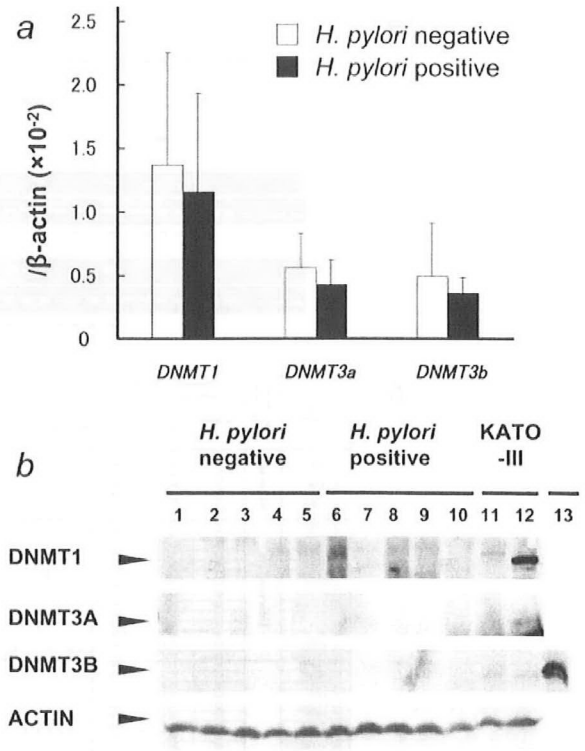

FIGURE 4 – The mRNA and protein expression levels of three DNA methyltransferases (*DNMT1*, *DNMT3A* and *DNMT3B*) in noncancerous gastric mucosae of young healthy individuals with and without *H. pylori* infection. (*a*) mRNA expression levels of *DMNTs*. Closed columns, individuals with *H. pylori* infection; open columns, those without. No significant increase was observed in the mRNA expression levels of these *DNMTs*. (*b*) Western blot analysis of DNMTs. For DNMT1 and DNMT3A, a stomach cancer cell line, KATO-III was used as a positive control (lane 12), and 5-aza-dC (1 μM)-treated KATO-III was used as a negative control (lane 11). ACTIN was used as a loading control. For DNMT3B, a commercially available positive control of DNMT3B (Santa Cruz, lane 13) was used. DNMT protein levels were below the detection limit in the noncancerous gastric mucosae of individuals without (lanes 1–5) and with (lanes 6–10) *H. pylori* infection, and no detectable increase was observed.

possibility that, even among CpG islands, their configurations (number of CpG sites, G+C content, and CpG score) might influence the susceptibility of individual genes to methylation induction by *H. pylori*. Therefore, we examined their configurations in 300 bp upstream regions from transcription start sites (Fig. 2*b*), which corresponded to the nucleosome-free region and whose methylation is critical for gene silencing.[11,12]

The number of CpG sites in the region was 29.2 ± 10.4 (mean ± standard deviation) and 25.4 ± 9.3 for the susceptible and resistant genes, respectively ($p = 0.38$). The G + C content was 68.4 ± 7.4 and 66.4 ± 7.9% for the susceptible and resistant genes, respectively ($p = 0.52$). The CpG score was 0.82 ± 0.18 and 0.75 ± 0.21 for the susceptible and resistant genes, respectively ($p = 0.40$). In short, no significant difference was present between the 2 groups.

Involvement of low mRNA expression levels in gene specificity in methylation induction

To investigate an association between the gene specificity in methylation induction and mRNA expression levels in gastric mucosae, we analyzed mRNA expression levels of all of the 7 methylation-resistant and 15 methylation-susceptible genes, which were randomly selected from the 26 methylation-susceptible genes. To compare mRNA expression levels among different genes, the numbers of cDNA molecules were measured by quantitative RT-PCR after accurate measurement of the weights (converted into the numbers of DNA molecules) of standard DNA samples of all the genes. mRNA expression levels were analyzed in the gastric mucosae of young healthy individuals with and without *H. pylori* infection, who were considered to have no or little methylation of the genes analyzed.

The average mRNA expression level of methylation-resistant genes was much higher than that of methylation-susceptible genes among individuals without *H. pylori* infection (4.3×10^{-2} *vs.* 7.3×10^{-3}; $p = 0.0008$) and also among individuals with *H. pylori* infection (2.9×10^{-2} *vs.* 5.1×10^{-3}; $p = 0.0012$) (Fig. 3). Three of the 7 resistant genes and 5 of the 15 susceptible genes showed a significant decrease of mRNA expression levels by *H. pylori* infection, but no genes showed significantly increased mRNA expression.

Expression levels of DNA methyltransferase

To gain an insight into how *H. pylori* infection induces aberrant methylation, we analyzed mRNA expression levels of maintenance DNA methyltransferase, *DNMT1*, and *de novo* methyltransferases, *DNMT3A* and *DNMT3B*, in the gastric mucosae with and without *H. pylori* infection. However, no significant increase in their mRNA expression levels was observed (Fig. 4*a*). Further, at the protein level, expression levels of DNMT1, DNMT3A and DNMT3B were below the detection limit even in the gastric mucosae with *H. pylori* (Fig. 4*b*), indicating no increase was induced by *H. pylori* infection.

Discussion

The presence of gene specificity for aberrant DNA methylation induction by a specific carcinogenic factor was demonstrated for the first time in this study. Also, genes susceptible to methylation had significantly lower mRNA expression levels than resistant genes. For clarification of the relationship between a methylation-inducing factor and gene specificity, use of noncancerous gastric tissue, which is polyclonal, was important because gene silencing due to promoter methylation can result in over- or under-presence of methylation in cancer tissues. Methylation in noncancerous tissues is also reported in the colonic mucosae of patients with ulcerative colitis[13,14] and liver tissues of patients with hepatocellular carcinomas,[15] but limited numbers of genes have been analyzed so far.

Methylation of specific genes can persist for a lifetime, and there is a possibility that the methylation profile can be used as a methylation fingerprint of *H. pylori* infection in the past, as specific *p53* and *APC* mutations are used to assess past exposure to specific carcinogens.[16,17] Use of DNA methylation has an advantage over mutations because methylation can be present in a significant fraction of cells in noncancerous tissues, and can be detected sensitively and reproducibly. The noncancerous gastric mucosae of cases with a gastric cancer without current *H. pylori* infection, most of which are considered to have had past exposure to *H. pylori*,[18] showed the same methylation profile as individuals with current *H. pylori* infection. This finding indicated that the methylation profile induced by *H. pylori* infection can persist even after discontinuation of *H. pylori* infection. Although eradication of *H. pylori* was reported to decrease incidences of individuals with methylation,[19,20] the decrease is only partial, not to zero, and highly variable among individuals (manuscript in preparation).

To establish a methylation profile as a fingerprint of *H. pylori* infection, the profile must be specific. Unfortunately, few gastric cancers can be considered as those induced solely by another carcinogenic factor, such as Epstein-Barr virus infection[21] or high salt intake,[22] and the specificity cannot be examined easily. However, since low mRNA expression levels are involved in gene specificity, there is a possibility that different carcinogenic factors induce different methylation profiles through induction of reduced mRNA expression of different genes. Once the specificity of a methylation profile is established, a methylation fingerprint will be very useful for clinicopathological analysis and epidemiology. Among the clinically used tests for *H. pylori* infection, the culture and rapid urease tests can detect only current *H. pylori* infection.[23,24] The serum antibody test remains positive in only half the patients as early as 1 year after successful eradication of *H. pylori*.[25,26]

The role of low mRNA expression in methylation induction has been reported.[4] De Smet *et al.* showed that weak transcriptional capacity leads to promoter remethylation by analysis of demethylation and mRNA expression of *MAGE-A1* in various cell lines.[27] Song *et al.* showed that decreased promoter activity leads to hypermethylation of a promoter CpG island of an exogenously introduced gene by disrupting its promoter activity.[28] We and others previously observed that most genes methylated in cancer tissues had no or little expression in cancer precursor cells.[29–32] This study showed that, in normal cells and *in vivo*, low mRNA expression is important for methylation induction. Also, it was suggested that downregulation by *H. pylori* infection precedes methylation since 8 of the 22 genes with expression analyses were downregulated by *H. pylori* infection, but none were upregulated. The 22 genes were selected from those that can be methylated in gastric cancer cell lines and even the resistant genes are considered to be relatively susceptible among the entire genes.

Even among the genes with similarly low mRNA expression levels, some genes were resistant and others were susceptible to methylation induction by *H. pylori*. As additional factors that determine the gene specificity of methylation induction, histone modification deregulation could be important. For example, a repressive histone modification, methylation at Lys27 of histone H3 (H3K27) induced by Polycomb group proteins, is associated with genes methylated in cancers.[33,34] Active chromatin marks, associated with active mRNA expression, could be important to protect DNA from methylation. At the same time, without *H. pylori* infection, even the susceptible genes were not methylated, indicating that abnormality in epigenetic regulation was induced by *H. pylori* infection. The final step of aberrant methylation must be mediated by DNA methyltransferases, and actually overexpression of *de novo* methyltransferases enhance methylation of specific genes in a mouse model.[35] Also, some inflammatory cytokines, such as IL-6, are reported to induce DNA methyltransferases.[36] However, contrary to initial expectations, *H. pylori* infection did not induce either mRNA or protein expression of DNMT1, DNMT3A and DNMT3B in gastric mucosae. Abnormalities in epigenetic regulation induced by *H. pylori* infection also need to be investigated.

In summary, methylation of specific genes was induced by *H. pylori* infection in noncancerous gastric membranes, and preceding low mRNA expression was suggested to be involved in the specificity. Use of the specific profile as a methylation fingerprint of past exposure to a specific carcinogenic factor was suggested.

References

1. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007;128:683–92.
2. Ushijima T, Sasako M. Focus on gastric cancer. Cancer Cell 2004;5:121–5.
3. Issa JP, Ottaviano YL, Celano P, Hamilton SR, Davidson NE, Baylin SB. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. Nat Genet 1994;7:536–40.
4. Ushijima T, Okochi-Takada E. Aberrant methylations in cancer cells: where do they come from? Cancer Sci 2005;96:206–11.
5. Maekita T, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M, Arii K, Kaneda A, Tsukamoto T, Tatematsu M, Tamura G, Saito D, 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.
6. Nakajima T, Maekita T, Oda I, Gotoda T, Yamamoto S, Umemura S, Ichinose M, Sugimura T, Ushijima T, Saito D. Higher methylation levels in gastric mucosae significantly correlate with higher risk of gastric cancers. Cancer Epidemiol Biomarkers Prev 2006;15:2317–21.
7. Kaneda A, Kaminishi M, Yanagihara K, Sugimura T, Ushijima T. Identification of silencing of nine genes in human gastric cancers. Cancer Res 2002;62:6645–50.
8. Yamashita S, Tsujino Y, Moriguchi K, Tatematsu M, Ushijima T. Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2′-deoxycytidine treatment and oligonucleotide microarray. Cancer Sci 2006;97:64–71.
9. Kaneda A, Kaminishi M, Sugimura T, Ushijima T. Decreased expression of the seven ARP2/3 complex genes in human gastric cancers. Cancer Lett 2004;212:203–210.
10. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007;39:457–66.
11. Li B, Carey M, Workman JL. The role of chromatin during transcription. Cell 2007;128:707–19.
12. Lin JC, Jeong S, Liang G, Takai D, Fatemi M, Tsai YC, Egger G, Gal-Yam EN, Jones PA. Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island. Cancer Cell 2007;12:432–44.
13. Hsieh CJ, Klump B, Holzmann K, Borchard F, Gregor M, Porschen R. Hypermethylation of the p16INK4a promoter in colectomy specimens of patients with long-standing and extensive ulcerative colitis. Cancer Res 1998;58:3942–45.
14. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res 2001;61:3573–77.

15. 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;32:970–9.
16. Hussain SP, Harris CC. Molecular epidemiology and carcinogenesis: endogenous and exogenous carcinogens. Mutat Res 2000;462:311–22.
17. Toyota M, Ushijima T, Kakiuchi H, Canzian F, Watanabe M, Imai K, Sugimura T, Nagao M. Genetic alterations in rat colon tumors induced by heterocyclic amines. Cancer 1996;77:1593–97.
18. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, Taniyama K, Sasaki N, Schlemper RJ. *Helicobacter pylori* infection and the development of gastric cancer. N Engl J Med 2001;345:784–9.
19. Leung WK, Man EP, Yu J, Go MY, To KF, Yamaoka Y, Cheng VY, Ng EK, Sung JJ. Effects of *Helicobacter pylori* eradication on methylation status of E-cadherin gene in noncancerous stomach. Clin Cancer Res 2006;12:3216–21.
20. Chan AO, Chu KM, Huang C, Lam KF, Leung SY, Sun YW, Ko S, Xia HH, Cho CH, Hui WM, Lam SK, Rashid A. Association between *Helicobacter pylori* infection and interleukin 1beta polymorphism predispose to CpG island methylation in gastric cancer. Gut 2007;56:595–7.
21. Takada K. Epstein-Barr virus and gastric carcinoma. Mol Pathol 2000;53:255–61.
22. Tsugane S, Sasazuki S, Kobayashi M, Sasaki S. Salt and salted food intake and subsequent risk of gastric cancer among middle-aged Japanese men and women. Br J Cancer 2004;90:128–34.
23. Kawakami Y, Akahane T, Gotoh A, Okimura Y, Oana K, Katsuyama T. Successful development of air-dried microplates (HP-Plates) for susceptibility testing against *Helicobacter pylori* isolates. Microbiol Immunol 1997;41:703–08.
24. Nishikawa K, Sugiyama T, Kato M, Ishizuka J, Kagaya H, Hokari K, Asaka M. A prospective evaluation of new rapid urease tests before and after eradication treatment of *Helicobacter pylori*, in comparison with histology, culture and 13C-urea breath test. Gastrointest Endosc 2000;51:164–8.
25. Kosunen TU, Seppala K, Sarna S, Sipponen P. Diagnostic value of decreasing IgG, IgA, and IgM antibody titres after eradication of. *Helicobacter pylori*. Lancet 1992;339:893–5.
26. Marchildon P, Balaban DH, Sue M, Charles C, Doobay R, Passaretti N, Peacock J, Marshall BJ, Peura DA. Usefulness of serological IgG antibody determinations for confirming eradication of *Helicobacter pylori* infection. Am J Gastroenterol 1999;94:2105–08.
27. De Smet C, Loriot A, Boon T. Promoter-dependent mechanism leading to selective hypomethylation within the 5′ region of gene MAGE-A1 in tumor cells. Mol Cell Biol 2004;24:4781–90.
28. Song JZ, Stirzaker C, Harrison J, Melki JR, Clark SJ. Hypermethylation trigger of the glutathione-S-transferase gene (GSTP1) in prostate cancer cells. Oncogene 2002;21:1048–61.
29. Hagihara A, Miyamoto K, Furuta J, Hiraoka N, Wakazono K, Seki S, Fukushima S, Tsao MS, Sugimura T, Ushijima T. Identification of 27 5′ CpG islands aberrantly methylated and 13 genes silenced in human pancreatic cancers. Oncogene 2004;23:8705–10.
30. Furuta J, Nobeyama Y, Umebayashi Y, Otsuka F, Kikuchi K, Ushijima T. Silencing of Peroxiredoxin 2 and aberrant methylation of 33 CpG islands in putative promoter regions in human malignant melanomas. Cancer Res 2006;66:6080–86.
31. Ushijima T. Detection and interpretation of altered methylation patterns in cancer cells. Nat Rev Cancer 2005;5:223–31.
32. Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E, Pikarski E, Young RA, Niveleau A, Cedar H, Simon I. Evidence for an instructive mechanism of de novo methylation in cancer cells. Nat Genet 2006;38:149–53.
33. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W, Berman DM, Jenuwein T, et al. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 2007;39:237–42.
34. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, Laird PW. Epigenetic stem cell signature in cancer. Nat Genet 2007; 39:157–8.
35. Linhart HG, Lin H, Yamada Y, Moran E, Steine EJ, Gokhale S, Lo G, Cantu E, Ehrich M, He T, Meissner A, Jaenisch R. Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. Genes Dev 2007;21:3110–22.
36. Hodge DR, Xiao W, Clausen PA, Heidecker G, Szyf M, Farrar WL. Interleukin-6 regulation of the human DNA methyltransferase (HDNMT) gene in human erythroleukemia cells. J Biol Chem 2001; 276:39508–11.

The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands

Hideyuki Takeshima, Satoshi Yamashita, Taichi Shimazu, et al.

Genome Res. 2009 19: 1974-1982 originally published online August 3, 2009
Access the most recent version at doi:10.1101/gr.093310.109

Letter

The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands

Hideyuki Takeshima,[1] Satoshi Yamashita,[1] Taichi Shimazu,[2] Tohru Niwa,[1] and Toshikazu Ushijima[1,3]

[1]*Carcinogenesis Division, National Cancer Center Research Institute, 104-0045 Tokyo, Japan;* [2]*Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, 104-0045 Tokyo, Japan*

Instructive mechanisms are present for induction of DNA methylation, as shown by methylation of specific CpG islands (CGIs) by specific inducers and in specific cancers. However, instructive factors involved are poorly understood, except for involvement of low transcription and trimethylation of histone H3 lysine 27 (H3K27me3). Here, we used methylated DNA immunoprecipitation (MeDIP) combined with a CGI oligonucleotide microarray analysis, and identified 5510 and 521 genes with promoter CGIs resistant and susceptible, respectively, to DNA methylation in prostate cancer cell lines. Expression analysis revealed that the susceptible genes had low transcription in a normal prostatic epithelial cell line. Chromatin immunoprecipitation with microarray hybridization (CHiP-chip) analysis of RNA polymerase II (Pol II) and histone modifications showed that, even among the genes with low transcription, the presence of Pol II was associated with marked resistance to DNA methylation (OR = 0.22; 95% CI = 0.12–0.38), and H3K27me3 was associated with increased susceptibility (OR = 11.20; 95% CI = 7.14–17.55). The same was true in normal human mammary epithelial cells for 5430 and 733 genes resistant and susceptible, respectively, to DNA methylation in breast cancer cell lines. These results showed that the presence of Pol II, active or stalled, and H3K27me3 can predict the epigenetic fate of promoter CGIs independently of transcription levels.

[Supplemental material is available online at http://www.genome.org. The microarray data from this study have been submitted to Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo) under accession no. GSE15154.]

Epigenetic alterations, along with genetic alterations, are known to play critical roles in human carcinogenesis and other acquired diseases (Laird and Jaenisch 1996; Robertson 2005; Jones and Baylin 2007). Especially, DNA methylation of promoter CpG islands (CGIs) has been known to be involved in silencing of tumor-suppressor and other genes (Ushijima 2005; Eckhardt et al. 2006; Jones and Baylin 2007). In addition, a critical role of methylation of the nucleosome-free region (NFR) just upstream of a transcription start site (TSS) was recently demonstrated in nucleosome occupation and thus in gene silencing (Li et al. 2007; Lin et al. 2007).

Epigenetic alterations, different from genetic alterations, have unique natures, such as gene specificity (Costello et al. 2000; Esteller et al. 2001; Keshet et al. 2006; Nakajima et al. 2009; Oka et al. 2009), high levels of accumulation in normal-appearing tissues (Kondo et al. 2000; Maekita et al. 2006; Ushijima 2007), and deep involvement of inflammation in their induction (Issa et al. 2001; Ushijima and Okochi-Takada 2005; Maekita et al. 2006). Especially, the presence of gene specificity, originally suggested by the presence of tumor type-specific DNA methylation patterns (Costello et al. 2000; Esteller et al. 2001), is now confirmed by methylation of specific genes in non-cancerous tissues exposed to specific carcinogenic factors (Nakajima et al. 2009; Oka et al. 2009). Selection biases for genes with growth advantage can be avoided by analysis of non-cancerous, therefore polyclonal, tissues (Mihara et al. 2006). The gene specificity of DNA methylation induction depending on cell types and carcinogenic factors shows that there are instructive mechanisms for DNA methylation induction, in contrast to the random nature of mutation induction.

As mechanisms for instructive induction, limited information is available so far, including low transcription levels and some histone modifications. Exogenous and endogenous genes are likely to become methylated only when they have low transcription levels (Song et al. 2002; De Smet et al. 2004). Most genes methylated in cancer tissues had no or low transcription in their normal counterpart cells (Ushijima 2005; Keshet et al. 2006). Transcription factors, such as SP1/SP3 and MLL, protected CpG sites from becoming methylated, independent of and dependent on transcription levels, respectively (Boumber et al. 2008; Erfurth et al. 2008). In addition, trimethylation of histone H3 lysine 27 (H3K27me3), a target of Polycomb repressive complex (PRC) 2 (Hansen et al. 2008), was enriched in normal cells and embryonic stem (ES) cells at genes that can be methylated in cancers (Ohm et al. 2007; Schlesinger et al. 2007; Widschwendter et al. 2007; Hahn et al. 2008; Rodriguez et al. 2008). Nevertheless, at a genome level, many genes have low transcription levels and H3K27me3 but are still resistant to DNA methylation induction, indicating that some critical factors are likely to be still missing.

In this study, we hypothesized that RNA polymerase II (Pol II) binding around TSSs can function as a protective factor for DNA methylation induction. Accumulation of Pol II at genes with low transcription levels (stalled Pol II) was recently found in as high as 12% of protein-coding genes in *Drosophila melanogaster* (Muse et al. 2007; Zeitlinger et al. 2007) and in humans (Guenther et al. 2007). We demonstrate in a genome-wide manner that Pol II binding, active or stalled, and histone modifications in normal cells predict genes resistant and susceptible to DNA methylation in cancers.

Results

Identification of genes with promoter CGIs resistant and susceptible to DNA methylation

To identify genes with promoter CGIs resistant and susceptible to induction of DNA methylation in human prostate cancers, four

[3]**Corresponding author.**
E-mail tushijim@ncc.go.jp; fax 81-3-5565-1753.
Article published online before print. Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.093310.109.

19:1974–1982 © 2009 by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/09; www.genome.org

prostate cancer cell lines (PC3, LNCaP, 22Rv1, and Du145), along with a normal prostatic epithelial cell line (RWPE1), were analyzed using methylated DNA immunoprecipitation (MeDIP) combined with a human CGI oligonucleotide microarray that covered 27,800 CGIs (MeDIP-CGI microarray analysis).

First, appropriate cutoff values of our original output values "DNA methylation values" (Me values) were determined using 145 samples (29 CGIs in five cell lines) (Supplemental Table S1). As cutoff values with high specificity and little compromise of sensitivity, cutoff values of 0.6 and 0.4 were selected for methylated and unmethylated CGIs, respectively (Supplemental Fig. S1). The specificity and sensitivity for methylated (unmethylated) CGIs with these values were 0.95 (0.96) and 0.85 (0.82), respectively. DNA methylation status of a CGI or putative NFR was judged as unmethylated (UM), moderately methylated (MM), and highly methylated (HM) when the average of Me values of the probes within the region was 0–0.4, 0.4–0.6, and 0.6–1.0, respectively. The validity of our methods was also supported by the fact that promoter CGIs were more likely to be unmethylated (68%–82%) than those in gene bodies (54%–63%), which conformed with previous observations (Supplemental Table S2; Ushijima et al. 2003; Eckhardt et al. 2006; Rakyan et al. 2008).

The susceptibility of genes was determined by methylation analysis of 8930 NFRs (Li et al. 2007). Genes with NFRs unmethylated (Me value, 0–0.4) in the normal cell line and all the four cancer cell lines were defined as DNA methylation-resistant genes. On the other hand, those unmethylated in the normal cell line but highly methylated (Me value, 0.6–1.0) in at least one of the four cancer cell lines were defined as DNA methylation-susceptible genes (Fig. 1A). Susceptible genes were further divided into S1, S2, S3, and S4 subclasses according to the DNA methylation frequency in cancer cell lines (highly methylated in one, two, three, and four, respectively, of the four cancer cell lines). In addition, genes unmethylated in the normal cell line but moderately methylated (Me value, 0.4–0.6) in at least one of the four cancer cell lines were defined as genes with intermediate susceptibility (intermediate genes). In prostate cancers, 5510, 1330, and 521 genes with promoter CGIs were classified as resistant, intermediate, and susceptible genes, respectively (Fig. 1B). DNA methylation levels of NFRs were largely consistent with those of further upstream regions up to −800 bp, and downstream regions up to +800 bp (Fig. 1C).

To avoid any tissue bias and statistical errors, we also analyzed three human breast cancer cell lines (MCF7, ZR-75-1, and MDA-MB-468), along with normal human mammary epithelial cells (HMEC). As in the prostate, the promoter CGIs were more likely to be unmethylated (68%–90%) than the CGIs located in gene bodies (52%–70%) (Supplemental Table S2). Using the same definition as in the prostate cancers, 5430, 1913, and 733 genes with promoter CGIs were classified as resistant, intermediate, and susceptible genes, respectively (Fig. 1B). As in prostate cancers, DNA methylation levels were also largely consistent among the NFRs, further upstream regions, and downstream regions in human breast cancers (Supplemental Fig. S2). Between breast and prostate cancers, only 261 genes, 36% of the susceptible genes in breast cancers and 50% of those in prostate cancers, were commonly susceptible, showing the presence of tissue specificity.

To explore possible selection bias for the resistant and susceptible genes due to gene functions, functional annotation analysis of resistant and susceptible genes was performed. In the prostate, 203 and 154 processes out of 16,621 biological processes were enriched among the resistant and susceptible genes, respectively. Among the resistant genes, processes involved in basic cellular processes such as metabolic process, RNA processing, and RNA splicing were enriched. In contrast, among the susceptible genes, biological processes involved in the developmental processes of

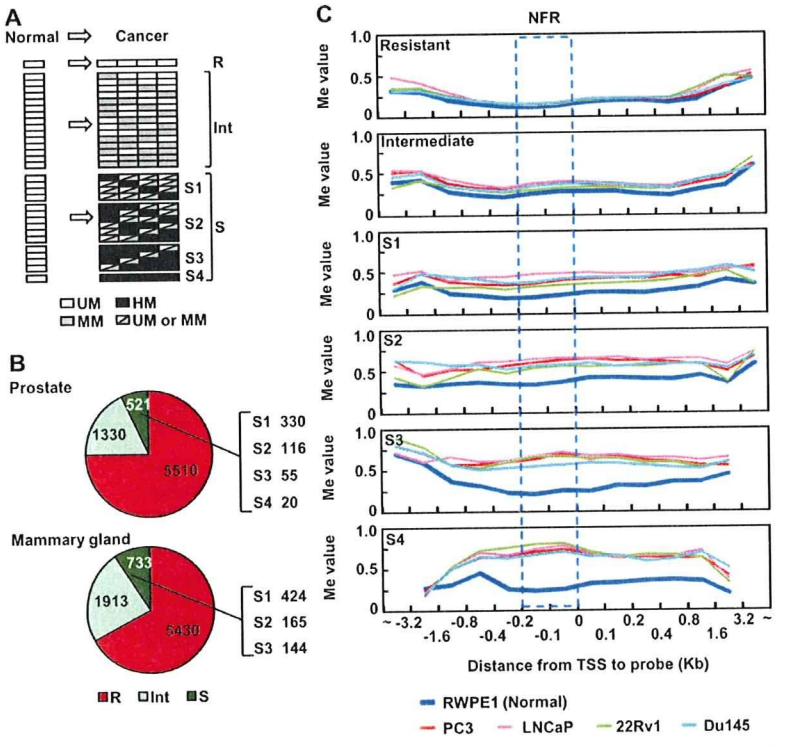

Figure 1. Identification of methylation-resistant and methylation-susceptible genes and their methylation profiles in various genomic regions against TSSs. (*A*) Definition of genes resistant and susceptible to induction of DNA methylation. Genes unmethylated (UM) (white) in the normal cell line (cells) and all cancer cell lines were defined as resistant genes (R). Genes unmethylated in the normal cell line (cells) but highly methylated (HM) (black) in at least one of the four cancer cell lines were defined as susceptible genes (*S*). Susceptible genes were further divided into four subclasses according to DNA methylation frequency in the cancer cell lines (S1–S4). Genes unmethylated in the normal cell line (cells) but moderately methylated (MM) (gray) in the cancer cell lines were defined as genes with intermediate susceptibility (intermediate genes: Int). (*B*) The fractions of resistant (red), intermediate (light green), and susceptible (green) genes in the prostate and the mammary gland. (*Right* side of the pie graph) Numbers of susceptible genes in each subclass (S1–S4). (*C*) DNA methylation levels at various positions against the TSSs in the normal prostatic cell line and four cancer cell lines. Average Me values of CGIs continuous from their NFRs are shown. (Blue dotted rectangle) The NFRs. Methylation levels of the NFRs were similar to those of upstream regions up to −800 bp and downstream regions up to +800 bp.

specific cells or tissues, such as nervous system development, and embryonic development, were enriched (Table 1). Similar enrichment of genes involved in specific biological processes was also observed in the mammary glands.

Low transcription levels of DNA methylation-susceptible genes in normal cell lines

For a limited number of genes, the susceptibility of genes with low transcription to DNA methylation has been reported in cell lines (Song et al. 2002; De Smet et al. 2004) and in human tissue (Ushijima and Okochi-Takada 2005; Nakajima et al. 2009). To analyze this susceptibility in a genome-wide manner, we performed expression analysis in the normal prostatic cell line using a GeneChip oligonucleotide microarray. Owing to the difference of array platforms between the CGI oligonucleotide microarray and the GeneChip oligonucleotide expression microarray, we were able to measure transcription levels of the 7574 genes out of 8930 genes with promoter CGIs in the normal prostatic cell line. The accuracy of the transcription levels obtained by the GeneChip oligonucleotide microarray was validated by observing a strong correlation between the microarray data and mRNA levels obtained by quantitative RT-PCR (correlation coefficient = 0.95 and 0.97 in RWPE1 and HMEC, respectively) (Supplemental Fig. S3). When the transcription levels were analyzed according to the DNA methylation status in the normal prostatic cell line itself, as expected, highly methylated genes had remarkably low transcription levels (Supplemental Fig. S4).

Genes highly methylated in prostate cancer cell lines had low transcription levels in the normal prostatic cell line (Fig. 2A). When transcription levels of resistant, intermediate, and susceptible genes were compared, susceptible genes had lower transcription levels than resistant genes. Even among the susceptible genes, genes with frequent DNA methylation had lower transcription levels than those with infrequent DNA methylation (Fig. 2B). When fractions of genes with high, moderate, and low transcription levels were analyzed in the 7574 total, 4567 resistant, and 479 susceptible genes, the susceptible genes had a significantly larger fraction of genes with low transcription (63%) than the total genes (38%; $P < 0.001$, χ^2 test) (Fig. 2C). Even among the susceptible genes, genes with more frequent DNA methylation had the larger fraction of genes with low transcription (Supplemental Fig. S5). These results showed that aberrant DNA methylation is preferentially induced in genes with low transcription, as previously reported (Song et al. 2002; De Smet et al. 2004; Ushijima 2005; Keshet et al. 2006; Nakajima et al. 2009), in a genome-wide manner.

In the mammary glands, the susceptible genes also had a significantly larger fraction of genes with low transcription (74%) than the total genes (37%; $P < 0.001$, χ^2 test) (Supplemental Figs. S5, S6).

Levels of histone modifications and Pol II binding were associated with DNA methylation susceptibility

Although most genes susceptible to DNA methylation in cancers had low transcription in the normal cell line (cells), the converse was not true: 1237 of 2852 (prostate) and 1048 of 2750 (breast) genes with low transcription in the normal cell line (cells) were still resistant to DNA methylation in cancers (Fig. 2C; Supplemental Fig. S6). This indicated that factors besides low transcription are also involved in DNA methylation susceptibility. To address this issue, we analyzed both active (acetylation of histone H3 [H3Ac] and trimethylation of histone H3 lysine 4 [H3K4me3]) and inactive (trimethylation of histone H3 lysine 9 [H3K9me3] and H3K27me3) histone modifications and Pol II binding at and adjacent to the NFRs in a genome-wide manner. Since the length of sheared DNA used for chromatin immunoprecipitation (ChIP) analysis ranged mainly from 200 to 1000 bp, analysis of probes within the NFRs automatically reflected histone modifications adjacent to the NFRs even if nucleosomes were absent in the NFRs.

Table 1. Functional annotation analysis of genes with different DNA methylation susceptibility

	Prostate		Mammary gland	
Category	Term	*P*-value	Term	*P*-value
Resistant	Primary metabolic process	3.72E-22	Cellular metabolic process	3.16E-20
	Macromolecule metabolic process	6.27E-22	Metabolic process	1.74E-19
	Cellular metabolic process	1.08E-21	Primary metabolic process	4.50E-19
	Metabolic process	5.98E-21	Macromolecule metabolic process	2.33E-17
	Biopolymer metabolic process	2.92E-15	RNA processing	9.75E-17
	RNA processing	1.34E-14	RNA splicing	6.54E-14
	Nucleobase, nucleoside, nucleotide, and nucleic acid metabolic process	1.49E-14	Nucleobase, nucleoside, nucleotide, and nucleic acid metabolic process	1.30E-13
	mRNA metabolic process	9.65E-14	Macromolecule localization	1.51E-13
	RNA splicing	3.94E-13	mRNA metabolic process	1.29E-12
	Protein transport	7.19E-13	Biopolymer metabolic process	1.29E-12
Susceptible	Multicellular organismal process	2.41E-16	Multicellular organismal process	8.20E-30
	Multicellular organismal development	3.33E-12	Multicellular organismal development	4.77E-23
	System development	4.77E-11	System development	2.57E-18
	Anatomical structure development	1.10E-09	Anatomical structure development	6.61E-17
	System process	3.09E-08	Developmental process	5.18E-16
	Nervous system development	9.59E-08	Nervous system development	3.96E-13
	Developmental process	4.48E-07	Cell-cell signaling	2.27E-12
	Organ development	7.93E-07	Organ development	2.85E-12
	Cell-cell signaling	3.03E-06	Embryonic development	6.96E-11
	Biological adhesion	9.69E-06	System process	1.01E-10

Enrichment of specific biological processes in Gene Ontology criteria among resistant and susceptible genes was analyzed by DAVID bioinformatics resources. The top 10 significantly enriched biological processes in each gene category are listed. The significance (*P*-value) of enrichment is shown.

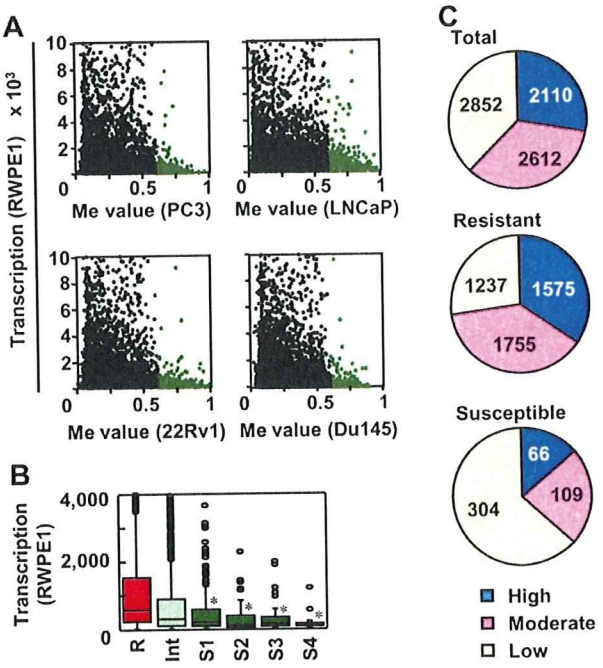

Figure 2. Low transcription levels of DNA methylation-susceptible genes in the normal prostatic cell line (RWPE1). (*A*) The association between DNA methylation levels (Me value of the NFRs) in each of the four prostate cancer cell lines (PC3, LNCaP, 22Rv1, and Du145) and transcription levels in RWPE1. (Green dots) Genes highly methylated in a cancer cell line. Genes highly methylated in a cancer cell line had low transcription levels in the normal cell line. (*B*) Transcription levels of resistant (R), intermediate (Int), and susceptible (S1–S4) genes in RWPE1. The boxes represent the 75th and 25th percentiles, and the line in the box represents the 50th percentile (the median). Whiskers represent the maximum data within (75th percentile + 1.5 × [75th percentile − 25th percentile]) and the minimum data within (25th percentile − 1.5 × [75th percentile − 25th percentile]). (Dots) The data not included between the whiskers. Transcription levels of Int, S1, S2, S3, and S4 were compared to that of R by the Mann-Whitney *U*-test (*$P < 1 \times 10^{-5}$). Susceptible genes had significantly lower expression levels than resistant genes. (*C*) The fraction of genes with high (blue) (signal intensity > 1000), moderate (pink) (250–1000), and low (yellow) (<250) transcription. Susceptible genes had a significantly larger fraction of genes with low transcription than the total genes.

The data obtained by the ChIP with microarray hybridization (ChIP-chip) analysis were validated by analyzing correlations between the signal ratio (immunoprecipitated DNA [IP]/whole cell extract [WCE]) obtained by ChIP-chip and those obtained by quantitative ChIP-PCR (Supplemental Fig. S7).

Using only genes with low transcription, we analyzed the association between the candidate instructive factors in the normal prostatic cell line and susceptibility to DNA methylation in prostate cancer cell lines. It was clear that H3Ac and H3K4me3 were elevated in resistant genes, and H3K27me3 was elevated in susceptible genes (Fig. 3A). In contrast, the H3K9me3 level was not different between resistant and susceptible genes. Notably, Pol II binding was remarkably higher in resistant genes (Fig. 3B). When further upstream regions and downstream regions were analyzed, resistant genes had elevated H3Ac and H3K4me3 mainly in their downstream regions, and susceptible genes had elevated H3K27me3 in their downstream regions and further upstream regions (Fig. 3C). Pol II binding was elevated mainly in the NFRs and then in downstream regions of resistant genes (Fig. 3C). In the mammary glands, exactly the same tendency was observed (Supplemental Fig. S8).

Next, within the normal prostatic cell line, the association between histone modifications and transcription levels was analyzed. Conforming to previous reports (Barski et al. 2007; Wang et al. 2008), genes with high and low transcription had elevated active and inactive histone modifications (Supplemental Fig. S9). Notably, among genes with low transcription, those without DNA methylation had elevated H3K27me3, confirming a previous report that H3K27me3 is involved in gene silencing independent of DNA methylation (Kondo et al. 2008). Within the normal mammary epithelial cells, the same tendency was observed.

Strongest association of Pol II binding with resistance to DNA methylation

The combination effect of H3K27me3 and one of the three active factors (H3Ac, H3K4me3, and Pol II binding) on DNA methylation susceptibility was then examined (Fig. 3D). All the three combinations were informative in distinguishing the resistant and susceptible genes, while Pol II binding gave the clearest discrimination. Multivariate logistic regression analysis was then performed to compare precisely the independent effects of H3Ac, H3K4me3, H3K9me3, H3K27me3, and Pol II binding on DNA methylation susceptibility. The genes with low transcription in the normal cell line (cells) were divided into quintiles according to the amounts of H3Ac, H3K4me3, H3K9me3, H3K27me3, and Pol II binding at the NFRs. Compared with the genes in the lowest quintile, multivariate-adjusted odds ratios (ORs) of genes in the other quintiles to become moderately or highly methylated in cancers (Int, and S1–S4 for the prostates; Int, and S1–S3 for the mammary glands) were calculated (Table 2). In the prostates, Pol II binding had the strongest independent association with resistance, and H3K27me3 had a strong and significant association with susceptibility. In the mammary glands, similar associations were observed. If the analysis was performed for the multivariate-adjusted odds ratio of genes to become highly methylated (S1–S4 for the prostates; and S1–S3 for the mammary glands), the association of Pol II binding became even clearer (Supplemental Table S3).

Finally, regardless of their transcription levels, all the genes were classified into genes with "active Pol II" (high/moderate transcription, high Pol II), those with "stalled Pol II" (low transcription, high Pol II), and those with "low Pol II" (low Pol II). The group of genes with low Pol II was further subdivided into those with and without H3K27me3. In the normal prostatic cell line, 47%, 13%, and 40% of genes had active, stalled, and low Pol II, respectively (Fig. 4A). Both genes with active Pol II and genes with stalled Pol II consisted mostly of resistant genes (Fig. 4B). In contrast, genes with low Pol II contained larger fractions of susceptible genes, and the presence of H3K27me3 remarkably increased the fraction. Similar results were obtained also in the mammary glands (Supplemental Fig. S10).

Discussion

In this study, we showed that Pol II binding in the NFRs in normal cell lines (cells) was closely associated with resistance to DNA methylation in cancer cell lines (cells) for the first time. The association between Pol II binding and resistance to DNA methylation was independent of transcriptional levels. It was also independent from the promoting effect of H3K27me3, and the combination of Pol II binding and H3K27me3 could explain a large part of the

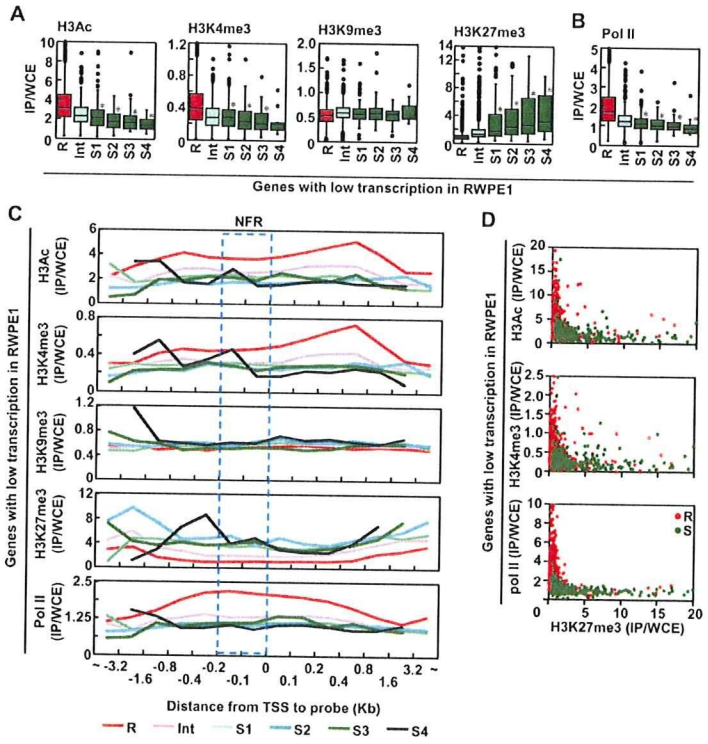

Figure 3. The association between the levels of candidate instructive factors in RWPE1 and DNA methylation susceptibility, among genes with low transcription in RWPE1. (*A*) Histone modification levels of genes with different susceptibilities to DNA methylation. For the box plot and statistical methods, refer to the legend to Figure 2B. Active histone modifications were elevated in resistant genes, and H3K27me3 was elevated in susceptible genes. (*B*) The association between Pol II binding and DNA methylation susceptibility. Pol II binding was associated with resistance even among genes with low transcription. (*C*) Levels of histone modifications and Pol II binding at various positions against the TSSs in RWPE1. Average levels of histone modifications and Pol II binding of CGIs continuous from their NFRs are shown. (Blue dotted rectangle) The NFRs. (*D*) The combination effect of one of the three active factors (H3Ac, H3K4me3, and Pol II binding) (*y*-axis) and H3K27me3 (*x*-axis) on resistance and susceptibility of genes with low transcription. (Red dots) DNA methylation-resistant genes; (green dots) DNA methylation-susceptible genes; they were separated by any of the three combinations.

instructive mechanisms for induction of DNA methylation. These data provided fundamental information on how the epigenetic fate of promoter CGIs is determined. The association between Pol II binding and resistance to DNA methylation can be potentially useful in the prediction of genes that will become silenced in cancer and other diseases.

Our multivariate analysis involving Pol II binding and histone modifications showed that the association between active histone modifications and resistance to DNA methylation was mostly overridden by that of Pol II binding, while the association between H3K27me3 and susceptibility to DNA methylation remained. It was reported that active histone modifications are involved in anchoring of the basal transcription factor TFIID (Vermeulen et al. 2007), which forms a transcription complex with Pol II. H3K4me3 is recognized by the PHD domain of TFIID, and acetylation of histone H3 lysine 9 and lysine 14 potentiates this interaction. It was therefore suggested that Pol II binding more directly works as a protection mechanism than active histone modifications, and that H3K27me3 has an independent mode of action.

Pol II forms a huge transcription complex of ~3 MDa with general transcription factors and other proteins (Boeger et al. 2005), and such a huge complex around promoter CGIs is expected to compete with DNA methyltransferases and their associated proteins. On the other hand, H3K27me3 is recognized by PRC2/3 (Hansen et al. 2008), which contains EZH2. Since EZH2 interacts with DNMT3A and DNMT3B (Vire et al. 2006), H3K27me3 is expected to signal binding of DNMT3A and DNMT3B. Taken together, Pol II binding and H3K27me3 are likely to function by preventing and promoting, respectively, recruitment of DNA methylation complexes.

Functional annotation analysis revealed that most of the susceptible genes were involved in the developmental processes of specific cells or tissues. Genes in this category were considered unnecessary for normal cells that have already differentiated. This raised alternative possibilities: The lack of current need for a gene is one of the instructive factors, or an unnecessary gene has a low level of Pol II, which is associated with methylation susceptibility. To distinguish these two possibilities, we examined overrepresentation of susceptible genes among genes with low Pol II levels after classification of genes by their function (Supplemental Table S4). As a result, in any categories of genes, susceptible genes were overrepresented among the genes with low Pol II levels, showing that the presence of Pol II was an independent factor for resistance to DNA methylation from functions of genes.

Specific genome structures are also known to be involved in the specificity of genes methylated, in addition to the instructive factors analyzed here. The presence of a repetitive sequence has been reported to be capable of functioning as a source of aberrant DNA methylation (Yates et al. 1999). In addition to methylation induction of individual genes, a cluster of genes can be methylated simultaneously in a cancer (Frigola et al. 2006). In this study, 64% and 50% of the susceptible genes in breast and prostate cancers, respectively, were unique to individual tumors. The susceptibility specific to a tissue is more likely to be due to Pol II binding and H3K27me3 rather, while susceptibility common to different tissues can be due to specific genome structures.

Genes moderately methylated were considered to be methylated in a fraction of cancer cells and thus to have been methylated after clonal expansion started. Genes highly methylated were considered to be present in all the cancer cells, and thus to have been methylated before clonal expansion. Therefore, DNA methylation susceptibility in normal cell line (cells) might be more precisely measured using genes highly methylated (Supplemental Table S3) than using genes highly and moderately methylated (Table 2).

As materials, we used normal and cancer cell lines to perform efficient and precise ChIP experiments. It is known that cancer cell

Table 2. The association between the levels of candidate instructive factors and susceptibility to DNA methylation (Int and S)

	Lowest quintile	2nd quintile	3rd quintile	4th quintile	Highest quintile
Prostate					
H3Ac	1	0.78 (0.54–1.12)	0.86 (0.56–1.31)	0.86 (0.54–1.37)	0.91 (0.53–1.57)
H3K4me3	1	0.99 (0.70–1.41)	1.09 (0.74–1.61)	0.92 (0.61–1.38)	0.52 (0.34–0.82)
Pol II	1	0.83 (0.58–1.18)	0.78 (0.52–1.17)	0.40 (0.25–0.62)	0.22 (0.12–0.38)
H3K9me3	1	1.47 (1.00–2.15)	1.26 (0.85–1.86)	1.22 (0.82–1.80)	1.20 (0.81–1.78)
H3K27me3	1	1.41 (0.92–2.17)	2.88 (1.89–4.40)	5.95 (3.87–9.13)	11.20 (7.14–17.55)
Mammary gland					
H3Ac	1	0.95 (0.68–1.35)	0.63 (0.43–0.91)	0.44 (0.30–0.66)	0.42 (0.26–0.67)
H3K4me3	1	0.96 (0.68–1.34)	1.02 (0.71–1.47)	0.59 (0.40–0.87)	0.49 (0.31–0.75)
Pol II	1	1.22 (0.88–1.71)	1.29 (0.90–1.86)	1.14 (0.77–1.68)	0.67 (0.43–1.04)
H3K9me3	1	1.03 (0.76–1.41)	1.07 (0.78–1.47)	1.43 (1.03–1.99)	0.89 (0.64–1.25)
H3K27me3	1	1.61 (1.20–2.18)	2.44 (1.78–3.34)	3.96 (2.86–5.48)	6.44 (4.56–9.10)

Multivariate-adjusted odds ratio (OR) (95% confidence interval; 95% CI) to become methylated (Int, and S1–S4 for the prostates; and Int, and S1–S3 for the mammary glands) is shown for each group. The multivariate-adjusted OR (95% CI) was derived from analyses in which all other listed variables were included into the mode.

lines generally show a larger number of methylated genes than primary tumor cells when a single cancer cell line and a primary tumor sample are compared. However, when a large number of primary tumor samples are analyzed, most DNA methylation found in cancer cell lines is also observed in at least one of the primary tumor samples (Sato et al. 2003; Lodygin et al. 2005; Yamashita et al. 2006). Therefore, it is considered that DNA methylation susceptibility identified in cancer cell lines reflects that in the primary cancer cells as a whole.

In summary, Pol II binding and H3K27me3 in normal cell lines (cells) could predict the epigenetic fate of genes with promoter CGIs in cancer cell lines independently of transcription activity and are major components of instructive mechanisms of DNA methylation induction.

Methods

Cell culture

PC3, LNCaP, 22Rv1, Du145, MCF7, ZR-75-1, and MDA-MB468 (American Type Culture Collection) were maintained in RPMI1640. RWPE1 (American Type Culture Collection) was maintained in keratinocyte-SFM containing 5 ng/mL rEGF, 50 μg/mL bovine pituitary extract (Invitrogen). HMEC (Clonetics) was maintained in mammary epithelial cell serum-free growth medium containing 1% growth supplement (CELL Applications).

ChIP assay

About 1×10^7 cells were cross-linked with 1% formaldehyde for 10 min at room temperature, and washed with ice cold 1× PBS (−) twice. Cells were re-suspended in lysis buffer (50 mM Tris-HCl at pH 8.0, 1 mM EDTA, 1% [w/v] SDS), incubated for 10 min on ice, and then sonicated to shear DNA to an average length ranging from 200 to 1000 bp with a Bioruptor UCD-250 (Cosmo Bio). After DNA shearing, the lysate was centrifuged at 13,000 rpm for 10 min, and supernatant was recovered. The volume of supernatant containing 30 μg of sheared DNA was adjusted to 100 μL with lysis buffer, and then was diluted with 900 μL of dilution buffer (50 mM Tris-HCl at pH 8.0, 167 mM NaCl, 1.1% [w/v] Triton X-100, 0.11% [w/v] sodium deoxycholate [DOC]). Twenty microliters of sheared chromatin was recovered and was used as input DNA.

Diluted lysate was incubated with 2 μg of antibody against H3K4me3 (07-473; Millipore), H3K9me3 (07-442; Millipore), H3K27me3 (07-449; Millipore), H3Ac (06-599; Millipore), or Pol II (ab5095; Abcam), which was reported to be capable of detecting stalled Pol II (Muse et al. 2007) overnight at 4°C with rotation, and then immuno-complexes were collected with 25 μL of Dynabeads Protein A (Invitrogen Dynal AS). Collected beads were washed with 1× RIPA buffer (50 mM Tris-HCl at pH 8.0, 150 mM NaCl, 1 mM EDTA, 1% [w/v] Triton X-100, 0.1% [w/v] SDS, 0.1% [w/v] DOC) containing 150 mM NaCl twice, 1× RIPA buffer containing 500 mM NaCl twice, LiCl wash buffer (10 mM Tris-HCl at pH 8.0, 0.25 M LiCl, 1 mM EDTA, 0.5% [w/v] NP-40, 0.5% [w/v] DOC), and 1× TE containing 50 mM NaCl. Beads were re-suspended with 1× TE, and the cross-links were reversed in the presence of 200 mM NaCl overnight at 65°C. DNA was recovered with RNase A and proteinase K treatment, followed by phenol extraction and ethanol precipitation, and dissolved in 100 μL of 1× TE. One microliter of DNA was used for quantitative ChIP-PCR to confirm the specificity of our

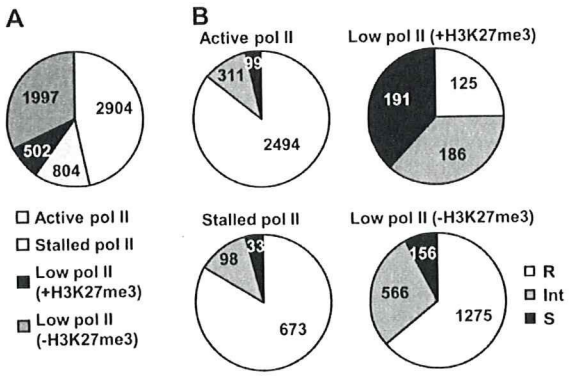

Figure 4. The association between Pol II binding and DNA methylation resistance in the total 6207 genes, regardless of transcription levels. (*A*) Classification of genes by Pol II status and H3K27me3 in the normal prostatic cell line. We were able to analyze transcription levels for 4567 of 5510 resistant, 1161 of 1330 intermediate, and 479 of 521 susceptible genes (total 6207 of 7361 genes) due to a difference in microarray platforms. Genes with high Pol II levels and high/moderate transcription levels were considered as those with "active Pol II." Genes with high Pol II levels but low transcription levels were considered as those with "stalled Pol II." Genes with low Pol II were further subdivided into those with and without H3K27me3. The numbers of genes with active, stalled, and low Pol II are shown. (*B*) The fractions of resistant, intermediate, and susceptible genes according to the Pol II and H3K27me3 statuses. Genes with either active or stalled Pol II had a larger fraction of resistant genes, and genes with low Pol II had a larger fraction of susceptible and intermediate genes.

ChIP technique (Supplemental Fig. S11) or to validate microarray results (Supplemental Fig. S7). Quantitative ChIP-PCR was performed using SYBR Green I (BioWhittaker Molecular Applications) and an iCycler Thermal Cycler (Bio-Rad Laboratories) as described previously (Nakajima et al. 2009). The primers used in quantitative ChIP-PCR are listed in Supplemental Table S5 (Kirmizis et al. 2004).

MeDIP

Five micrograms of genomic DNA was sheared by sonication using a VP-5s homogenizer (TAITEC) to a length of ~300 bp (Supplemental Fig. S12). Generally, there are nine to 53 CpG sites in 300-bp regions of promoter CGI (Nakajima et al. 2009), and this number of CpG sites is sufficient for efficient immunoprecipitation by MeDIP (Keshet et al. 2006). After heat denaturation for 10 min at 95°C, DNA was incubated with 5 μg of antibody against 5-methyl cytidine (Diagnode) in 1× IP buffer (10 mM Na-phosphate at pH 7.0, 140 mM NaCl, 0.05% [w/v] Triton X-100) overnight at 4°C with rotation. Immuno-complexes were collected with 70 μL of Dynabeads Protein A, washed with 1× IP buffer four times, and were recovered by Proteinase K treatment, followed by phenol extraction and ethanol precipitation. DNA was dissolved in 26 μL of 1× TE.

CGI oligonucleotide microarray analysis

Genome-wide analysis of DNA methylation, histone modifications, and Pol II binding was carried out using a human CGI oligonucleotide microarray (Agilent technologies) that contained 237,220 probes in or within 95 bp of CGI covering 27,800 CGIs, with an average probe spacing of 100 bp.

For MeDIP-CGI microarray analysis, immunoprecipitated DNAs from 4.33 μg of sonicated DNA and 0.96 μg of input DNA, without any amplification, were labeled with Cy5 and Cy3, respectively, using an Agilent Genomic DNA Labeling kit PLUS (Agilent technologies). Labeled DNA was hybridized to the microarray for 40 h at 67°C with constant rotation (20 rpm), and then scanned with an Agilent G2565BA microarray scanner (Agilent Technologies). The scanned data were processed using Feature Extraction Ver.9.1 (Agilent Technologies), and the IP (Cy5) and WCE (Cy3) signal values were obtained. These two values were normalized using background subtraction, and signal log ratio [$\log_2$(IP/WCE)] and *P*[Xbar] were obtained using Agilent G4477AA ChIP Analytics 1.3 software (Agilent Technologies). Xbar is a signal value for a probe that takes account of signals for neighboring probes (within 1 kb), and *P*[Xbar] is a probability of how the Xbar value is deviated from a normal distribution of Xbar values of the entire genome of a sample.

For ChIP-chip analysis, 500 ng of immunoprecipitated and input DNA, without any amplification, was labeled with Cy5 and Cy3, respectively, and then hybridized with the microarray. A scan of the microarray and the data processing were performed as described above. The levels of each histone modification or Pol II binding were assessed by the signal ratio (IP/WCE). Genes were classified into those with high and low levels of each histone modification or Pol II binding when they had signal intensities higher and lower, respectively, than the average signal intensity of total probes. The microarray data (MeDIP-CGI microarray and ChIP-chip analyses) were submitted to the GEO database under accession no. GSE15154.

Calculation of Me value

The Me value of each probe was calculated as Me value = [signal log ratio × (1 − *P*[Xbar]) − 1.3]/2.6 + 0.5. The Me value was developed to give a value between 0 and 1 that linearly correlates with the amount of methylated DNA molecules at a specific locus and is not influenced by the genome-overall methylation levels. The Me value of a single probe is known to correlate well with an average DNA methylation level of CpG sites within 200 bp from the probe (Yamashita et al. 2009).

Definition of genomic regions

The position of each probe against a TSS was determined using UCSC hg18 (NCBI Build 36.1, March 2006). A CGI was defined as an assembly of probes with intervals <500 bp. CGIs were classified into four categories, promoter CGIs (within 10 kb upstream of the TSS), divergent CGIs (within 10 kb upstream of the TSSs of two genes that are transcribed in opposite directions), gene body CGIs, and downstream CGIs (within 10 kb downstream from genes). A CGI spanning both a promoter region and gene body was split into a promoter CGI and a gene body CGI. A putative NFR was defined as a region between a TSS, determined by UCSC hg18 (NCBI Build 36.1, March 2006), and its 200 bp upstream. Since TSSs are inherently variable for some genes (Suzuki et al. 2001), and the size of NFRs are different according to studies (Yuan et al. 2005; Gal-Yam et al. 2006), the locations are approximate, but expected to be correct as a whole. According to these definitions, 34,697 assemblies of probes were defined as CGIs, and 9624 assemblies were defined as NFRs. Genes with multiple NFRs because of their multiple TSSs were analyzed as different genes. DNA methylation status and histone modifications/Pol II binding in each CGI (or NFR) were assessed by an average Me value and signal ratio, respectively, of the probes located within each CGI (or NFR). A single CGI (or NFR) contains 6.8 (2.0) probes on average.

Gene expression analysis by oligonucleotide microarray

Expression microarray analysis was performed by a GeneChip Human Genome U133 Plus 2.0 expression microarray (Affymetrix) that contained 54,000 probe sets from 39,000 genes. From 8 μg of total RNA, the first-strand cDNA was synthesized with SuperScript III reverse transcriptase (Invitrogen) and a T7-(dT) 24 primer (Amersham Bioscience). Double-stranded cDNA was then synthesized, and biotin-labeled cRNA was synthesized using a BioArray HighYield RNA transcript labeling kit (Enzo). Twenty micrograms of labeled cRNA was fragmented and hybridized to the GeneChip oligonucleotide microarray. The microarray was stained and scanned according to the protocol from Affymetrix. The scanned data were processed using GeneChip operating software (ver. 1.4). The signal intensity of each probe was normalized so that the average signal intensity of all the probes on a microarray would be 500. Average signal intensity of all the probes for a gene was used as its transcription level. Genes were classified into those with high (>1000), moderate (250–1000), and low (<250) transcription according to their signal intensities.

Multivariate analysis and other statistical tests

To evaluate the independent contribution of each predictor variable (H3Ac, H3K4me3, Pol II binding, H3K9me3, or H3K27me3 level) in relation to the other four predictor variables on DNA methylation susceptibility (an outcome variable), multivariate logistic regression analysis was performed. Susceptible genes were defined as (1) those moderately and highly methylated in cancer cell lines (Int, and S1–S4 for the prostates; Int, and S1–S3 for the mammary glands), or (2) those highly methylated in cancer cell lines (S1–S4 for the prostates; and S1–S3 for the mammary glands). The predictor variables were classified into quintiles according to

H3Ac, H3K4me3, Pol II binding, H3K9me3, or H3K27me3 levels of the NFRs to create dummy variables. This was done because a log linear relationship was unclear between the raw value (signal ratio of each gene) and DNA methylation susceptibility. Multivariate-adjusted ORs and 95% confidence intervals (CIs) of genes in each quintile for DNA methylation susceptibility were calculated, including all predictor variables simultaneously in the model using SAS software, ver. 9.1 (SAS Institute Inc, *SAS/STAT 9.1 User's Guide*, SAS Institute Inc., Cary, NC). Using the lowest quintile as a reference, we calculated multivariate-adjusted ORs of genes in each quintile, which reflect DNA methylation susceptibility relative to the reference while controlling for the simultaneous effect of all the other predictor variables included in the model.

The fractions of genes with low transcription were compared between different groups of genes by the χ^2-test. The transcription, histone modification, and Pol II binding levels were compared between two groups of genes by the Mann-Whitney's *U*-test.

Functional annotation analysis

Functional annotation analysis was performed by DAVID bioinformatics resources (Dennis et al. 2003; Huang et al. 2009). The enrichment of genes in a biological process (a Gene Ontology criterion) was analyzed by comparing a fraction of genes with an ontology among the resistant (or susceptible) genes with that among all the genes.

Acknowledgments

We thank Hiroyuki Sasaki for his critical reading of this manuscript. H.T. is a recipient of Research Resident Fellowships from the Foundation for Promotion of Cancer Research. This study was supported by Grants-in-Aid for the Third-Term Comprehensive Cancer Control Strategy from the Ministry of Health, Labour and Welfare, Japan; for the Priority-area Research from the Ministry of Education, Science, Culture, and Sport, Japan; and a grant from Uehara Life Science Foundation.

References

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129:** 823–837.

Boeger H, Bushnell DA, Davis R, Griesenbeck J, Lorch Y, Strattan JS, Westover KD, Kornberg RD. 2005. Structural basis of eukaryotic gene transcription. *FEBS Lett* **579:** 899–903.

Boumber YA, Kondo Y, Chen X, Shen L, Guo Y, Tellez C, Estecio MR, Ahmed S, Issa JP. 2008. An Sp1/Sp3 binding polymorphism confers methylation protection. *PLoS Genet* **4:** e1000162. doi: 10.1371/journal.pgen.1000162.

Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, Wright FA, Feramisco JD, Peltomaki P, Lang JC, et al. 2000. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat Genet* **24:** 132–138.

Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* **4:** R60. doi: 10.1186/gb-2003-4-9-r60.

De Smet C, Loriot A, Boon T. 2004. Promoter-dependent mechanism leading to selective hypomethylation within the 5′ region of gene MAGE-A1 in tumor cells. *Mol Cell Biol* **24:** 4781–4790.

Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, et al. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* **38:** 1378–1385.

Erfurth FE, Popovic R, Grembecka J, Cierpicki T, Theisler C, Xia ZB, Stuart T, Diaz MO, Bushweller JH, Zeleznik-Le NJ. 2008. MLL protects CpG clusters from methylation within the *Hoxa9* gene, maintaining transcript expression. *Proc Natl Acad Sci* **105:** 7517–7522.

Esteller M, Corn PG, Baylin SB, Herman JG. 2001. A gene hypermethylation profile of human cancer. *Cancer Res* **61:** 3225–3229.

Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ. 2006. Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. *Nat Genet* **38:** 540–549.

Gal-Yam EN, Jeong S, Tanay A, Egger G, Lee AS, Jones PA. 2006. Constitutive nucleosome depletion and ordered factor assembly at the GRP78 promoter revealed by single molecule footprinting. *PLoS Genet* **2:** e160. doi: 10.1371/journal.pgen.0020160.

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130:** 77–88.

Hahn MA, Hahn T, Lee DH, Esworthy RS, Kim BW, Riggs AD, Chu FF, Pfeifer GP. 2008. Methylation of Polycomb target genes in intestinal cancer is mediated by inflammation. *Cancer Res* **68:** 10280–10289.

Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, Rappsilber J, Lerdrup M, Helin K. 2008. A model for transmission of the H3K27me3 epigenetic mark. *Nat Cell Biol* **10:** 1291–1300.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4:** 44–57.

Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. 2001. Accelerated age-related CpG island methylation in ulcerative colitis. *Cancer Res* **61:** 3573–3577.

Jones PA, Baylin SB. 2007. The epigenomics of cancer. *Cell* **128:** 683–692.

Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E, Pikarski E, Young RA, Niveleau A, Cedar H, et al. 2006. Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat Genet* **38:** 149–153.

Kirmizis A, Bartley SM, Kuzmichev A, Margueron R, Reinberg D, Green R, Farnham PJ. 2004. Silencing of human Polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes & Dev* **18:** 1592–1605.

Kondo Y, Kanai Y, Sakamoto M, Mizokami M, Ueda R, Hirohashi S. 2000. 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* **32:** 970–979.

Kondo Y, Shen L, Cheng AS, Ahmed S, Boumber Y, Charo C, Yamochi T, Urano T, Furukawa K, Kwabi-Addo B, et al. 2008. Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nat Genet* **40:** 741–750.

Laird PW, Jaenisch R. 1996. The role of DNA methylation in cancer genetic and epigenetics. *Annu Rev Genet* **30:** 441–464.

Li B, Carey M, Workman JL. 2007. The role of chromatin during transcription. *Cell* **128:** 707–719.

Lin JC, Jeong S, Liang G, Takai D, Fatemi M, Tsai YC, Egger G, Gal-Yam EN, Jones PA. 2007. Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island. *Cancer Cell* **12:** 432–444.

Lodygin D, Epanchintsev A, Menssen A, Diebold J, Hermeking H. 2005. Functional epigenomics identifies genes frequently silenced in prostate cancer. *Cancer Res* **65:** 4218–4227.

Maekita T, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M, Arii K, Kaneda A, Tsukamoto T, Tatematsu M, et al. 2006. High levels of aberrant DNA methylation in *Helicobacter pylori*-infected gastric mucosae and its possible association with gastric cancer risk. *Clin Cancer Res* **12:** 989–995.

Mihara M, Yoshida Y, Tsukamoto T, Inada K, Nakanishi Y, Yagi Y, Imai K, Sugimura T, Tatematsu M, Ushijima T. 2006. Methylation of multiple genes in gastric glands with intestinal metaplasia: A disorder with polyclonal origins. *Am J Pathol* **169:** 1643–1651.

Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K. 2007. RNA polymerase is poised for activation across the genome. *Nat Genet* **39:** 1507–1511.

Nakajima T, Yamashita S, Maekita T, Niwa T, Nakazawa K, Ushijima T. 2009. The presence of a methylation fingerprint of *Helicobacter pylori* infection in human gastric mucosae. *Int J Cancer* **124:** 905–910.

Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W, et al. 2007. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat Genet* **39:** 237–242.

Oka D, Yamashita S, Tomioka T, Nakanishi Y, Kato H, Kaminishi M, Ushijima T. 2009. The presence of aberrant DNA methylation in non-cancerous esophageal mucosae in association with smoking history: A target for risk diagnosis and prevention of esophageal cancers. *Cancer* **115:** 3412–3426.

Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Graf S, Tomazou EM, Backdahl L, Johnson N, Herberth M, et al. 2008. An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res* **18:** 1518–1529.

Robertson KD. 2005. DNA methylation and human disease. *Nat Rev Genet* **6:** 597–610.