

DNA Methylation and Expression Patterns of Key Tissue-specific Genes in Adult Stem Cells and Stomach Tissues

CpG-island margins and non-island-CpG sites round the transcription start sites of CpG-island-positive and -negative genes are methylated to various degrees in a tissue-specific manner. These methylation-variable CpG sites were analyzed to delineate a relationship between the methylation and transcription of the tissue-specific genes. The level of tissue-specific transcription was estimated by counting the number of the total transcripts in the SAGE (serial analysis of gene expression) database. The methylation status of 12 CpG-island margins and 21 non-island CpG sites near the key tissue-specific genes was examined in pluripotent stromal cells obtained from fat and bone marrow samples as well as in lineage-committed cells from marrow bulk, stomach, colon, breast, and thyroid samples. Of the 33 CpG sites examined, 10 non-island-CpG sites, but none of the CpG-island margins were undermethylated concurrent with tissue-specific expression of their nearby genes. The net methylation of the 33 CpG sites and the net amount of non-island-CpG gene transcripts were high in stomach tissues and low in stromal cells. The present findings suggest that the methylation of the non-island-CpG sites is inversely associated with the expression of the nearby genes, and the concert effect of transitional-CpG methylation is linearly associated with the stomach-specific genes lacking CpG-islands.

Key Words : DNA Methylation; Stem Cells; Stomach; Tissue-Specific Gene; CpG Islands

Seung-Jin Hong^{1,*}, Moo-Il Kang^{2,*},
Jung-Hwan Oh², Yu-Chae Jung¹,
Young-Ho Kim¹, Sung-Ja Kim¹,
Seung-Hye Choi³, Eun-Joo Seo⁴,
Sang-Wook Choi², and Mun-Gan Rhyu¹

Departments of Microbiology¹, Internal Medicine²,
Surgery³, and Clinical Pathology⁴, College of Medicine,
The Catholic University of Korea, Seoul, Korea

*Seung-Jin Hong and Moo-Il Kang contributed equally
to this work.

Received : 4 February 2009
Accepted : 1 July 2009

Address for correspondence

Mun-Gan Rhyu, M.D.
Department of Microbiology, College of Medicine,
The Catholic University of Korea, 505 Banpo-dong,
Seocho-gu, Seoul 137-701, Korea
Tel : +82.2-590-1215, Fax : +82.2-596-8969
E-mail : rhyumung@catholic.ac.kr

This work was supported by grants 2006005041 from
Catholic Institute of Cell therapy Basic Science Programs
Foundation made in the program year of 2006.

INTRODUCTION

Mammalian development initiates transcription factor cascades and DNA methylation to maintain the orderly transcription of housekeeping and tissue-specific genes (1). During this process, the DNA methyltransferase 3 establishes a genome-wide methylation pattern to lock the inactive state of the tissue-specific genes, while the DNA methyltransferase 1 plays a vital role in the maintenance of a tissue-specific methylation pattern (1). In addition, in vivo and vitro adult tissue studies have shown the multi-lineage differentiation of stromal cells derived from fat and bone marrow tissues (2, 3). Specifically, marrow stem cells have been found to migrate to gastric mucosa infected with *Helicobacter pylori*, where they then undergo gastric epithelial differentiation (4). Errors in the maintenance of DNA methylation, which occur much more frequent than errors in DNA replication, are believed to create variations in methylation patterns associated with the process of tissue development and renewal (5).

CpG-rich islands in germ-line and adult somatic tissues remain unmethylated throughout the lifetime of individuals,

while promoter regions lacking CpG islands are methylated to various degrees (6, 7). Accordingly, the CpG islands usually overlap with the unmethylated promoters of the housekeeping genes, whereas the tissue-specific genes that do not contain CpG islands have methylation-variable sites in their promoter regions (8). The methylation-variable sites that are prone to replication errors have been expected to induce the dynamic process of development and differentiation (9, 10). However, the relationship between the methylation and transcription of individual genes has remained elusive due to the complexity of the methylated CpG position and content (7).

Alu and L1 retroelements are highly repetitive sequences that are dispersed throughout the human genome as a result of long-term evolution (11). Several studies have suggested that the densely methylated retroelements influence the gene expression profiles via methylation-based mechanism (12, 13). In fact, it has been demonstrated that the transitional area between the unmethylated promoters and nearby methylated retroelements, such as the margin of the CpG islands and the non-island-CpG sites near the transcription start sites, is methylated in a tissue-specific manner (13, 14) (Fig. 1). There-

fore, the variable methylation of the transitional-CpG sites is likely to mediate the regulation of cell differentiation.

To delineate the methylation patterns of the key tissue-specific genes in pluripotent cells and lineage-committed tissues, the methylation-variable sites of housekeeping genes with CpG islands and tissue-specific genes without CpG islands were examined using seven adult tissues including fat and bone marrow stromal cells and gastrointestinal tissues. The non-island-CpG sites of the genes lacking CpG islands in stomach tissues tended to be undermethylated in a pattern consistent with the stomach-specific strong expression of their nearby genes. In addition, there was a linear association between the net methylation of the transitional-CpG sites and the expression of tissue-specific genes without CpG islands. This suggests that the methylation of the transitional-CpG sites has both negative individual and positive concert effects on the expression of the stomach-specific genes lacking CpG islands.

MATERIALS AND METHODS

Adult stromal and glandular tissue preparation

Stromal cells from 15 fat and 18 bone marrow samples, 7 marrow bulk samples, and glandular tissue samples from 12 thyroids, 16 breasts, 15 colons, and 26 stomachs were collected from 41 males and 68 females between 18 and 66 yr old during September 2005 and January 2007 (Table 1). The Institutional Review Board approved this study and the written informed consent was obtained from each individual that



Fig. 1. A schematic diagram of the methylation-variable CpG sites in transitional-CpG sites close to transcriptional start sites (TSS). The transitional-CpG sites are distinguished from so-called T-DMRs (tissue-dependent differentially methylated regions) widespread in the human genome regardless of gene-control regions.

provided samples (Document number 48, 'Genetic study of gastrointestinal cancer', January 28, 2005, St. Paul's Hospital, The Catholic University of Korea, Seoul, Korea). Fat tissues and bone marrow aspiration samples were cultured for the preparation of stromal cells using a previously described method (10). The cell pellets isolated from marrow aspiration specimens were used as marrow bulk cells. The glandular tissues of the thyroid and breast were dissected microscopically from surgically resected adenoma tissues. Normal mucosa of the colon and stomach antrum and body separated by a 5-cm distance were obtained by endoscopic biopsy. The presence of *H. pylori* in normal stomach mucosa was confirmed using the Warthin-Starry silver impregnation method. Frozen or fresh tissues were then digested in a Tween 20-Proteinase K lysis buffer for 12 hr at 50°C, after which the genomic DNA was extracted using a DNA isolation Kit (A1120, Promega, Madison, WI, U.S.A.) according to the manufacturer's instruction.

DNA modification with sodium bisulfite

Bisulfite conversion of the genomic DNA extracted from the fresh tissue was performed as described previously (10, 14). Briefly, 1 μ g of genomic DNA was denatured with 10 μ L of 3 M NaOH for 15 min at 37°C prior to modification with sodium bisulfite. Next, 100 μ L of the denatured DNA was treated with 1,040 μ L of 2.3 M sodium bisulfite and 60 μ L of 10 mM hydroquinone for 12 hr at 50°C. The modified DNA was then purified using a Wizard DNA purification resin (Promega), after which it was precipitated with ethanol and then dissolved in 35 μ L of 5 mM Tris buffer (pH 8.0). A 1 μ L aliquot of the modified DNA solution was then placed in a 200 μ L microcentrifuge tube and stored at -20°C until further analysis.

Radioisotope-labeling methylation-specific PCR analysis

Methylation-specific PCR (MSP) primer sets designed to contain a GC content >40% and 3 to 5 CpG dinucleotides near the 3'-position. The MSP sites are shown in Fig. 2. All the MSP primer sets covered a short amplicon of \leq 150 bp and a small number of CpG dinucleotides. These CpG-poor MSP sites reduced the difference in GC content between the

Table 1. Adult stromal cells and lineage-committed tissues examined for tissue-specific methylation

	Stromal cells		Lineage-committed tissues				
	Fat	Marrow	Marrow bulk	Thyroid	Breast	Colon	Stomach
No. of samples	15	18	7	12	16	15	26
Age (yr)	Mean	33	31	46	39	46	43
	Range	23-58	22-45	22-45	23-66	18-49	27-56
Sex	Male	2	16	5		9	9
	Female	13	2	2	12	16	6

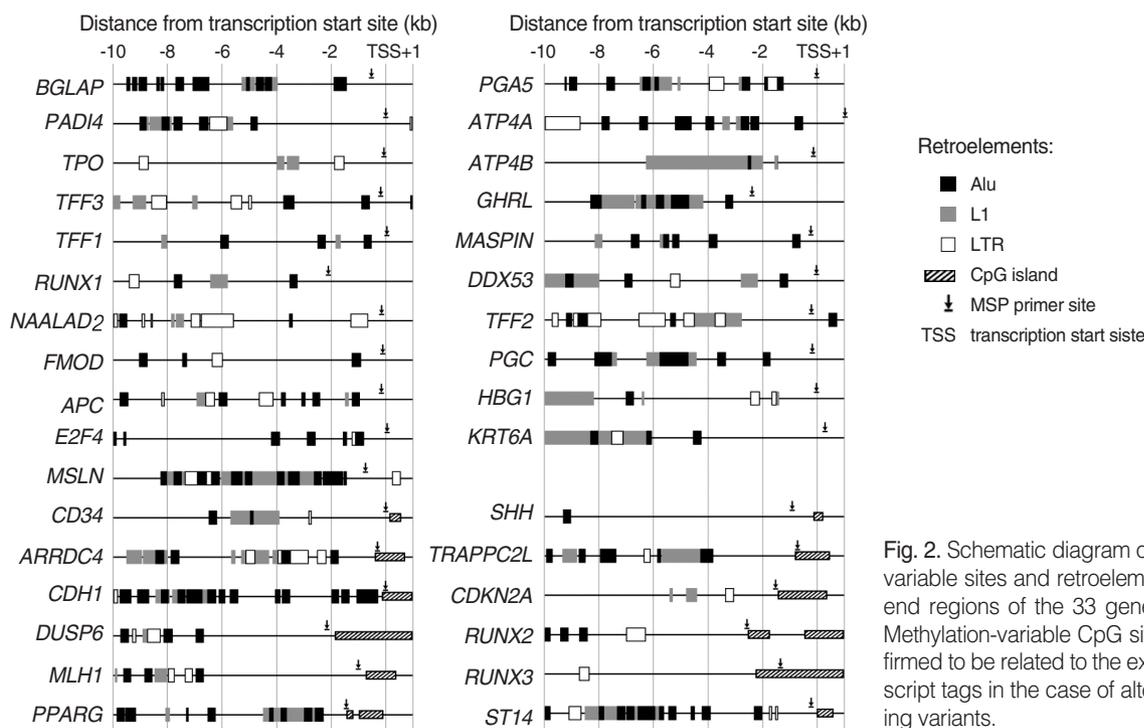


Fig. 2. Schematic diagram of methylation-variable sites and retroelements in the 5'-end regions of the 33 genes examined. Methylation-variable CpG sites were confirmed to be related to the expressed transcript tags in the case of alternative splicing variants.

unmethylated and methylated DNA following bisulfite modification, and then allowed the linear amplification of the unmethylated and methylated sequences under a stringent PCR condition (Fig. 3A).

Repeated radioisotope-labeling experiments have demonstrated that the dTTP-radioisotope, in conjunction with a low concentration of dNTP, produces a balance of PCR band intensities from the same amount of unmethylated and methylated DNA (14-16). Finally, a 1 μ L of bisulfite-modified DNA was labeled with 1 μ Ci of α - 32 P dTTP (PerkinElmer, Boston, MA, U.S.A.) in 10 μ L PCR mixture that contained 62.5 μ M of dATP, dCTP and dGTP, 25 μ M of dTTP, 0.5 pM of each primer, 0.1% Tween 20, and 0.06 μ L of *Taq* polymerase (5 U/ μ g, R001, Takara, Shiga, Japan). This mixture was subjected to 32 PCR cycles under a hot start condition to reach sub-plateau DNA amplification. The band intensity was measured by repeated autoradiography using a radioluminograph scanner (BAS 2500, Fuji Photo Film, Kanakawa, Japan) and the TINA image software (Raytest Isotopenme β gerate, Straubenhardt, Germany).

The common-PCR primer sets were designed to encompass both the unmethylated and methylated CpGs in the CpG amplicons. Each CpG amplicon was analyzed by cloning and sequencing the common-PCR DNA as previously described (10, 14, 15). The methylation-variable sites estimated by the radioisotope method were similarly confirmed by cloning and sequencing of common PCR products (16) (Fig. 3B). The PCR specificity of each MSP primer set was then validated by plotting a standard curve for a mixture containing different ratios of the methylated and unmethylated DNA clones.

Next, the density of methylation was calculated as the relative proportion of the methylated-band intensity to the total intensity of unmethylated and methylated bands following a method that has been described elsewhere (10, 14-17). The density of methylation was then reported using the following scale: level 1 (0-20% methylation), level 2 (21-40% methylation), level 3 (41-60% methylation), level 4 (61-80% methylation), and level 5 (81-100% methylation). The band sharpness and reproducibility of wide-range variable methylation were improved with the radioisotope-labeling protocol when compared with the non-radioisotope PCR bands including real-time PCR and with the results of common PCR DNA sequencing (Fig. 3, 4). The radioisotope-labeling amplification of bisulfite-modified DNA was found to reproduce the same levels of band intensities in >95% of 20 experiments conducted using the duplicated DNAs for each MSP primer set.

Tissue-specific transcription estimated using *in-silico* data

A total of 17,723 reference genes identified in a public database (<http://genome.ucsc.edu/>, March 2006 assembly) were analyzed to calculate the GC percent and CpG content in a 6-kb segment spanning from 5-kb upstream to 1-kb downstream of the transcription start site. Genes with a CpG ratio of >0.75 were categorized as high-CpG genes, whereas those with a CpG ratio of <0.48 were categorized as low-CpG genes and those with a CpG ratio of 0.48 to 0.75 were classified as intermediate-CpG genes. There were 12,220 high-CpG genes (69%), 2,685 intermediate-CpG genes (15%), and 2,818 low-CpG genes (16%).

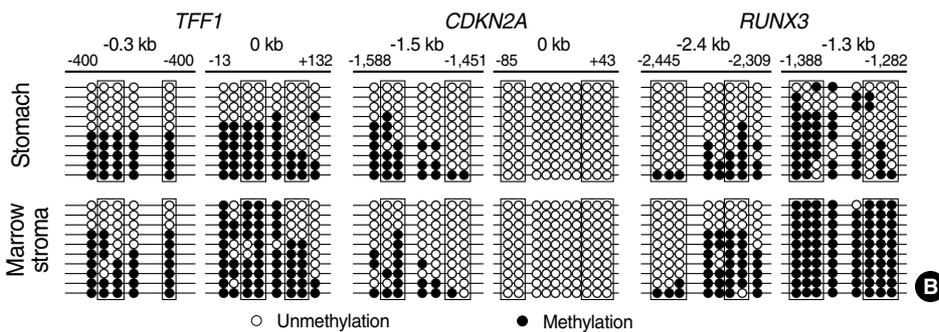
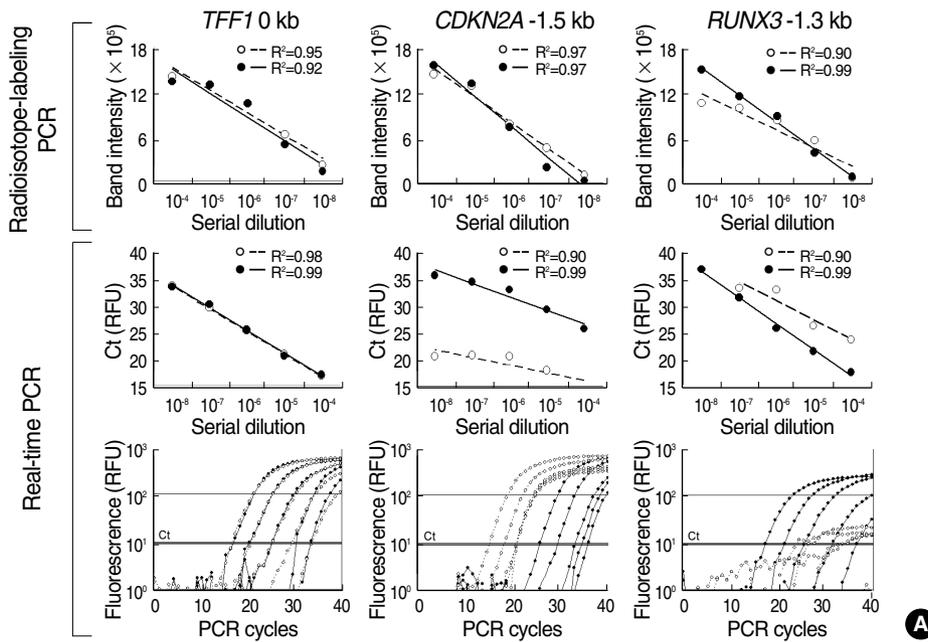


Fig. 3. Standard curves (A) and common PCR DNA sequencing (B) at methylation-variable sites nearest to the *TFF1*, *CDKN2A*, and *RUNX3* genes. (A) The cycle threshold (Ct) is calculated with relative fluorescence unit (RFU). (B) The box indicates the MSP primer position.

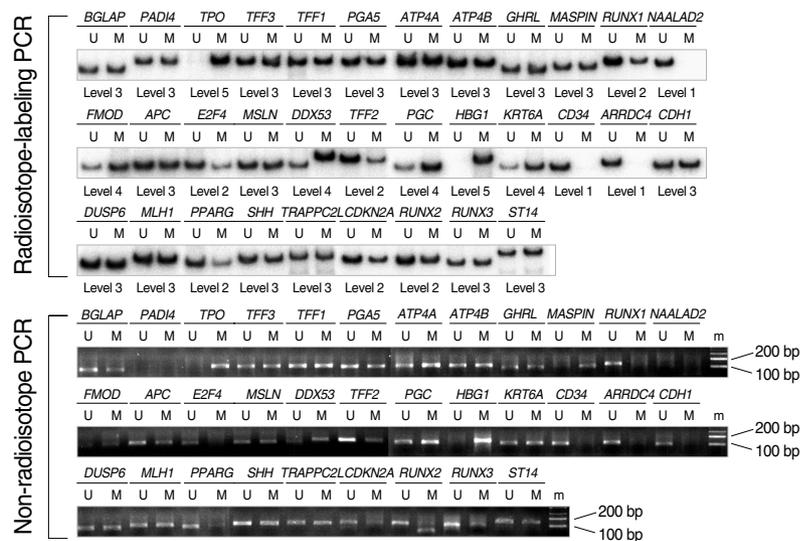


Fig. 4. Autoradiograph and ethidium bromide staining of electrophoretic bands of 33 methylation-variable sites examined in stomach tissue. The level of methylation is indicated below each pair of unmethylation (U) and methylation (M) PCR bands.

To clarify the activity of tissue-specific transcription, we analyzed the SAGE data estimating the number of expressed sequence tags (i.e. transcript) as well as the microarray data estimating the intensity of transcript hybridization signal.

The SAGE database (<http://www.ncbi.nlm.nih.gov/SAGE/>) provided the absolute number of expressed tags for genes in embryonic stem cells, bone marrow cells with CD34+/CD38-, marrow bulk, thyroid, mammary gland, colon, and stomach

antrum and body. The sum of the expressed tags was directly used as a quantitative measure for the corresponding gene activity. Microarray data (GEO, <http://www.ncbi.nlm.nih.gov/geo/> and MIAME, <http://www.ebi.ac.uk/arrayexpress/>) were used to analyze the gene expression profiles of fat and bone marrow stromal cells and breast, thyroid, colon and stomach tissues. For the microarray analysis, each dataset was median centered and the standard deviation was then normalized to one per array. Based on the highest number of transcripts in a given tissue as well as by reviewing available literature regarding their roles in cell differentiation, the 33 genes for MSP study were classified into 27 tissue-specific genes and 5 housekeeping genes (Table 2, 3). The housekeeping genes were found to be similarly expressed or inactive in various tissue

types.

Statistical analysis

An independent *t* test was used to determine if there were any significant differences in the level of methylation between 1) the CpG sites associated and not associated with their nearby gene expression, 2) *H. pylori*-infected and non-infected gastric mucosa, and 3) stromal cells and lineage-committed tissues. Statistical analysis was conducted using the SPSS ver. 11.0 software. Two-sided *P* values <0.05 were considered to be statistically significant.

Table 2. Expression patterns of key tissue-specific genes in pluripotent cells (embryonic stem cell and hematopoietic progenitor cells) and lineage-committed tissues (marrow bulk, thyroid, breast, colon, and stomach)

Gene	CpG island*	CpG content in 6-kb promoter region [†]	No. of expressed sequences tags									
			Embryonic stem cell	Marrow CD34+/CD38-	Marrow CD34+/CD38+	Marrow bulk	Thyroid	Breast	Colon	Stomach antrum	Stomach body	
<i>BGLAP</i>	No	Low	1				2					
<i>PADI4</i>	No	Low		2		1						
<i>TPO</i>	No	Intermediate					108					1
<i>TFF3</i>	No	Low	3			1	329	14	103	14	14	1
<i>TFF1</i>	No	Intermediate			1			165	186	11	11	95
<i>PGA5</i>	No	Low	2		1		2			734	4,145	
<i>ATP4A</i>	No	Low	1	3	1	1	1			175	83	
<i>ATP4B</i>	No	Intermediate	1							43	48	
<i>GHRL</i>	No	Low								3	30	
<i>MASPIN</i>	No	Intermediate										1
<i>RUNX1</i>	No	Low		1	5	1	2	11			2	1
<i>NAALAD2</i>	No	Intermediate	4	2	2		39	3	1	3	3	1
<i>FMOD</i>	No	Low				1	25		5	30	6	
<i>APC</i>	No	Intermediate	3		1	6	2			2		
<i>E2F4</i>	No	Intermediate	6	9	9	4	8	2	3	13	5	
<i>MSLN</i>	No	Intermediate							1	50		
<i>DDX53</i>	No	Intermediate	2				1					
<i>TFF2</i>	No	Low			1			1		632	116	
<i>PGC</i>	No	Low		2	3	2	1	1		33	348	
<i>HBG1</i>	No	Low	1	2			2					
<i>KRT6A</i>	No	Low	4	1		1	8	2	3	2	1	
<i>CD34</i>	Yes	High	4	1		1	8	2	3	2	1	
<i>ARRDC4</i>	Yes	High	4	1	1	1	1		5	7		
<i>CDH1</i>	Yes	High	31	9	1		41	19	43	19	15	
<i>DUSP6</i>	Yes	High	7	5	4	5	27	3	8	8	1	
<i>MLH1</i>	Yes	High	5	1			1	1				
<i>PPARG</i>	Yes	High						1	5	3		
<i>SHH</i>	Yes	High	2									
<i>TRAPPC2L</i>	Yes	High	21	10	5		9	3		14	1	
<i>CDKN2A</i>	Yes	Intermediate	2			1	2	1		1	1	
<i>RUNX2</i>	Yes	High	1									
<i>RUNX3</i>	Yes	Intermediate			3			1				
<i>ST14</i>	Yes	High	3	2	3		2		1	41	1	

*More than 50% GC percent, greater than 200 bps length, and over 0.6 ratio of observed to expected number of CpG dinucleotide to the expected G+C number (<http://genome.ucsc.edu>). The length and position of CpG island are different at each gene; [†]Classification of CpG content is described in Materials and Methods section.

Table 3. Methylation patterns of 33 genes examined in fat and bone marrow stromal cells and lineage-committed tissues (marrow bulk, thyroid, breast, colon, and stomach)

Gene	Tissue specificity	Mean level of methylation estimated by 5-level classification							
		Fat stroma	Marrow stroma	Marrow bulk	Thyroid	Breast	Colon	Stomach antrum	Stomach body
Inverse association between gene-specific methylation and expression									
<i>BGLAP</i>	Bone	3.0	2.3	2.7	3.5	2.8	3.3	2.7	2.7
<i>PADI4</i>	Marrow bulk	3.6	2.8	2.0	3.8	3.6	3.3	3.6	3.5
<i>TPO</i>	Thyroid	4.9	5.0	5.0	2.8	5.0	4.8	5.0	5.0
<i>TFF3</i>	Thyroid	2.9	2.9	3.0	2.7	3.6	3.7	3.3	3.5
<i>TFF1</i>	Colon, Stomach, Breast	4.3	4.8	5.0	4.8	2.6	1.8	2.2	2.2
<i>PGA5</i>	Stomach	3.2	3.1	3.1	3.1	3.3	3.0	2.6	2.5
<i>ATP4A</i>	Stomach	4.1	4.2	4.4	4.4	4.4	4.1	3.5	3.7
<i>ATP4B</i>	Stomach	5.0	4.8	5.0	5.0	4.8	4.7	3.3	3.1
<i>GHRL</i>	Marrow bulk, Stomach	4.0	4.3	2.3	3.7	4.0	3.9	3.2	3.2
<i>MASPIN</i>	Breast	5.0	5.0	5.0	4.6	1.3	2.9	2.7	2.8
No association between expression and methylation									
<i>RUNX1</i>	Marrow bulk	1.3	1.0	1.0	1.2	2.9	1.3	1.2	1.3
<i>NAALAD2</i>	Thyroid	1.0	1.0	1.0	1.0	1.5	1.5	1.9	1.8
<i>FMOD</i>	Thyroid, Stomach	1.0	1.3	3.0	1.3	3.4	4.9	3.2	3.4
<i>APC</i>	Colon, Stomach	1.1	1.3	1.0	1.6	1.3	1.2	2.5	2.7
<i>E2F4</i>	Housekeeping	2.1	1.8	2.1	3.8	2.4	1.5	2.5	2.3
<i>MSLN</i>	Stomach	3.7	4.3	4.6	3.7	3.9	4.4	3.6	3.7
<i>DDX53</i>	Testis	3.3	4.7	4.4	5.0	3.4	4.2	4.1	4.3
<i>TFF2</i>	Stomach	2.9	2.2	4.1	3.9	2.9	3.0	2.8	2.6
<i>PGC</i>	Stomach	4.3	4.3	4.9	4.9	3.0	4.2	2.8	3.0
<i>HBG1</i>	Fetus	4.0	4.2	4.6	4.1	4.1	3.4	4.7	4.6
<i>KRT6A</i>	Tongue, Esophagus	3.9	3.9	4.6	4.6	3.9	4.0	4.0	4.1
<i>CD34</i>	Hematopoietic stem cell	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
<i>ARRDC4</i>	Housekeeping	1.0	1.0	1.0	1.0	1.0	1.1	1.7	1.5
<i>CDH1</i>	Non-neural epithelial cells	1.0	1.0	1.0	1.0	1.0	1.9	1.8	1.6
<i>DUSP6</i>	Housekeeping	1.0	1.0	1.3	1.0	1.3	1.6	2.0	2.2
<i>MLH1</i>	Housekeeping	1.1	1.1	1.3	1.3	1.1	2.9	2.4	2.3
<i>PPARG</i>	Adipose tissue	1.3	1.0	1.1	1.3	2.1	1.8	1.9	1.7
<i>SHH</i>	Fetus	1.5	1.0	1.3	1.6	3.1	2.7	3.3	3.3
<i>TRAPPC2L</i>	Marrow bulk	1.9	2.1	3.4	2.6	2.0	2.7	2.9	2.8
<i>CDKN2A</i>	Housekeeping	2.3	1.2	1.7	1.8	3.3	1.7	1.9	1.8
<i>RUNX2</i>	Marrow bulk	3.3	1.0	1.0	1.0	2.4	1.5	1.8	2.0
<i>RUNX3</i>	Colon, Stomach	1.2	5.0	4.9	2.5	3.4	2.9	2.7	2.5
<i>ST14</i>	Pancreas, Kidney	3.8	3.1	2.6	3.6	1.9	3.6	2.7	3.0

Methylation level of tissue-specific gene in the corresponding tissue type is indicated by boldface.

RESULTS

Expression profiles of key tissue-specific genes

The number of transcripts of the 33 genes evaluated by analyzing the SAGE data is listed in Table 2. The eight genes lacking CpG islands (*TPO*, *PGA5*, *ATP4A*, *ATP4B*, *NAA-LAD2*, *MSLN*, *TFF2*, and *PGC*) were strongly expressed in a single type of glandular tissues derived from the endoderm layer. Both the *TFF1* and *TFF3* genes without CpG islands tended to be commonly expressed in glandular tissues. Although the *BGLAP*, *PADI4*, *MASPIN*, *RUNX1*, and *APC* genes lacking CpG islands and the *CD34*, *PPARG*, *SHH*, *RUNX2*, and *RUNX3* genes containing CpG islands are well-

known to be expressed in a tissue-specific manner, the SAGE data showed a low-number or no expressed tags corresponding to these genes in any of the tissues examined here. Consequently, the SAGE data counting the absolute number of transcripts demonstrated that the expression of stomach-specific genes lacking CpG islands was remarkably strong when compared with other tissue-specific genes.

Tissue-specific methylation associated with tissue-specific gene expression

The methylation density of methylation-variable sites was calculated using the 5-level classification of CpG amplicon (Fig. 5). The 21 non-island-CpG sites close to the transcrip-

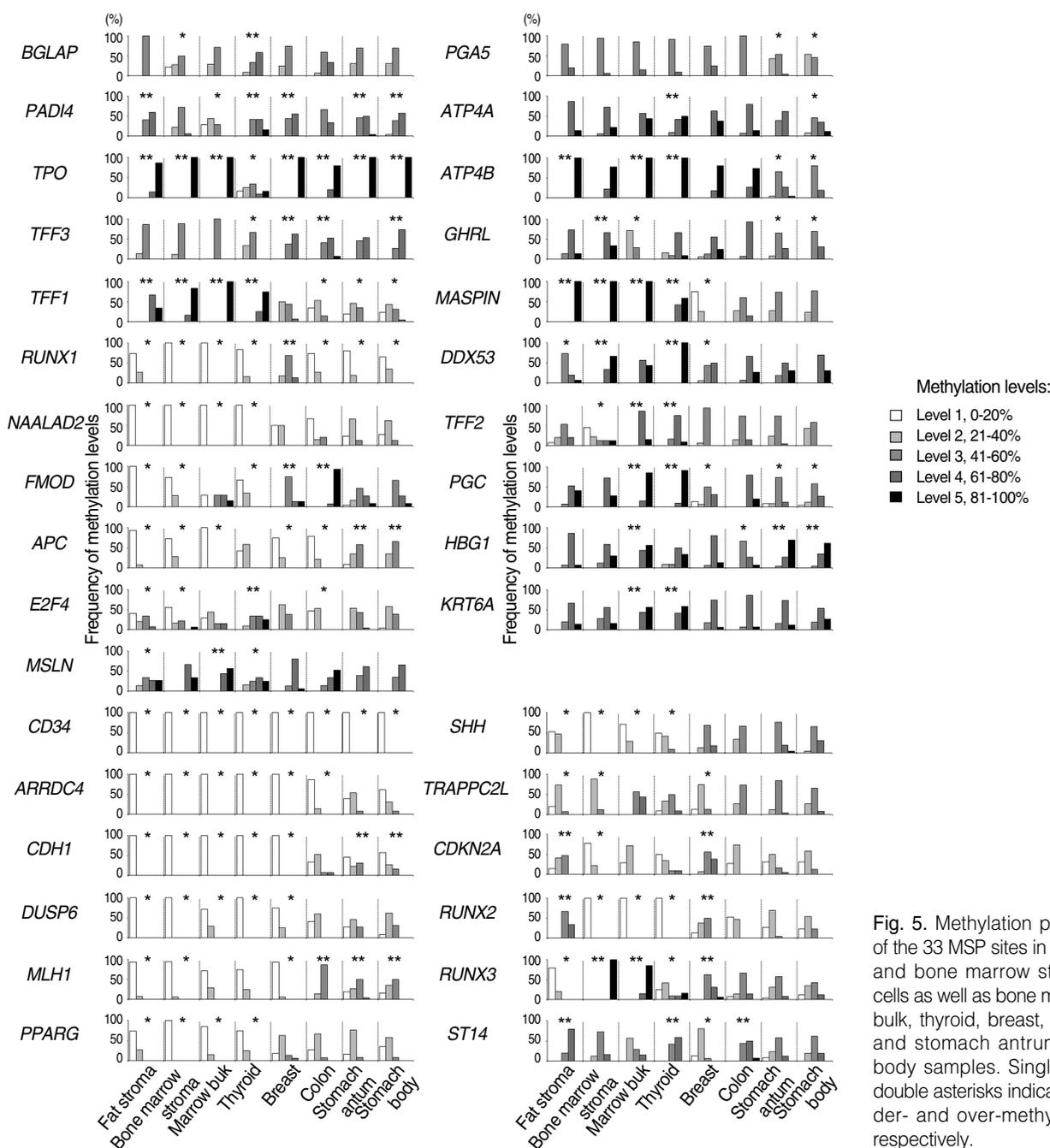


Fig. 5. Methylation profiles of the 33 MSP sites in the fat and bone marrow stromal cells as well as bone marrow bulk, thyroid, breast, colon, and stomach antrum and body samples. Single and double asterisks indicate under- and over-methylation respectively.

tion start sites of the genes lacking CpG islands tended to be densely methylated at a level of ≥ 3 , whereas the 12 genes containing CpG islands were weakly methylated at a level of ≤ 3 in all tissue types. The mean methylation level of the non-island-CpG genes was high when compared with that of the CpG-island genes (3.32 vs. 1.95) (Fig. 6A).

The methylation status of each gene in a given tissue type was then scored based on the mean level of methylation estimated in each tissue-type samples. If the mean methylation was lowest in the corresponding tissue type, DNA methylation and transcription were considered to have an inverse asso-

ciation. An inverse association between DNA methylation and transcription was observed in 10 non-island-CpG genes, *BGLAP*, *PADI4*, *TPO*, *TFF3*, *TFF1*, *PGA5*, *ATP4A*, *ATP4B*, *GHRL*, and *MASPIN*. However, no inverse association between methylation and transcription was found in any of the CpG-island genes.

These 10 non-island-CpG sites in all of the tissues tended to be more densely methylated when compared with the 11 non-island-CpG sites that were not associated with the expression of their nearby genes (mean level of methylation, 3.55 vs. 2.95, $P < 0.001$) (Fig. 6A). Therefore, the 33 CpG sites were

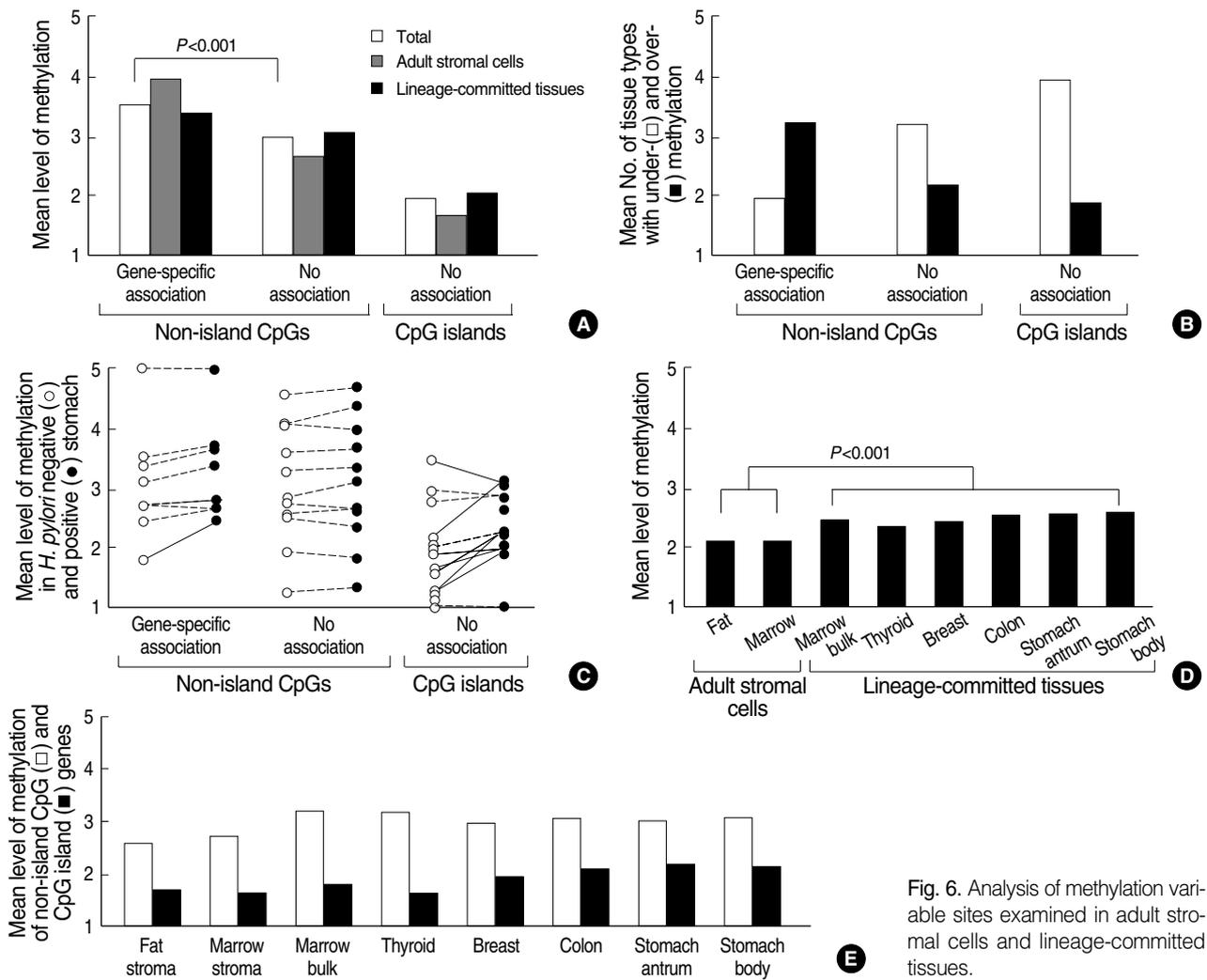


Fig. 6. Analysis of methylation variable sites examined in adult stromal cells and lineage-committed tissues.

divided into 10 dense-methylation sites inversely associated with the expression of their nearby genes and 23 weak-methylation sites, including 11 non-island-CpGs and 12 CpG islands, not associated with the expression of their nearby genes. When the methylation of each gene was categorized as under-methylated or overmethylated based on the median level of methylation observed in the seven tissue types (Fig. 5), both the weak-methylation genes with and without CpG islands were overmethylated in one or two tissue types (mean number of over- vs. under-methylated tissue types; 1.1 vs. 3.7 and 1.5 vs. 2.7, respectively, Fig. 6B). One (*TFF1*) of the 21 non-island-CpG genes and six (*ARRDC4*, *CDH1*, *MLH1*, *PPARG*, *CDKN2A*, and *RUNX3*) of the 12 CpG-island genes were densely methylated in gastric mucosa infected with *H. pylori* when compared with non-infected mucosa (Fig. 6C).

Methylation-transcription relations in stromal cells and lineage-committed tissues

The mean methylation level of the 23 weak-methylation

sites was significantly lower in the fat and marrow stromal cells than in lineage-committed tissues ($P < 0.001$) (Fig. 6D). In addition, the methylation level of these weak-methylation sites in various tissue samples was high in the following order: body and antrum of the stomach > colon > bone marrow bulk > breast > thyroid > stromal cells of fat and bone marrow. The histogenetically related cells such as marrow stroma vs. bulk showed a difference in the methylation level of both non-island-CpG and CpG islands (Fig. 6E). A comparison of glandular tissues illustrated that the methylation of CpG islands was high in the following order: stomach > colon > breast > thyroid.

Due to differences in the method of transcription measurement between SAGE and microarray analyses, all active genes were sorted in descending order based on their transcription activities and top 2,000 active genes were selected from the SAGE and microarray database (Fig. 7A). The proportion of CpG-island gene transcripts in the total transcripts was higher in embryonic and hematopoietic stem cells than in marrow bulk and four glandular tissues, stomach, colon, breast, and

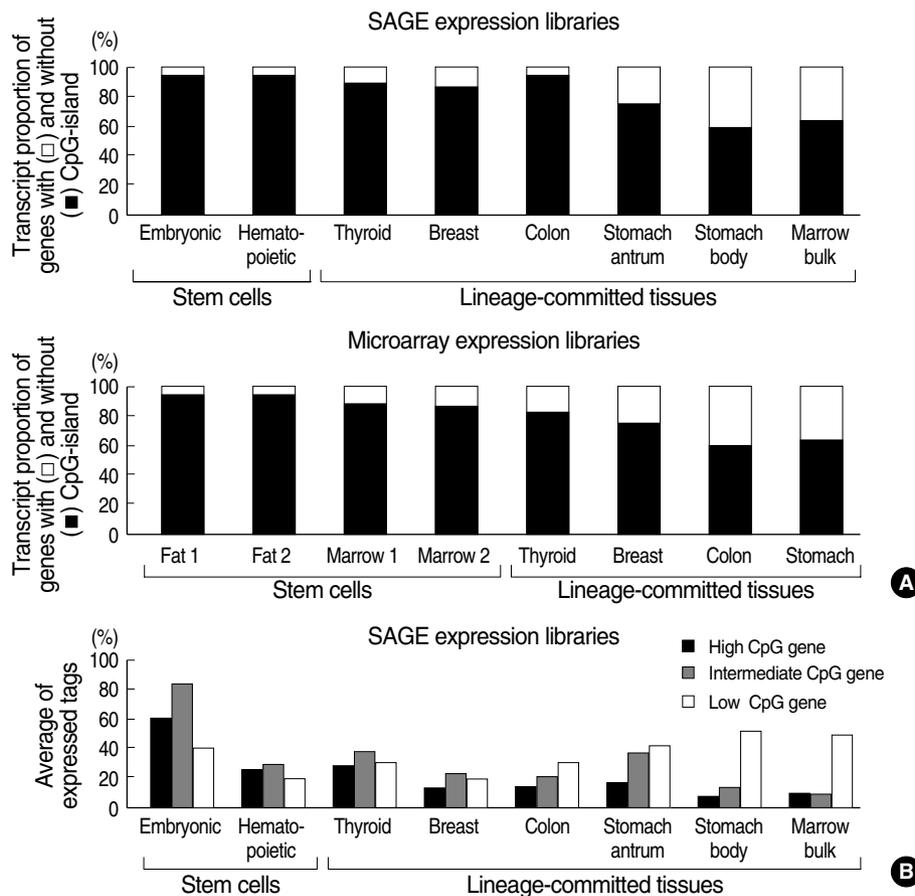


Fig. 7. The proportion of CpG-island and non-island-CpG gene transcripts in stem cells and lineage-committed tissues.

thyroid. In addition, the microarray data revealed that the expression of CpG-island genes was higher in marrow and fat stromal cells than in the four glandular tissues. The transcription of the non-island-CpG genes in the glandular tissues that were estimated using both the SAGE and microarray data was high in the following order: stomach>colon>breast>thyroid. When the CpG content of the genes was divided into high-, intermediate-, and low-CpG genes (Fig. 7B), the average transcription of low-CpG genes was highest in the stomach and marrow bulk. Consequently, analyses of tissue-specific gene methylation and transcription and global gene transcription demonstrated that the expression of non-island-CpG genes was relatively decreased with the net methylation of both non-island-CpGs and CpG islands in bone marrow tissues and was relatively increased with that of CpG islands in glandular tissues.

DISCUSSION

A series of studies have reported that there is an inverse correlation between tissue-specific expression and methylation in approximately 30% of genes, most of which have no CpG islands in the gene-control regions (6, 7). However, a large proportion of the high- and low-CpG promoters are unmethylated and methylated, respectively, regardless of their transcriptional activities (7). In this study, the 10 key tissue-specific genes that were less methylated and more strongly expressed in a given tissue type were always found to have no CpG islands in the gene-control regions. The mean methylation of the key tissue-specific genes estimated in extraneous tissue types was denser than that of the remaining 11 non-island-CpG sites and 12 CpG-island margins. Interestingly, the weak methylated of non-island-CpG sites and CpG island margins that were not associated with their nearby gene expression was the lowest in the pluripotent stromal cells and was linearly associated with the relative proportion of non-island-CpG gene expression to CpG-island gene expression in lineage-committed tissues.

It has been suggested that the methylation of retroelements associated with human genome evolution initiates the spread of DNA methylation into the adjacent gene control regions during embryogenesis (12). The tissue-specific genes without CpG islands that were close to the retroelements (14) showed an inverse association between gene-specific methylation and expression, and were densely methylated when compared with the genes that were expressed irrespective of their nearby CpG methylation (Table 3 and Fig. 6A). The chromatin structure is closely related to DNA methylation and organizes an intranuclear dynamic network by which the nucle-

ar processes for transcription and replication rapidly respond to cell-intrinsic and -extrinsic stimuli (18). The initial transcriptional activation of key tissue-specific genes results in chromatin being remodeled after which transcription factors are recruited to the key gene where a group of genes initiates the formation of a transcription hub (19). The relatively undermethylated tissue-specific genes with dense-methylation-CpG sites then likely nucleate the master transcription center to coordinate gene expression for the developmental transcription cascade and cell lineage specification.

The mean level of weak methylated non-island-CpG sites and CpG island margins was found to increase in differentiated cells in comparison of marrow stromal versus bulk cells (Fig. 6E) in consistent with a previous study (20). Additionally, each weak-methylation site was tissue-specifically overmethylated in one or two tissue types whereas the dense methylation site was tissue-specifically undermethylated (Fig. 6B). Therefore, the increased methylation of weak-methylation sites indicated a late epigenetic process following initial activation of master genes with dense-methylation sites. When considering that Polycomb proteins in adult stem cells promote DNA methylation to induce multilineage differentiation, which is also known as epigenetic memory module (21), the weak-methylation sites appeared to be further methylated in adult tissues that were under the influence of the epigenetic memory module.

In this study, the overmethylation of the CpG islands was common in gastric mucosa infected with *H. pylori* (Fig. 6C) in consistent with a previous report (22). An analysis of the surface-to-gland axis of the *H. pylori*-infected mucosa revealed that all major gastric genes lacking CpG islands were weakly expressed in middle stem-line cells and strongly expressed in surface differentiated cells (23). Given that pluripotent stem cells derived from the bone marrow often undergo gastric epithelial differentiation in gastric mucosa infected with *H. pylori* (4), a linear association between weak-methylation sites and non-island-CpG genes observed in this study suggests that the increased methylation of the CpG-island margins up-regulates the inactive or weakly active gastric genes lacking CpG islands in the *H. pylori*-infected mucosa containing marrow cells.

The marrow and fat stromal cells up-regulated the house-keeping genes with CpG island and down-regulated the tissue-specific genes without CpG islands in consistent with the transcription-ready state of embryonic stem cells (Fig. 7A), which was also described in a previous study (24). Previous studies have suggested that random changes in the neutral CpG methylation that result from replication errors during DNA methylation are asymmetrically divided among two distinct progeny cells, stem-line and differentiated cells (5, 25). The error-prone changes in gene-regulatory methylation sites might result in an increase or decrease in DNA methylation and were selected and expanded depending on their effects on the expression profiles. It is possible that progeny

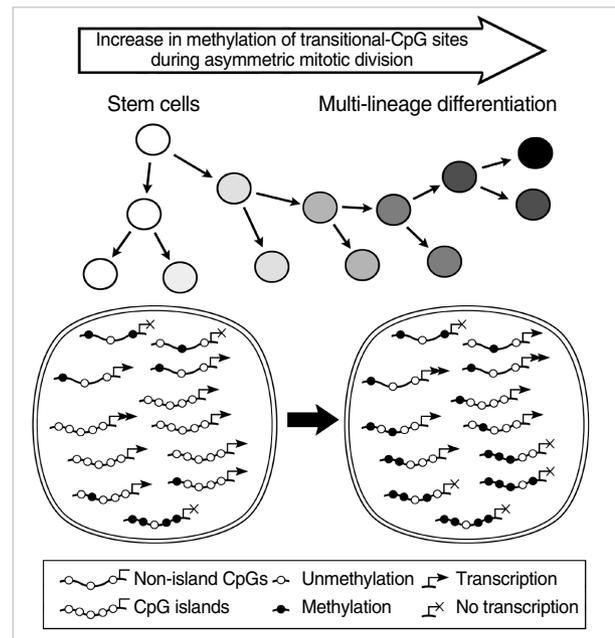


Fig. 8. A schematic diagram of variable methylation created during asymmetric stem cell division. A progeny cell that attains a low or high level of variable methylation via asymmetric cell division remains a stem-line cell or activates the CpG-poor tissue-specific genes, respectively.

cells that attain a high level of methylation in weak methylated CpG sites up-regulate the gastric genes lacking CpG islands during marrow-derived cell differentiation in the gastric mucosa, whereas progeny cells with a low level of weak methylation remain in the stem-line lineage (Fig. 8).

The *PGA5*, *PGC*, *TFF1*, *TFF2*, *TFF3*, and *GHRL* genes were expressed more strongly in one of the antrum and body sites (Table 2). In addition, the *MUC5AC*, *MUC5B*, and *MUC6* genes lacking CpG islands were found to be weakly expressed in the stem-line cells and strongly expressed in either the surface or bottom differentiated cells of the gastric mucosa (23). Consequently, all the gastric-specific genes had no CpG islands and were complementarily expressed in the two distinctive subsites. This suggests that in the early developmental stage as well as adult tissue environment, the asymmetrically increased methylation of methylation-variable sites up-regulates the non-island CpG gastric genes in the distinctive subsites (Fig. 8). Although the Polycomb proteins promote histone modifications in adult stem cells, there is a little evidence of asynchronous histone modifications (26). Moreover, the newly synthesized DNA in adult tissues was methylated largely independent of histone modifications (27). Taken together, the fine regulation of the non-CpG-island gastric genes for the bifurcation of the stem-line and differentiated cells appears to depend on a DNA methylation-based mechanism.

Highly expressed genes can organize a master chromatin compartment in the nucleus, in which CpG-poor tissue-spe-

cific and CpG-rich housekeeping genes share the transcription machinery for their coordinate regulation (18). In a nuclear space that contains the limited amount of nuclear proteins, the low-CpG genes may have a higher probability of using the nuclear proteins in a manner complementary to the high-CpG genes. The increased methylation of weak-methylation sites is likely to decrease the accessibility of a chromatin compartment to nuclear proteins, thereby allowing the CpG-poor genes to use more nuclear proteins than the high-CpG genes. Only 16% of the total genes have a low CpG content (based on analysis of SAGE data), which indicates that a large number of weak methylation segments has a strong concert effect on a small number of the tissue-specific genes without CpG islands. In fact, the CpG-poor genes on the active X chromosome are hypertranscribed when compared with those on the autosomes (28). Therefore, the non-island-CpG gene linearly associated with the mean methylation of the 33 transitional-CpG sites (Fig. 6E), which was less methylated and inactive in the marrow stem cells, might be strongly expressed in gastric mucosa and marrow bulk cells as a result of the concerted effect of the weak-methylation sites.

We compared DNA methylation profiles of 33 selected genes examined in seven tissue types to SAGE data estimating the entire gene transcripts in other samples of the same tissue types. This experimental design is perhaps a potential flaw in credible data showing a direct relationship between the methylation and expression of individual and global genes. However, the relative proportion of non-island-CpG gene expression to CpG-island gene expression was relatively increased with the mean methylation of the 33 transitional CpG sites near key tissue-specific and housekeeping genes in lineage-committed tissues. Both self-organization and dosage-compensation models strongly suggest epigenetic chromosome structure in which CpG-poor tissue-specific genes and CpG-rich housekeeping genes reciprocally interact with each other. Therefore, at least, it is likely that a set of the selected transitional CpG sites partially reflects tissue-type-dependent methylation levels in different tissue types (12, 13).

A low GC content and repetitive sequence in the methylation-variable site limited the template-primer and amplicon-probe specificities due to the reduced sequence complexity of heterogeneously methylated DNA following bisulfite treatment (29). Accordingly, non-radioisotope PCR often showed weak or smearing bands when amplifying the variable-methylation sites (Fig. 4). Additionally, the spotted and variegated methylation of transitional-CpG sites varied over a wide range and could not be estimated by high-resolution methods. The 5-level classification of transitional methylation was found to be reproducible using the radioisotope-labeling protocol that was conducted under a stringent PCR condition.

In conclusion, the methylation of tissue-specific genes without CpG islands is inversely associated with the expression of their nearby dense-methylation sites. The weak-methylation sites are less methylated concurrent with the weak ex-

pression of the tissue-specific genes in the fat- and marrow-derived stromal cells, and are the most methylated in stomach tissues that strongly expressed the non-island-CpG genes. These findings suggest that the methylation-variable sites can exert the individual and concert effects on the strong expression of the stomach-specific genes without CpG islands.

ACKNOWLEDGMENTS

We wish to thank Tae-Min Kim of Department of Microbiology of College of Medicine, The Catholic University of Korea, for help with mining and interpretation of microarray datasets and comments on the manuscript. The authors have no competing interests.

REFERENCES

1. Li E. *Chromatin modification and epigenetic reprogramming in mammalian development*. *Nat Rev Genet* 2002; 3: 662-73.
2. Zuk PA, Zhu M, Mizuno H, Huang J, Futrell JW, Katz AJ, Benhaim P, Lorenz HP, Hedrick MH. *Multilineage cells from human adipose tissue: implications for cell-based therapies*. *Tissue Eng* 2001; 7: 211-28.
3. Jiang Y, Jahagirdar BN, Reinhardt RL, Schwartz RE, Keene CD, Ortiz-Gonzalez XR, Reyes M, Lenvik T, Lund T, Blackstad M, Du J, Aldrich S, Lisberg A, Low WC, Largaespada DA, Verfaillie CM. *Pluripotency of mesenchymal stem cells derived from adult marrow*. *Nature* 2002; 418: 41-9.
4. Houghton J, Stoicov C, Nomura S, Rogers AB, Carlson J, Li H, Cai X, Fox JG, Goldenring JR, Wang TC. *Gastric cancer originating from bone marrow-derived cells*. *Science* 2004; 306: 1568-71.
5. Yatabe Y, Tavare S, Shibata D. *Investigating stem cells in human colon by using methylation patterns*. *Proc Natl Acad Sci U S A* 2001; 98: 10839-44.
6. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, Haefliger C, Horton R, Howe K, Jackson DK, Kunde J, Koenig C, Liddle J, Niblett D, Otto T, Pettett R, Seemann S, Thompson C, West T, Rogers J, Olek A, Berlin K, Beck S. *DNA methylation profiling of human chromosomes 6, 20 and 22*. *Nat Genet* 2006; 38: 1378-85.
7. 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.
8. Larsen F, Gundersen G, Lopez R, Prydz H. *CpG islands as gene markers in the human genome*. *Genomics* 1992; 13: 1095-107.
9. Noer A, Sorensen AL, Boquest AC, Collas P. *Stable CpG hypomethylation of adipogenic promoters in freshly isolated, cultured, and differentiated mesenchymal stem cells from adipose tissue*. *Mol Biol Cell* 2006; 17: 3543-56.
10. Kang MI, Kim HS, Jung YC, Kim YH, Hong SJ, Kim MK, Baek KH, Kim CC, Rhyu MG. *Transitional CpG methylation between promot-*

- ers and retroelements of tissue-specific genes during human mesenchymal cell differentiation. *J Cell Biochem* 2007; 102: 224-39.
11. Medstrand P, van de Lagemaat LN, Mager DL. Retroelement distributions in the human genome: variations associated with age and proximity to genes. *Genome Res* 2002; 12: 1483-95.
 12. Arnaud P, Goubely C, Pelissier T, Deragon JM. SINE retroposons can be used in vivo as nucleation centers for de novo methylation. *Mol Cell Biol* 2000; 20: 3434-41.
 13. Turker MS. Gene silencing in mammalian cells and the spread of DNA methylation. *Oncogene* 2002; 21: 5388-93.
 14. Kang MI, Rhyu MG, Kim YH, Jung YC, Hong SJ, Cho CS, Kim HS. The length of CpG islands is associated with the distribution of Alu and L1 retroelements. *Genomics* 2006; 87: 580-90.
 15. Kim YH, Hong SJ, Jung YC, Kim SJ, Seo EJ, Choi SW, Rhyu MG. The 5'-end transitional CpGs between the CpG islands and retroelements are hypomethylated in association with loss of heterozygosity in gastric cancers. *BMC Cancer* 2006; 6: 180.
 16. Jung YC, Hong SJ, Kim YH, Kim SJ, Kang SJ, Choi SW, Rhyu MG. Chromosomal losses are associated with hypomethylation of the gene-control regions in the stomach with a low number of active genes. *J Korean Med Sci* 2008; 23: 1068-89.
 17. Hong SJ, Kim YH, Choi YD, Min KO, Choi SW, Rhyu MG. Relationship between the extent of chromosomal losses and the pattern of CpG methylation in gastric carcinomas. *J Korean Med Sci* 2005; 20: 790-805.
 18. Misteli T. Beyond the sequence: cellular organization of genome function. *Cell* 2007; 128: 787-800.
 19. Chakalova L, Debrand E, Mitchell JA, Osborne CS, Fraser P. Replication and transcription: shaping the landscape of the genome. *Nat Rev Genet* 2005; 6: 669-77.
 20. Tagoh H, Melnik S, Lefevre P, Chong S, Riggs AD, Bonifer C. Dynamic reorganization of chromatin structure and selective DNA demethylation prior to stable enhancer complex formation during differentiation of primary hematopoietic cells in vitro. *Blood* 2004; 103: 2950-5.
 21. Rajasekhar VK, Begemann M. Concise review: roles of polycomb group proteins in development and disease: a stem cell perspective. *Stem Cells* 2007; 25: 2498-510.
 22. Yoo EJ, Park SY, Cho NY, Kim N, Lee HS, Kang GH. Helicobacter pylori-infection-associated CpG island hypermethylation in the stomach and its possible association with Polycomb repressive marks. *Virchows Arch* 2008; 452: 515-24.
 23. Van De Bovenkamp JH, Korteland-Van Male AM, Buller HA, Einderhand AW, Dekker J. Infection with Helicobacter pylori affects all major secretory cell populations in the human antrum. *Dig Dis Sci* 2005; 50: 1078-86.
 24. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. *Hum Mol Genet* 2005; 14 Spec No 1: R47-58.
 25. Beckmann J, Scheitza S, Wernet P, Fischer JC, Giebel B. Asymmetric cell division within the human hematopoietic stem and progenitor cell compartment: identification of asymmetrically segregating proteins. *Blood* 2007; 109: 5494-501.
 26. Kiel MJ, He S, Ashkenazi R, Gentry SN, Teta M, Kushner JA, Jackson TL, Morrison SJ. Haematopoietic stem cells do not asymmetrically segregate chromosomes or retain BrdU. *Nature* 2007; 449: 238-42.
 27. Easwaran HP, Schermelleh L, Leonhardt H, Cardoso MC. Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO Rep* 2004; 5: 1181-6.
 28. Lin H, Gupta V, Vermilyea MD, Falciani F, Lee JT, O'Neill LP, Turner BM. Dosage compensation in the mouse balances up-regulation and silencing of X-linked genes. *PLoS Biol* 2007; 5: 326.
 29. Tusnady GE, Simon I, Varadi A, Aranyi T. BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. *Nucleic Acids Res* 2005; 33: 9.