Chromosomal Losses are Associated with Hypomethylation of the Gene-Control Regions in the Stomach with a Low Number of Active Genes

Transitional-CpG methylation between unmethylated promoters and nearby methylated retroelements plays a role in the establishment of tissue-specific transcription. This study examined whether chromosomal losses reducing the active genes in cancers can change transitional-CpG methylation and the transcription activity in a cancer-type-dependent manner. The transitional-CpG sites at the CpG-island margins of nine genes and the non-island-CpG sites round the transcription start sites of six genes lacking CpG islands were examined by methylation-specific polymerase chain reaction (PCR) analysis. The number of active genes in normal and cancerous tissues of the stomach, colon, breast, and nasopharynx were analyzed using the public data *in silico*. The CpG-island margins and non-island CpG sites tended to be hypermethylated and hypomethylated in all cancer types, respectively. The CpG-island margins were hypermethylated and a low number of genes were active in the normal stomach compared with other normal tissues. In gastric cancers, the CpG-island margins and non-island-CpG sites were hypomethylated in association with high-level chromosomal losses, and the number of active genes increased. Colon, breast, and nasopharyngeal cancers showed no significant association between the chromosomal losses and methylation changes. These findings suggest that chromosomal losses in gastric cancers are associated with the hypomethylation of the gene-control regions and the increased number of active genes.

Key Words : *Chromosomal Loss; Loss of Heterozygosity; CpG Methylation; Methylation-Specific PCR; Tissue Expression Profiles*

Yu-Chae Jung, Seung-Jin Hong,
Young-Ho Kim, Sung-Ja Kim,
Seok-Jin Kang*, Sang-Wook Choi†,
and Mun-Gan Rhyu

Departments of Microbiology, Clinical Pathology*, and Internal Medicine†, College of Medicine, The Catholic University of Korea, Seoul, Korea

Address for correspondence
Mun-Gan Rhyu, M.D.
Department of Microbiology, The Catholic University of Korea, 505 Banpo-dong, Seocho-gu, Seoul 135-701, Korea
Tel : +82.2-590-1215, Fax : +82.2-596-8969
E-mail : rhyumung@catholic.ac.kr

INTRODUCTION

The establishment of tissue-specific gene expression profiles relies on a variety of transcription factor cascades, which initiate embryogenesis and development, and on DNA methylation, which prevents the reactivation of inactive genes (1). The promoter regions of the housekeeping genes always overlap with the CpG islands, whereas approximately 60% of the genes lacking CpG islands are expressed in a tissue-specific manner (2, 3). However, an inverse correlation between the methylation state and gene expression is found in only 30% of the genes randomly selected (4). Therefore, it is believed that rather than simply acting as an on-off switch in gene regulation, DNA methylation is essential for maintaining the well-organized gene expression in a variety of somatic tissues (5), and that changes in the normal methylation state might play a role in the genesis and development of cancers (6).

The loss-of-heterozygosity (LOH) event detected by an analysis of microsatellite sequences represents the unilateral chromosomal loss, major genetic alterations observed in human solid tumors (7). Although the LOH event may be a part of bi-allelic gene inactivation, it is still unclear if a decrease in the chromosomal dose plays a role in carcinogenesis. Assuming the DNA methylation profile of a given tissue is established with respect to the global gene expression profiles, it is possible that the gene-copy number reduction caused by LOH can influence the methylation and transcriptional status of the genes on the chromosomal copies retained. However, there are a few reports describing the functional relationship between the genetic and epigenetic features of cancer cells, especially on how the LOH events and methylation changes collectively influence the global gene expression profiles of cancer cells. We previously observed that the methylation-variable CpG sites between unmethylated promoters and nearby methylated retroelements, including the CpG-island margins and the non-island-CpG sites the genes lacking CpG islands, are associated with the global as well as individual gene expression patterns (8-10). These findings lead us to propose an assumption that a decrease in the number of active gene copy caused by LOH might facilitate the methylation changes in the transitional-CpG sites as well as the changes in the total number of active genes.

In this study, we examined the chromosomal losses events, methylation changes, and global expression profiles in gastric, colonic, mammary and nasopharyngeal cancers in order to further delineate the interaction of genetic, epigenetic, and transcriptional alterations using LOH, methylation-specific polymerase chain reaction (PCR) and Serial Analysis of Gene Expression (SAGE) data.

MATERIALS AND METHODS

Collection of cancer and normal tissues

Formalin-fixed, paraffin-embedded tissue sections of four cancer types (stomach, colon, breast, and nasopharynx) were obtained from 100 patients who had undergone a surgical resection between March 2004 and July 2006 at St. Paul's Hospital and Kangnam St. Mary's Hospital, The Catholic University of Korea. The Institutional Review Board approved this study, and written informed consent was obtained from each patient before surgery. A single 10-mm-diameter site containing a homogeneous cell population was selected from each section that contained a representative of paired normal and tumor tissues. Seven-μm-thick hematoxylin-eosin-stained sections were microdissected under a $40\times$ stereomicroscope using a surgical scalpel. In the microscopic examination, the tumor specimens consisted mainly of tumor tissue and the normal tissues did not show any evidence of tumor cell invasion or significant inflammatory involvement. Approximately 50 microdissected cells were digested in 1 μL of a Tween 20-Proteinase K lysis buffer.

PCR-based microsatellite analysis

A pair of normal and cancer DNAs was examined using a panel of PCR primers that covered 40 microsatellite loci on eight cancer-associated chromosomes, 3p, 4p, 5q, 8p, 9p, 13q, 17p, and 18q, as reported elsewhere (11-14). The allelic profiles of the 40 microsatellite sequences were initially analyzed for any microsatellite instability (MSI) at the homozygous markers showing a few stutter bands in a pair of normal and cancer tissues. If the cancer DNA showed novel bands that were absent in normal DNA, in >40% of the homozygous microsatellite alleles, they were interpreted as being a MSI. The difference in the allelic intensity between the normal and cancer DNAs was scored as the relative allelic ratio calculated by dividing the intensity ratio of the cancer by the normal allelic ratio. A relative allelic ratio >1.5 in the heterozygous case without a MSI was interpreted as a LOH on the basis of the distribution of relative allelic ratios, as this provided the best discrimination between wild-type heterozygosity and LOH.

Methylation-specific PCR and sequencing of bisulfite-modified DNA

Ninety microliters of genomic DNA was denatured with 10 μL of 3 M NaOH for 15 min at 37°C and modified with 1,040 μL of 2.3 M sodium bisulfite and 60 μL of 10 mM hydroquinone for 12 hr at 50°C, as described elsewhere (12, 13). The methylation-variable CpG sites reported in previous study of 11 somatic tissue types were chosen in the transitional area between promoters and retroelements (9). A total of 15 CpG sites selected from six CpG-island-negative genes (*MAGEA2*, *TFF2*, *DDX53*, *MSLN*, *MASPIN*, and *BGLAP*) and nine CpG-island-positive genes (*MUC8*, *KIAA1752*, *CDKN2A*, *ESR2*, *PPARG*, *MLH1*, *CDH1*, *VDR*, and *RUNX3*) were examined using methylation-specific PCR (MSP) primer sets (Supplementary Table 1 and 2). Supplementary Fig. 1 shows the coverage of CpG islands along with

Table 1. The number of active genes and the number of tags expressed per active gene in embryos and somatic tissues

| Tissue | Genes with CpG islands | | | | | | Genes without CpG islands | | | |
| | L1 retroelement | | *Alu* and L1 retroelement | | *Alu* retroelement | | L1 retroelement | | *Alu* retroelement | |
	Active genes	Tags per gene	Active genes	Tags per gene	Active genes	Tags per gene	Active genes	Tags per gene	Active genes	Tags per gene
Embryo	513	7.1	3,576	15.3	5,817	16.7	999	5.7	810	6.7
Placenta	395	4.6	2,885	7.8	4,652	8.1	870	10.3	704	11.9
Normal tissues										
Stomach	213	2.3	1,676	4.3	2,636	4.2	444	11.4	368	5.0
Colon	246	3.9	2,108	5.0	3,529	5.1	543	7.3	480	3.8
Breast	332	3.5	2,386	5.6	3,988	5.9	672	4.6	543	5.2
Cancer tissues										
Stomach	315	3.0	2,273	5.7	3,736	6.0	624	4.8	540	5.5
Colon	289	2.5	2,382	4.2	3,910	4.5	643	4.0	541	3.0
Breast	334	3.1	2,474	5.4	4,122	5.9	641	3.9	552	4.0

The mean numbers of active genes and the expressed tags were calculated using two to eight SAGE libraries for each tissue type. Individual genes were classified according to on the presence or absence of CpG islands and the type of nearby retroelements in the 5′-end regions.

the distribution of retroelements in the 5´-end regions of the 15 genes examined.

For semiquantitative MSP analysis, a minimum number of PCR rounds to reach sub-plateau DNA amplification were performed using a radioisotope. The bisulfite-modified DNA was amplified and labeled by a hot-start PCR containing α-^{32}P dTTP (PerkinElmer, Boston, MA, U.S.A.) and dNTP mixture through 32 PCR cycles. The PCR products were loaded onto a nondenaturing polyacrylamide gel and visualized by repeated autoradiography using a radioluminograph scanner (BAS 2500, Fuji Photo Film, Kanakawa, Japan) and analyzed with TINA image software (Raytest Isotopenmeβgerate, Straubenhardt, Germany). The specificity of each MSP primer set was validated using a standard curve for the universal methylated and unmethylated DNAs, as described elsewhere (9).

The relative proportion of the methylated CpG band to the total intensity of the methylated and unmethylated CpG bands was calculated from the MSP bands amplified by the MSP primer set. The proportion of methylated CpGs was divided into 5 levels; 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 results of the LOH and MSP analyses on the four cancer types are listed in Supplementary Table 3.

The methylation-variable CpGs of the *TFF2*, *MSLN*, *BGLAP*, *CDKN2A*, and *MLH1* genes were analyzed by cloning and sequencing of the common PCR DNA (Supplementary Fig. 2). The common PCR primer sets were designed to span both the unmethylated and methylated CpGs in the CpG amplicons. The PCR product of each common primer set was cloned into the T&A cloning vector (Real Biotech, Taipei, Taiwan). The DNA sequencing was performed using a BigDye Terminator Kit (PE Biosystems,

Table 2. Statistical analysis (R values) of the number of expressed tags for gene expression similarities between the embryo and somatic tissues

		Genes with CpG islands						Genes without CpG islands			
		L1 element		*Alu* and L1		*Alu* element		L1 element		*Alu* element	
		Embryo	Placenta	Embryo	Placenta	Embryo	Placenta	Embryo	Placenta	Embryo	Placenta
Normal tissues											
Stomach	1	**0.48**	0.16	0.33	*0.59*	**0.44**	**0.45**	0.05	0.01	**0.20**	0.10
	2	**0.39**	0.17	0.31	*0.59*	*0.51*	*0.54*	0.04	0.02	**0.21**	0.07
	3	*0.57*	0.27	0.48	*0.72*	*0.63*	*0.81*	0.09	0.03	0.17	0.11
	4	**0.47**	0.42	*0.50*	*0.73*	**0.47**	*0.54*	0.10	0.27	**0.20**	0.12
Colon	1	0.34	0.10	0.20	0.43	0.20	0.32	0.33	0.14	0.02	0.05
	2	*0.62*	0.07	0.24	0.48	0.27	0.35	0.22	0.06	0.07	0.13
Breast	1	0.08	0.32	*0.69*	*0.59*	*0.68*	*0.75*	*0.73*	0.20	0.05	0.01
	2	0.15	0.19	*0.55*	0.44	*0.63*	*0.70*	*0.61*	0.17	0.06	0.01
	3	0.11	0.36	*0.57*	0.46	*0.70*	*0.70*	*0.51*	0.19	0.25	0.39
	4	0.17	0.34	*0.61*	*0.58*	*0.61*	*0.66*	*0.57*	0.21	0.19	0.90
	5	0.11	0.30	*0.61*	0.45	*0.70*	*0.67*	0.39	0.10	0.31	0.18
	6	0.15	0.24	*0.64*	*0.51*	*0.75*	*0.56*	*0.64*	0.21	0.24	0.08
Cancer tissues											
Stomach	1	0.26	0.17	*0.68*	*0.78*	*0.53*	*0.66*	*0.68*	0.15	0.29	0.19
	2	0.28	0.35	*0.66*	*0.57*	*0.53*	*0.70*	*0.55*	0.21	0.17	*0.79*
	3	0.21	0.21	*0.65*	*0.58*	*0.50*	*0.64*	0.41	0.14	0.20	*0.75*
	4	0.25	0.26	*0.59*	*0.57*	*0.68*	*0.50*	*0.64*	0.12	0.12	0.25
	5	0.33	0.30	*0.63*	*0.79*	*0.54*	*0.60*	*0.66*	0.22	0.38	*0.53*
	6	0.31	0.37	*0.66*	*0.83*	*0.72*	*0.73*	*0.73*	0.18	0.32	*0.78*
Colon	1	**0.45**	0.12	0.29	*0.50*	0.37	*0.54*	*0.55*	0.29	0.09	0.02
	2	0.29	0.13	0.38	0.35	0.24	0.46	0.44	0.14	0.35	0.28
Breast	1	0.34	*0.50*	*0.65*	*0.78*	0.46	*0.55*	*0.54*	0.17	0.21	*0.51*
	2	0.19	0.37	*0.60*	*0.77*	*0.56*	*0.53*	*0.56*	0.18	0.28	*0.58*
	3	0.14	0.13	*0.70*	*0.51*	*0.63*	0.45	*0.56*	0.08	0.30	0.07
	4	0.34	0.20	*0.69*	*0.72*	*0.72*	*0.67*	*0.59*	0.48	0.09	0.01
	5	0.17	0.28	*0.70*	*0.63*	*0.57*	*0.73*	0.39	0.08	0.31	0.18
	6	0.33	0.42	*0.64*	*0.76*	0.47	*0.54*	0.46	0.13	0.18	*0.56*
	7	0.18	0.35	*0.61*	*0.72*	0.40	0.44	*0.62*	0.18	0.20	0.31
	8	**0.40**	0.18	*0.70*	*0.73*	*0.58*	*0.67*	*0.54*	0.13	0.32	*0.82*

Using the SAGE data *in silico*, the number of expressed tags of individual genes was compared between the embryo (embryonic cell line of passage 16 and the first trimester placenta) and somatic tissues. Total genes were classified according to the presence or absence of CpG islands and the type of the nearby retroelements. Pearson's correlation coefficients of $p<0.05$ are marked in bold. R values ≥ 0.5 are indicated by italics.

Foster City, CA, U.S.A.) and an ABI automated DNA sequencer (PE Biosystems, Warrington, U.K.). The intensity of the MSP bands was compared with the distribution of methylated CpGs, which were determined by cloning and sequencing of the common PCR amplicons.

Methylation analysis of normal somatic tissues

The methylation status of the transitional CpGs examined in 11 normal somatic tissues was previously reported (9). This study extended the previous data by adding six transitional sites (*TFF2*, *MSLN*, *BGLAP*, *MUC8*, *KIAA1752*, and *PPARG*) and three somatic tissues (fat, ovary, and testis). A total of eleven tissue types were collected from 51 individuals. The bone marrow and fat stromal cells were obtained from three individuals (10) and each of the remaining nine tissue types were obtained from five individuals. The stomach and colon was divided into proximal and distal portions

according to the anatomic landmarks, gastric body and the splenic flexure of the colon.

Analysis of *in-silico* data

A total of 32 SAGE libraries for the embryonic stem cells, placenta, stomach, colon, and breast tissues were obtained from a public database (http://cgap.nci.nih.gov/SAGE/). Six expressed sequence tag (EST) libraries for the nasopharyngeal tissues were collected from a public database (http://cgap.nci.nih.gov/Tissues) because of no SAGE data available. The SAGE and EST libraries analyzed are listed in Supplementary Table 4. The SAGE tags and EST tags were assigned using UniGene cluster ID by accumulating the total expressed tags to the matched genes at each tissue library. A total of 434,325 expressed tags in the 32 SAGE libraries corresponded to the 15,770 gene symbols through tag-to-gene matching.
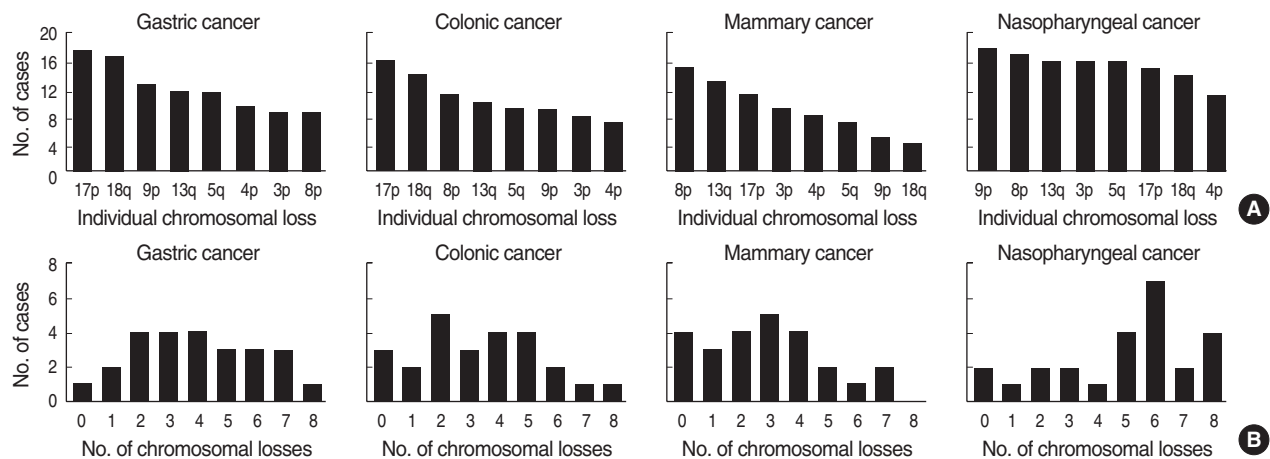
The genomic location of the NCBI RefSeq cDNA sequ-

Fig. 1. Chromosomal losses detected in the gastric, colonic, mammary, and nasopharyngeal cancers (25 cases for each cancer type). (A) Individual chromosomal losses and (B) the number of chromosomal losses were evaluated by PCR-based analysis using 40 microsatellite markers on chromosomes 3p, 4p, 5q, 8p, 9p, 13q, 17p, and 18q.

Table 3. The methylation status of CpG-island margins, the number of active genes, and the transcript number per active gene in embryo and somatic tissues

	Embryo	Placenta	Stomach	Colon	Breast	Nasopharynx
CpG-island margin methylation	+	++	++++	++++	+++	+++
Association between high-level LOH and hypomethylation	NE	NE	Yes	No	No	No
Genes with CpG islands						
No. of active genes	+++++	++++	+	++	+++	NE
Transcript number per gene	+++++	++++	+	++	+++	NE
Genes without CpG islands						
No. of active genes	+++++	++++	+	++	+++	NE
Transcript number per gene	++	++++	+++	++	+	NE
L1-close transcript proportion	+	+++	++++	+++	++	++

The methylation of the CpG-island margins were divided into four levels (+, ++, +++, and ++++) according to the genome-wide DNA methylation of the embryo (27, 28) and the placenta (29) and the CpG-island margin methylation estimated in somatic tissues (Fig. 4). The number of active genes and the transcript number per active gene were divided into four or five levels according to the ranks observed in the embryo and somatic tissues (Fig. 5). Hypomethylation changes in cancer tissues are detailed in Fig. 3. NE, not examined.

ences was obtained from the genome web site (UCSC Golden Path May, 2004 assembly, http://genome.ucsc.edu/). A 3-kb sized non-overlapping window was used to analyze a DNA segment upstream and downstream of the transcription start site. The coverage of CpG islands and the distribution of retroelements compiled from searches of the genome database were evaluated by inputting the sequence data that was delimited from the 5′-end regions into a local program. The annotation of retroelements was made using the Repeat-Masker program (http://ftp.genome.washington.edu/RM/RepeatMasker.html). A total of 15,770 active genes showing the tag-and-gene match were demarcated by the presence or absence of CpG islands at the transcription sites as well as the types of nearby retroelements existing in a 3-kb window. For the fidelity of the genome-wide expression data, we compared SAGE and Affymetrix GeneChip (http://symatlas.gnf.org/SymAtlas), and there was strong agreement in major mRNA content of the analyzed tissue types (data not shown) as previous reports (15, 16). However, the SAGE data was found to reliably reflect the wide-range of transcription level by counting the sequence-based 'digital' tags, while the microarray data based on the fluorescence signal was not suitable for defining the active or inactive transcription status as well as the estimation of strong transcription activities due to the limit of probe hybridization method.

Statistical analysis

A chi-square test was used to compare the methylation changes between gastric, colonic, mammary, and nasopharyngeal cancers. The Pearson's correlation coefficients of the expressed tag numbers in different tissues were calculated to determine the similarities of the individual gene expression. A two-sided p value <0.05 was considered significant.

RESULTS

The level of chromosomal losses estimated in four cancer types

The LOH events in each cancer were determined using 40
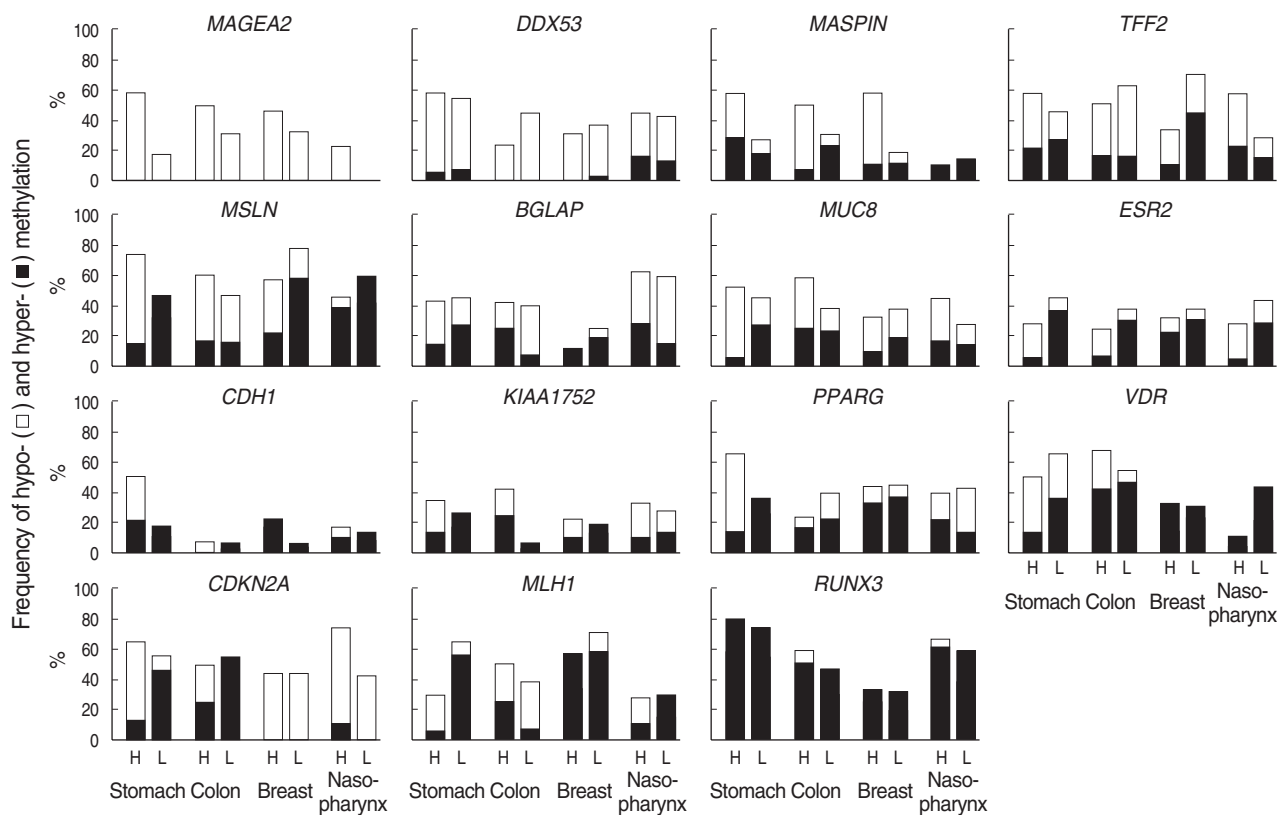
Fig. 2. Methylation changes in the transitional-CpG sites of the 15 selected genes examined in four cancer types. The methylation status of the transitional-CpG sites was estimated using a semiquantitative methylation-specific PCR method. The methylation changes were scored based on the difference in the level of methylation between the normal and tumor tissues. The frequency of methylation changes in each transitional area is indicated as a percentage in 25 cancer cases. The level of chromosomal losses was evaluated by PCR-based loss-of-heterozygosity analysis. The cancer tissues were grouped into high-level chromosomal losses (H) four or more chromosomes and low-level chromosomal losses (L) involving less than four chromosomes.

microsatellite markers on eight cancer-associated chromosomes. In order to count the substantial loss of a chromosome, a unilateral chromosomal loss was defined when two or more allelic losses were detected on a single chromosome in cancer tissue. Fig. 1 shows the frequency of individual chromosomal losses and the number of chromosomal losses examined in four cancer types. The losses of chromosomes 17p and 18q were most frequent in gastric (72% and 68%) and colonic cancers (64% and 56%). An 8p loss was most common in mammary cancers (60%) and 9p loss (72%) most frequent in nasopharyngeal cancers. The level of LOH was categorized as high level (involving four or more chromosomes) and low level (involving less than four chromosomes). High- and low-level chromosomal losses had a similar frequency in gastric (56% vs. 44%) and colonic (48% vs. 52%) cancers. Mammary and nasopharyngeal cancers frequently showed low-level (64%) and high-level (72%) chromosomal losses, respectively.

Methylation changes in the transitional-CpG sites

A total of 15 transitional-CpG sites selected from six CpG-island-negative genes (*MAGEA2*, *TFF2*, *DDX53*, *MSLN*, *MASPIN*, and *BGLAP*) and nine CpG-island-positive genes (*MUC8*, *KIAA1752*, *CDKN2A*, *ESR2*, *PPARG*, *MLH1*, *CDH1*, *VDR*, and *RUNX3*) were examined using MSP primer sets (Supplementary Fig. 1, Supplementary Table 1-3). The methylation density of the CpG amplicon was divided into the following 5 levels: 1 (0-20%), 2 (21-40%), 3 (41-60%), 4 (61-80%), and 5 (81-100%). The frequency of the methylation changes was scored as the number of the

CpG amplicons showing a difference in the level of methylation between the normal and tumor DNAs.

Fig. 2 shows the methylation status of the transitional-CpG sites examined in the four cancer types. The non-island CpG sites round the transcription start sites of the *MAGEA2*, *TFF2*, and *DDX53* genes lacking CpG islands were hypomethylated at various frequencies (16-48%). The CpG-island margins of the *VDR*, *MLH1*, and *RUNX3* genes bordered by retroelements at a long distance had a tendency for hypermethylation in various frequencies (16-76%). A total of 1,500 CpG amplicon pairs were obtained from 100 pairs of the normal and cancer DNAs using 15 methylation primer sets, of which 617 (41%) showed a similar frequency of hypermethylation (320, 20%) or hypomethylation (297, 21%) changes in the cancer DNA. Nine CpG-island margins (900 CpG amplicon pairs) and six non-island CpG sites (600 CpG amplicon pairs) were mainly hypermethylated (hyper- vs. hypo-methylation, 25% vs. 16%) and hypomethylated (hyper- vs. hypo-methylation, 16% vs. 26%) in cancer tissues, respectively ($p<0.0001$).

A total of 225 CpG amplicon pairs from nine CpG-island margins and a total of 150 CpG amplicon pairs from six non-island CpG sites were analyzed in 25 cancer tissues tested for each cancer type. Fig. 3 shows the relationships between the transitional-CpG methylation status and the level of LOH analyzed in the four cancer types. A comparison of the four cancer types showed that hypomethylation of the CpG-island margins tend to be more common in gastric (21%) and nasopharyngeal (18%) cancers than in colonic (13%), and mammary (10%) cancers ($p=0.009$). The hypomethylation of the non-island CpG sites was more frequent
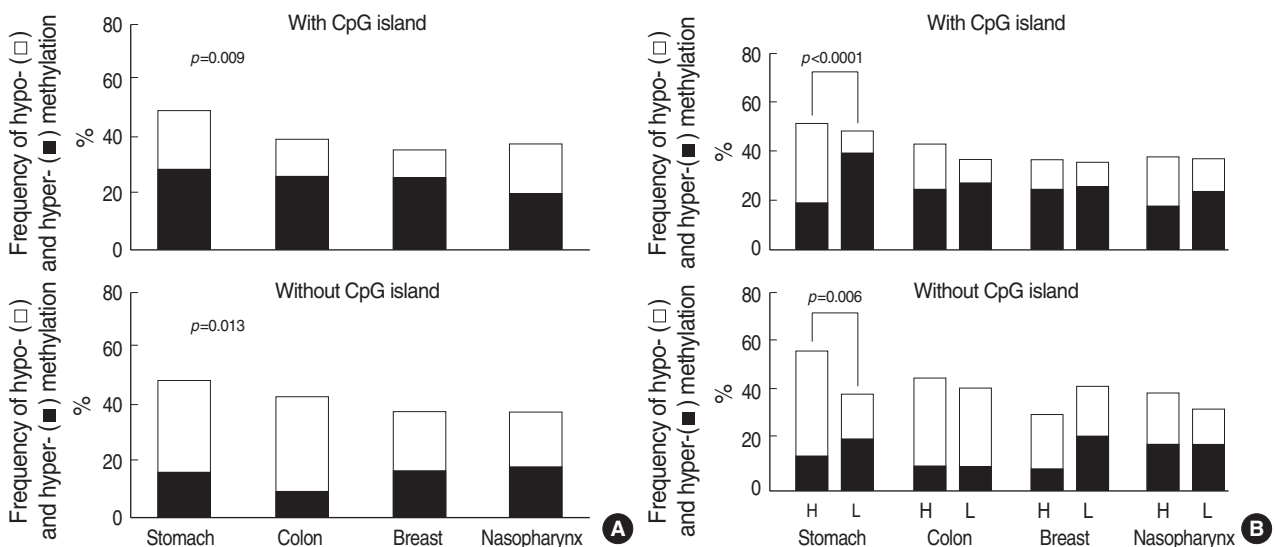
Fig. 3. Comparison of the transitional-CpG methylation changes between the different cancer types (A) and between cancers with high-level (H) and low-level (L) chromosomal losses (B). The criteria for the level of chromosomal losses are described in the legend of Fig. 2. The frequency of the methylation changes in the six non-island CpG sites and nine CpG-island margins are indicated as a percentage in 25 cancer cases. *p* values were calculated for the differences in the frequency of methylation changes between four cancer types by a chi-square test.

in the gastric (32%) and colonic (33%) cancers than in the mammary (21%) and nasopharyngeal (19%) cancers ($p=0.013$). Hypomethylation of the transitional-CpG sites frequent in gastric cancers was associated with high-level chromosomal losses (CpG-island margins, $p<0.0001$; non-island CpG sites, $p=0.006$).

Methylation status of transitional CpGs in normal somatic tissues

The methylation status of the 15 transitional-CpG sites, including the methylation data reported previously (9, 10), was analyzed comparatively in the 11 somatic tissue types (Fig. 4). The mean level of transitional-CpG methylation estimated in three or five individuals was calculated for each tissue type. The somatic tissues were classified into the three germ-layer lineages to determine the lineage-dependent pattern of transitional-CpG methylation. Eight of the nine CpG-island margins were completely unmethylated or most hypomethylated in the mesodermal lineage and all six non-island CpG sites were the most hypermethylated. Meanwhile, seven CpG-island margins were most hypermethylated in the endodermal lineage and three non-island CpG sites were the most hypomethylated. The CpG-island margins were slightly hypomethylated in the ectodermal lineage compared with the endodermal lineage.

Analysis of the transcript populations based on CpG islands and nearby retroelements

The number of active genes and the mean number of expressed tags (i.e. mean transcript number) per active gene were used to access the net output of transcription in each tissue (Fig. 5A). A total of 15,770 active genes obtained from the SAGE data were categorized according to the presence or absence of CpG islands. The number of active genes in both gene groups with and without CpG islands was higher in the embryonic stem cells than in the placenta. The transcript number per active gene in the gene group with CpG islands was higher in the embryonic stem cells while that in the gene group lacking CpG islands was higher in the placenta. The somatic tissues also showed an inverse correlation between the total number of active genes and the transcript number per active gene in the gene group lacking CpG islands. The genes with CpG islands showed a low number of active genes as well as the transcript number per active gene in the normal stomachs compared with the other normal tissues.
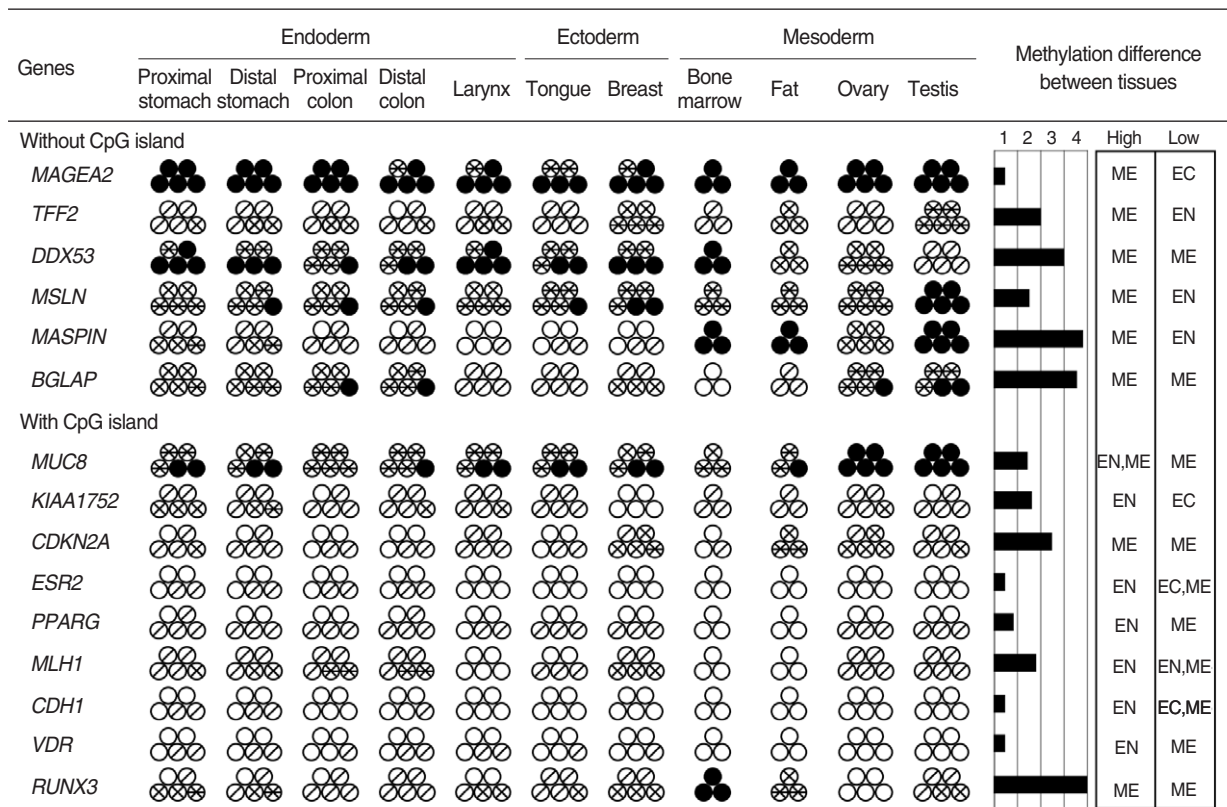
Fig. 4. Methylation profiles of 15 transitional CpG sites examined in 11 tissue types. The estimation of CpG methylation obtained using semiquantitative methylation-specific PCR were divided into five levels (○ , 0-20%; ◔ , 21-40%; ◒ , 41-60%; ◕ , 61-80%; ● , 81-100%). Differences in the mean level of methylation between the most highly methylated and least methylated tissue types are indicated by closed bars. Somatic tissues were classified into three germ-layer lineages, endoderm (EN), mesoderm (ME), and ectoderm (EC).

All cancers tended to show an increase in the number of active genes accompanying the down-regulation of tissue-specific strong gene expression. Gastric and colonic cancers increased the number of active genes with and without CpG islands. Gastric cancers increased the transcript number per active gene in the gene group containing CpG islands and decreased the transcript number in the gene group lacking CpG islands. Colonic cancers showed a lower transcript number per active gene in both the gene groups with and without CpG islands. Mammary cancers showed a slightly high-er number of the active genes in the gene group with CpG islands and a lower transcript number per active gene in the gene group lacking CpG islands.

The type of retroelements in close proximity to the promoter was demarcated using non-overlapping 3-kb windows moving away from the transcription start site. The CpG-island-positive genes were classified into three gene groups according to the type of retroelements occupying a 3-kb window as follows; L1, *Alu*, and L1-*Alu* combination. The CpG-island-negative genes with L1 or L1-*Alu* in a 3-kb window
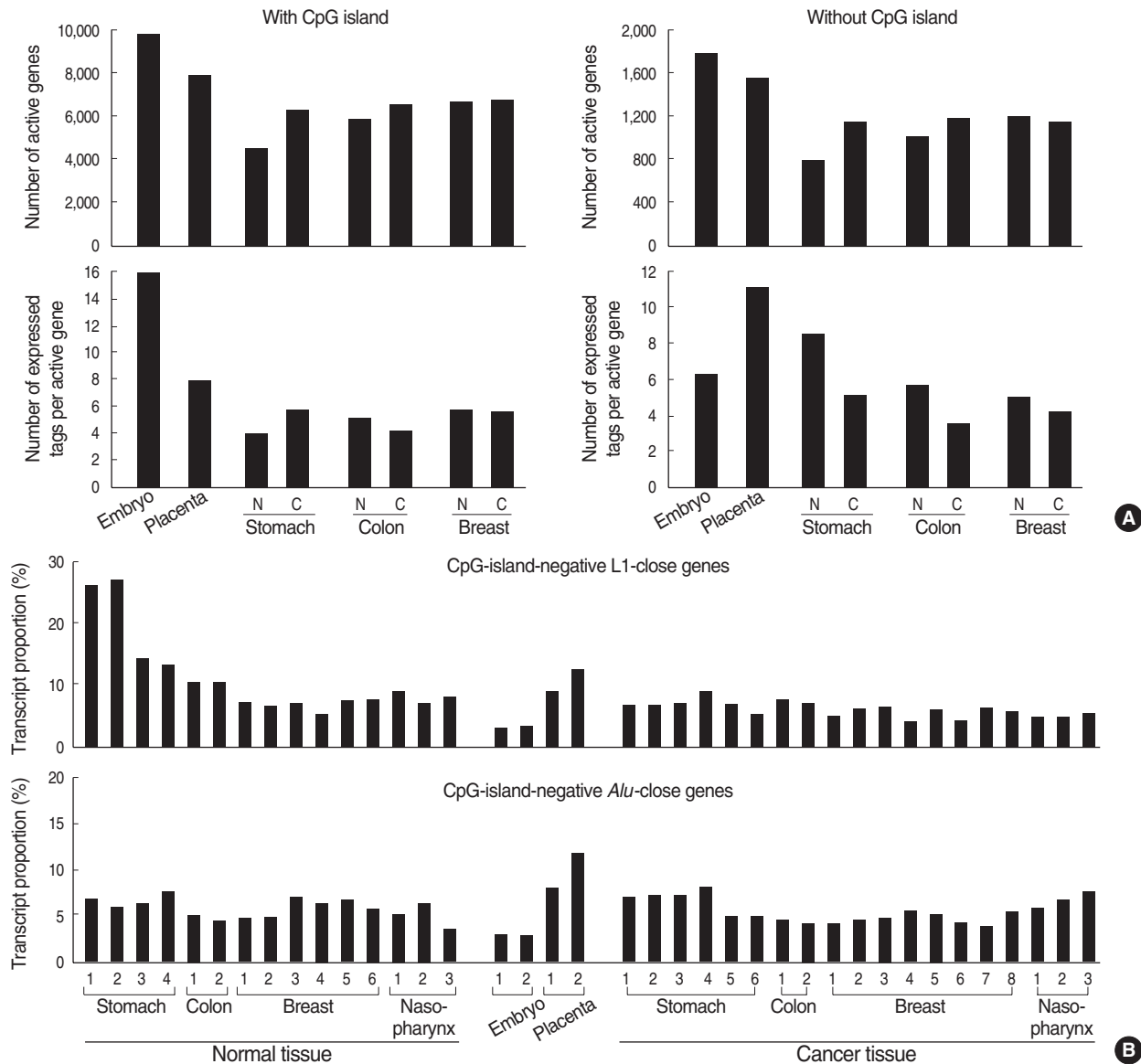
Fig. 5. Transcript populations of the embryonic stem cells, placenta, and somatic tissues in the SAGE libraries. (**A**) The number of the active genes and the number of tags expressed per active gene were calculated separately according to the presence or absence of CpG islands. The mean values of two to eight tissues are indicated for each tissue type. N, normal; C, cancer. All information about the SAGE data is listed in Supplementary Table 4. (**B**) Transcripts of the CpG-island-negative genes close to the L1 and *Alu* retroelements. The genes lacking the CpG islands were grouped according to the type of retroelements round the transcription start sites. The relative proportion of the gene-group transcripts in the total transcript population was calculated for each normal and cancer tissue. The transcript data of the nasopharynx was obtained from the EST libraries. All information regarding individual tissues is listed in Supplementary Table 4.

were collectively classified as the L1-close gene group. The number of active genes was high in order of the breast, colon, and stomach irrespective of the type of nearby retroelements and the presence or absence of CpG islands (Table 1). There was an inverse correlation between the number of active genes and the transcript number per active gene in the CpG-island-negative gene group close to L1 elements in normal tissues. There was a similar number of active genes and a similar transcript number per active gene in each gene group in all cancers.

SAGE or the expressed sequence tag (EST) data was used to calculate the relative proportion of CpG-island-negative gene transcripts in the total transcripts (Fig. 5B). The CpG-island-negative gene group close to the L1 elements was transcribed in the highest proportion (13-27%) in the stomach and in the lowest proportion in the embryonic stem cells (3-4%). In all cancers, the proportion of CpG-island-negative L1-close gene transcripts reached intermediate levels (4-9%) compared with the placenta (9-12%) and embryonic stem cells (3-4%).

Tissue-specific gene expression profiles in embryo and somatic tissues

The Pearson's correlation coefficients of the transcript numbers of the individual genes between the somatic tissues and either embryonic stem cells or the placenta were calculated to determine the similarities in the gene expression profiles (Table 2). The transcript numbers of the CpG-island-negative genes in the placenta and embryonic stem cells were not associated with those in the stomach but strongly or weakly associated with those in the colon and breast. There was no or weak association between the transcript numbers of the CpG-island-positive genes close to L1 elements in the gastrointestinal tissues and placenta or in the breast tissues and embryonic stem cells. In all cancer tissues, including gastrointestinal cancers, the transcript numbers of both the genes with and without CpG islands were strongly associated with those of the embryonic stem cells as well as the placenta.

DISCUSSION

In this study, we investigated the four types of human cancers-gastric, colonic, mammary, and nasopharyngeal cancers using LOH and MSP to delineate the relationship of chromosomal losses with the DNA methylation and transcription profiles in cancers. The four human cancers can be distinguished for their preferred types and extents of LOH (Fig. 1), providing potential clues on the presence of cancer-type-dependent tumor suppressor genes. With respect to methylation changes, there was a similar tendency for the 15 transitional CpG sites in the four cancers; the hypermethylation of nine CpG-island margins and the hypomethylation of six

non-island CpG sites (Fig. 2). A series of evidence proposes that retroelement methylation is prevented at the unmethylated CpG islands and promote the methylation of the non-island-CpG sites (1, 2, 17, 18). Both CpG-island margin hypermethylation and non-island CpG hypomethylation might be associated with the genome-wide pattern of retroelement methylation (6, 19). In addition, chromosomal losses reducing the active genes appear to influence the total number of active genes in the cancer genome through global methylation changes in the gene-control regions.

The hypomethylation of both the CpG-island margins and non-island-CpG sites was more common in gastric cancers than in other cancer types. The SAGE data demonstrated an increase in the number of active genes with and without CpG islands in gastric cancers and their expression profiles were similar to those of embryonic stem cells and the placenta (Table 1, 2). It is well-known that there are many similarities exist between carcinogenesis and embryogenesis, such as invasive growth and multi-lineage differentiation (20-23). The hypomethylation of the transitional-CpG sites in the gene-control regions may allow the reactivation of cell-intrinsic developmental programs that are repressed by transitional-CpG methylation in the adult somatic tissues.

In gastric cancer, the cases with high-level of chromosomal losses showed a significantly higher extent of hypomethylation for both the CpG-island margins and non-island CpG sites (Fig. 3). Meanwhile, the colon, breast, and nasopharynx with a higher number of active genes showed no significant association between the chromosomal losses and methylation changes in their cancer tissues. The different epigenetic and transcriptional changes in different cancer types are likely to commonly drive the universal malignant traits through the reactivation of the dormant cell-intrinsic programs for embryonic implantation and placentation (9, 12, 13). Therefore, the hypomethylation of the transitional-CpG sites associated with chromosomal losses might be uniquely prevalent in gastric cancers because the stomach with a small number of active genes needs an increase in the active genes for cancer evolution.

In the stomach, high-level chromosomal losses can cause the excessive reduction of active gene copies and have an adverse effect on cell viability. The hypomethylation of the transitional-CpG sites related to the global methylation pattern would increase the number of active genes as well as facilitate the reactivation of the inactive genes. In case of the normal colon with a large number of active genes, it is likely that the hypomethylation of the transitional-CpG sites in colonic cancers is not necessarily associated with high-level chromosomal losses. Similarly, the gene-dose-dependent methylation changes in both the genes with and without CpG islands appear to be relatively less dominant in mammary and nasopharyngeal cancers because their normal tissues already have a sufficient number of active genes.

When the 11 normal tissue types were categorized based

on the three germ-layer origins, the CpG-island margins, most of which are close to L1 elements or are bordered by retroelements at a long distance (Supplementary Fig. 1), were more methylated in the endoderm-derived tissues than in the other tissues (Fig. 4). In particular, the normal stomach had a lower number of active genes and higher transcript number per active genes in the CpG-island-negative gene group close to L1 elements. Previous studies suggest that during embryogenesis and development, highly repetitive *Alu* and L1 elements trigger genome-wide methylation as well as the spreading of methylation signals into the flanking CpGs (9, 24, 25). The dense methylation of the transitional-CpG markers indicates long-distance L1 methylation in the stomach, which is necessary for maintaining a small number of active genes as well as repressing a large number of genes (Table 3). Therefore, the tissue-specific methylation profiles established in consistent with the tissue-specific number of active genes are likely to be disturbed with chromosomal losses reducing the active genes in cancers.

With the accuracy of the SAGE data, the biased number of active genes is attributable to the total number of transcript tags counted in the SAGE library (26). However, both the SAGE and microarray data showed the strong expression of a small number of the active genes without CpG islands in the normal stomach (data not shown). The non-island-CpG sites were hypomethylated in the normal stomach whereas the hypermthylation of the CpG-island margins in the normal stomach (Fig. 4). This pattern of transitional-CpG methylation appears to be a useful epigenetic way for the maintenance of tissue-specific strong expression along with the inactivation of a large number of genes. Thus, it is possible that a set of the transitional-CpG sites examined in this study can serve as a global epigenetic marker for the total number of active genes established under the influence of genome-wide retroelement methylation.

Taken together, unilateral chromosomal losses, which reflect a reduction in the gene dose, might lead to the hypomethylation of the gene-control regions and an increase in the number of active genes in gastric cancers. The hypomethylation changes in the transitional CpG sites appear to facilitate the reactivation of cell-intrinsic embryogenesis-development programs in cancer cells as well. The results of this study explain how the LOH events initiate the invasive outgrowth of cancer cells similar to embryonic implantation and placentation.

REFERENCES

1. Walsh CP, Bestor TH. *Cytosine methylation and mammalian development. Genes Dev 1999; 13: 26-34.*

2. Gardiner-Garden M, Frommer M. *CpG islands in vertebrate genomes. J Mol Biol 1987; 196: 261-82.*

3. Larsen F, Gundersen G, Lopez R, Prydz H. *CpG islands as gene markers in the human genome. Genomics 1992; 13: 1095-107.*

4. 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.*

5. Jones PA. *The DNA methylation paradox. Trends Genet 1999; 15: 34-7.*

6. Feinberg AP, Ohlsson R, Henikoff S. *The epigenetic progenitor origin of human cancer. Nat Rev Genet 2006; 7: 21-33.*

7. Sieber OM, Heinimann K, Tomlinson IP. *Genomic instability--the engine of tumorigenesis? Nat Rev Cancer 2003; 3: 701-8.*

8. Turker MS. *Gene silencing in mammalian cells and the spread of DNA methylation. Oncogene 2002; 21: 5388-93.*

9. 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.*

10. Kang MI, Kim HS, Jung YC, Kim YH, Hong SJ, Kim MK, Baek KH, Kim CC, Rhyu MG. *Transitional CpG methylation between promoters and retroelements of tissue-specific genes during human mesenchymal cell differentiation. J Cell Biochem 2007; 102: 224-39.*

11. Hong SJ, Choi SW, Lee KH, Lee S, Min KO, Rhyu MG. *Preoperative genetic diagnosis of gastric carcinoma based on chromosomal loss and microsatellite instability. Int J Cancer 2005; 113: 249-58.*

12. 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.*

13. 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.*

14. Kim KM, Kwon MS, Hong SJ, Min KO, Seo EJ, Lee KY, Choi SW, Rhyu MG. *Genetic classification of intestinal-type and diffuse-type gastric cancers based on chromosomal loss and microsatellite instability. Virchows Arch 2003; 443: 491-500.*

15. Ishii M, Hashimoto S, Tsutsumi S, Wada Y, Matsushima K, Kodama T, Aburatani H. *Direct comparison of GeneChip and SAGE on the quantitative accuracy in transcript profiling analysis. Genomics 2000; 68: 136-43.*

16. van Ruissen F, Ruijter JM, Schaaf GJ, Asgharnegad L, Zwijnenburg DA, Kool M, Baas F. *Evaluation of the similarity of gene expression data estimated with SAGE and Affymetrix GeneChips. BMC Genomics 2005; 6: 91.*

17. Cross SH, Bird AP. *CpG islands and genes. Curr Opin Genet Dev 1995; 5: 309-14.*

18. Yoder JA, Walsh CP, Bestor TH. *Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 1997; 13: 335-40.*

19. Baylin SB, Belinsky SA, Herman JG. *Aberrant methylation of gene promoters in cancer---concepts, misconcepts, and promise. J Natl Cancer Inst 2000; 92: 1460-1.*

20. Gold P, Freedman SO. *Specific carcinoembryonic antigens of the human digestive system. J Exp Med 1965; 122: 467-81.*

21. Edynak EM, Old LJ, Vrana M, Lardis MP. *A fetal antigen associated with human neoplasia. N Engl J Med 1972; 286: 1178-83.*

22. Tanaka M, Sasaki H, Kino I, Sugimura T, Terada M. *Genes preferentially expressed in embryo stomach are predominantly expressed in gastric cancer. Cancer Res 1992; 52: 3372-7.*

23. Soundararajan R, Rao AJ. *Trophoblast 'pseudo-tumorigenesis': significance and contributory factors. Reprod Biol Endocrinol 2004; 2: 15.*

24. 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.*

25. Meunier J, Khelifi A, Navratil V, Duret L. *Homology-dependent methylation in primate repetitive DNA. Proc Natl Acad Sci USA 2005; 102: 5471-6.*

26. Sun M, Zhou G, Lee S, Chen J, Shi RZ, Wang SM. *SAGE is far more sensitive than EST for detecting low-abundance transcripts. BMC Genomics 2004; 5: 1.*

27. Fulka H, Mrazek M, Tepla O, Fulka J Jr. *DNA methylation pattern in human zygotes and developing embryos. Reproduction 2004; 128: 703-8.*

28. Razin A, Kafri T. *DNA methylation from embryo to adult. Prog Nucleic Acid Res Mol Biol 1994; 48: 53-81.*

29. Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, Gehrke C. *Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res 1982; 10: 2709-21.*

Supplementary Table 1 List of 15 genes examined by methylation analysis

Gene symbol	Gene description	Accession no.	Chromosome locus	MSP position (kb)	Methylation studies
MAGEA2	Melanoma antigen family A, 2	NM_005361	Xq28	0.0	(1)
TFF2	Trefoil factor 2 (spasmolytic protein 1)	NM_005423	21q22	-0.2	(2)
DDX53	DEAD (Asp-Glu-Ala-Asp) box polypeptide 53	NM_182699	Xq22.11	0.0	(1)
MSLN	Mesothelin isoform 2 precursor	NM_013404	16q13	-0.8	(2)
MASPIN	Serpin peptidase inhibitor, clade B (ovalbumin), member 5	NM_002639	18q21	-0.3	(1)
BGLAP	Osteocalcin	NM_199173	1q23.1	-0.4	(3)
MUC8	Mucin 8, tracheobronchial	U14383[a]	12q24	+0.2	(2)
KIAA1752	Homo sapiens mRNA for KIAA1752 protein	AB051539[a]	16q12	+0.4	(2)
CDKN2A	Cyclin-dependent kinase inhibitor 2A	NM_000077	9p21	-1.5	(1)
ESR2	Estrogen receptor 2 (ER beta)	NM_001437	14q23	-0.9	(2)
PPARG	Peroxisome proliferative activated receptor	NM_138711	3p25.1	-1.5	(3)
MLH1	MutL homolog 1, colon cancer, nonpolyposis type 2	NM_000249	3p22	-1.0	(1)
CDH1	Cadherin 1, type 1, E-cadherin (epithelial)	NM_004360	16q22	0.1	(1)
VDR	Vitamin D (1,25- dihydroxyvitamin D3) receptor	NM_000376	12q13	-0.7	(2)
RUNX3	Runt-related transcription factor 3	NM_004350	1p36	-1.7	(2)

[a], GenBank accession number.

1. 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.*

2. 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.*

3. Kang MI, Kim HS, Jung YC, Kim YH, Hong SJ, Kim MK, Baek KH, Kim CC, Rhyu MG. *Transitional CpG methylation between promoters and retroelements of tissue-specific genes during human mesenchymal cell differentiation. J Cell Biochem 2007 Mar 12.*

Supplementary Table 2 Sequences and PCR condition of the unmethylation (U) and methylation (M)-specific primer sets

CpG sites		Forward (5´ to 3´)	Reverse (5´ to 3´)	Amplicon size (bp)	Tm (℃)
MAGEA2	U	GTTAGGTTGTTGTTTAGGGT	CCAAAAAAATCACAAACCCA	92	59
	M	GCGTTTGTTTTTTTTCGTCGAC	AAATCACGAACCCGAATATAACG	108	61
TFF2	U	GGTAGTTGTGTTTTGTGTAGGT	CACATAACCAATTTTCCACA	130	56
	M	GGTAGTTGTGTTTTGTGTAGGC	CACGTAACCGATTTTCCACG	130	62
DDX53	U	TGGTTTTTGGGGTAATTTTTGT	CAAATCTACAACCTATTTCCCA	105	57
	M	TTTTATACGATTCGGAATTCGAC	CAAATCTACGACCTATTTCCCG	136	58
MSLN	U	GGAGAGATTAGAGATGATTGTTGT	CATAAACTCTTATCCCCAATACA	103	55
	M	GGAGAGATTAGAGATGATCGTCGC	CGTAAACTCTTATCCCCAATACG	103	60
MASPIN	U	GAATATTTTATTTTTTGGTTTTGTG	AAAAAACCTCCAACATATTCA	111	56
	M	TTATTTTTCGGTTTTGCG	AAAAAACCTCCAACATATTCG	104	54
BGLAP	U	AGGGTAGGGTTTGAGTTGTT	AATACCTCACAATACCCCCA	85	58
	M	AGGGTAGGGTTTGAGTCGTC	AATACCTCGCAATACCCCCG	85	58
MUC8	U	GGTAGGAGTTATTAGGAGAGTATT	AATACAAACACTCACCACCTAACC	140	55
	M	GGTAGGAGTTATTAGGAGAGTATC	AATACAAACGCTCACCGCCTAACC	140	60
KIAA1752	U	TAATGGTTTTTGAGGATTGAGATTG	CACAAACTATTATCAACCAATCAC	103	58
	M	TAATGGTTTTTGAGGATTGAGATC	CACAAACTATTATCAACCGATCAC	103	62
CDKN2A	U	TTGGGATTAGGTTTAGTTTTGG	CTATAAAACCCTATCAACTCACAC	130	58
	M	TCGGGATTAGGTTTAGTTTCG	AAACCCTATCGACTCACGCT	125	60
ESR2	U	TTTTTTTTAAGGATTTTGTGTGT	ACTAAAAATACACATTCCACCA	111	56
	M	TTTTTTTAAGGATTTCGCGCGC	CCAACTAAAAATACACGTTCCACC	113	58
PPARG	U	GGTTAGGTTTTGTGTTTTGATGT	CCTAACTACACACTCCATCCA	103	58
	M	GGTTAGGTTTTGTGTTTTGACGC	CCTAACTACGCGCTCCATCCG	103	56
MLH1	U	GATTTTAGGATTGTTGATATGAGT	AAACTACCTCCTAATCTTTATCCA	126	58
	M	GATTTTAGGATTGTCGATATGAGC	AACTACCTCCTAATCTTTATCCG	125	58
CDH	U	GGTGAATTTTTAGTTAATTAGTGGTAT	TCACAAATACTTTACAATTCCAAC	108	56
	M	TGAATTTTTAGTTAATTAGCGGTAC	ACAAATACTTTACAATTCCGACG	104	58
VDR	U	TGGTAGTGATTGTGGTTGATTAT	CCTCACACCAATACCACAAAACA	130	58
	M	GGTAGCGATCGCGGTTGATTAC	CTCACGCCGATACCACGAAACG	128	58
RUNX3	U	TGGGGTTAGATTTTTGTTGTTTTT	ATAAAATCTTACAACCACCATCA	107	56
	M	CGGGGTTAGATTTTCGTTGTTTTC	ATAAAATCTTACGACCACCGTCG	107	58

Supplementary Table 3 Results of LOH and MSP analyses on 100 patients

Tissue	Sample ID	Sex	Age	Genotype[a]	Analysis of loss of heterozygosity[†]							
					No. of LOH	Chromosome 3	Chromosome 4	Chromosome 5	Chromosome 8	Chromosome 9	Chromosome 13	Chromosome 17	Chromosome 18
Stomach	S1	Male	57	LOH-H	8	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	S2	Male	65	LOH-H	7	LOH+	LOH+	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+
	S3	Female	35	LOH-H	7	LOH+	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+	LOH+
	S4	Male	65	LOH-H	7	LOH+	LOH+	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+
	S5	Male	60	LOH-H	6	LOH-	LOH+	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+
	S6	Female	64	LOH-H	6	LOH+	LOH+	LOH-	LOH-	LOH+	LOH+	LOH+	LOH+
	S7	Male	74	LOH-H	6	LOH-	LOH+	LOH+	LOH+	LOH-	LOH+	LOH+	LOH+
	S8	Female	57	LOH-H	5	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+	LOH-	LOH-
	S9	Male	63	LOH-H	5	LOH-	LOH-	LOH+	LOH+	LOH+	LOH-	LOH+	LOH+
	S10	Male	67	LOH-H	5	LOH-	LOH-	LOH+	LOH+	LOH+	LOH-	LOH+	LOH+
	S11	Male	75	LOH-H	4	LOH+	LOH+	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+
	S12	Female	74	LOH-H	4	LOH-	LOH+	LOH+	LOH-	LOH-	LOH-	LOH+	LOH+
	S13	Male	76	LOH-H	4	LOH-	LOH-	LOH-	LOH+	LOH+	LOH-	LOH+	LOH+
	S14	Male	61	LOH-H	4	LOH-	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+	LOH-
	S15	Male	64	LOH-L	3	LOH+	LOH-	LOH-	LOH-	LOH-	LOH+	LOH-	LOH+
	S16	Male	49	LOH-L	3	LOH+	LOH-	LOH-	LOH+	LOH-	LOH-	LOH+	LOH-
	S17	Male	57	LOH-L	3	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-	LOH+	LOH+
	S18	Male	48	LOH-L	3	LOH-	LOH+	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+
	S19	Male	64	LOH-L	2	LOH-	LOH-	LOH-	LOH+	LOH+	LOH-	LOH-	LOH-
	S20	Female	66	LOH-L	2	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+	LOH-
	S21	Male	62	LOH-L	2	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+	LOH-
	S22	Male	64	LOH-L	2	LOH-	LOH-	LOH+	LOH-	LOH+	LOH-	LOH-	LOH-
	S23	Female	42	LOH-L	1	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+
	S24	Male	59	LOH-L	1	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+
	S25	Male	68	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
Colon	C1	Male	71	LOH-H	8	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	C2	Male	71	LOH-H	7	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	C3	Male	44	LOH-H	6	LOH+	LOH+	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+
	C4	Female	53	LOH-H	6	LOH+	LOH+	LOH+	LOH+	LOH-	LOH-	LOH+	LOH+
	C5	Male	61	LOH-H	5	LOH+	LOH+	LOH-	LOH+	LOH+	LOH-	LOH+	LOH-
	C6	Female	60	LOH-H	5	LOH+	LOH+	LOH-	LOH+	LOH-	LOH+	LOH+	LOH-
	C7	Female	40	LOH-H	5	LOH-	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+
	C8	Male	58	LOH-H	5	LOH-	LOH-	LOH-	LOH+	LOH+	LOH+	LOH+	LOH+
	C9	Male	67	LOH-H	4	LOH-	LOH-	LOH+	LOH+	LOH-	LOH-	LOH+	LOH+
	C10	Female	59	LOH-H	4	LOH+	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+	LOH+
	C11	Male	67	LOH-H	4	LOH-	LOH-	LOH+	LOH+	LOH-	LOH-	LOH+	LOH+
	C12	Female	73	LOH-H	4	LOH-	LOH+	LOH-	LOH-	LOH+	LOH-	LOH+	LOH+
	C13	Female	56	LOH-L	3	LOH-	LOH-	LOH+	LOH+	LOH+	LOH-	LOH-	LOH-
	C14	Male	63	LOH-L	3	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-	LOH+	LOH-
	C15	Male	64	LOH-L	3	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+	LOH+
	C16	Female	70	LOH-L	2	LOH+	LOH-	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-
	C17	Female	75	LOH-L	2	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+
	C18	Female	78	LOH-L	2	LOH-	LOH+	LOH-	LOH+	LOH-	LOH-	LOH-	LOH-
	C19	Female	29	LOH-L	2	LOH-	LOH-	LOH+	LOH-	LOH-	LOH+	LOH-	LOH-
	C20	Male	30	LOH-L	2	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+
	C21	Female	64	LOH-L	1	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+	LOH-	LOH-
	C22	Male	75	LOH-L	1	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH+
	C23	Female	46	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
	C24	Female	46	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
	C25	Female	75	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
Breast	B1	Female	41	LOH-H	7	LOH-	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	B2	Female	64	LOH-H	7	LOH-	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	B3	Female	47	LOH-H	6	LOH-	LOH+	LOH+	LOH+	LOH-	LOH+	LOH+	LOH+
	B4	Female	67	LOH-H	5	LOH-	LOH+	LOH+	LOH+	LOH+	LOH+	LOH-	LOH-
	B5	Female	50	LOH-H	5	LOH+	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+	LOH-

(Continued to the next page)

Supplementary Table 3 *(Continued from the previous page)* Results of LOH and MSP analyses on 100 patients

Tissue	Sample ID	Sex	Age	Genotype[a]	Analysis of loss of heterozygosity[†]								
					No. of LOH	Chromosome 3	Chromosome 4	Chromosome 5	Chromosome 8	Chromosome 9	Chromosome 13	Chromosome 17	Chromosome 18
	B6	Female	67	LOH-H	4	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+	LOH-	LOH-
	B7	Female	59	LOH-H	4	LOH-	LOH-	LOH+	LOH+	LOH-	LOH+	LOH+	LOH-
	B8	Female	31	LOH-H	4	LOH+	LOH+	LOH-	LOH-	LOH-	LOH+	LOH+	LOH-
	B9	Female	37	LOH-H	4	LOH-	LOH-	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+
	B10	Female	56	LOH-L	3	LOH+	LOH-	LOH-	LOH+	LOH-	LOH-	LOH+	LOH-
	B11	Female	55	LOH-L	3	LOH-	LOH-	LOH-	LOH+	LOH-	LOH+	LOH+	LOH-
	B12	Female	51	LOH-L	3	LOH-	LOH+	LOH-	LOH+	LOH-	LOH-	LOH+	LOH-
	B13	Female	33	LOH-L	3	LOH+	LOH-	LOH-	LOH+	LOH-	LOH-	LOH+	LOH-
	B14	Female	60	LOH-L	3	LOH+	LOH+	LOH-	LOH-	LOH-	LOH+	LOH-	LOH-
	B15	Female	56	LOH-L	2	LOH+	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-	LOH-
	B16	Female	52	LOH-L	2	LOH-	LOH-	LOH-	LOH+	LOH-	LOH+	LOH-	LOH-
	B17	Female	50	LOH-L	2	LOH+	LOH-	LOH-	LOH-	LOH-	LOH+	LOH-	LOH-
	B18	Female	42	LOH-L	2	LOH+	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-	LOH-
	B19	Female	55	LOH-L	1	LOH-	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-	LOH-
	B20	Female	39	LOH-L	1	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-	LOH-	LOH-
	B21	Female	44	LOH-L	1	LOH+	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
	B22	Female	50	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
	B23	Female	43	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
	B24	Female	63	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
	B25	Female	49	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
Nasopharynx	N1	Male	56	LOH-H	8	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	N2	Male	48	LOH-H	8	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	N3	Male	55	LOH-H	8	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	N4	Male	71	LOH-H	8	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	N5	Male	72	LOH-H	7	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH-
	N6	Male	50	LOH-H	7	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+
	N7	Female	71	LOH-H	6	LOH-	LOH+	LOH+	LOH+	LOH-	LOH+	LOH+	LOH+
	N8	Male	56	LOH-H	6	LOH+	LOH+	LOH+	LOH+	LOH-	LOH-	LOH+	LOH+
	N9	Male	62	LOH-H	6	LOH+	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+	LOH-
	N10	Male	65	LOH-H	6	LOH+	LOH+	LOH+	LOH+	LOH+	LOH+	LOH-	LOH-
	N11	Male	48	LOH-H	6	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+	LOH-	LOH+
	N12	Male	65	LOH-H	6	LOH+	LOH+	LOH+	LOH+	LOH+	LOH-	LOH+	LOH-
	N13	Male	66	LOH-H	6	LOH+	LOH+	LOH+	LOH+	LOH+	LOH-	LOH-	LOH+
	N14	Male	58	LOH-H	5	LOH-	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+
	N15	Male	64	LOH-H	5	LOH+	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+
	N16	Male	47	LOH-H	5	LOH+	LOH-	LOH-	LOH+	LOH+	LOH+	LOH+	LOH-
	N17	Male	67	LOH-H	5	LOH-	LOH-	LOH+	LOH-	LOH+	LOH+	LOH+	LOH+
	N18	Male	51	LOH-H	4	LOH+	LOH-	LOH-	LOH-	LOH-	LOH+	LOH+	LOH+
	N19	Female	62	LOH-L	3	LOH-	LOH-	LOH-	LOH+	LOH+	LOH-	LOH+	LOH-
	N20	Female	45	LOH-L	3	LOH-	LOH-	LOH-	LOH+	LOH+	LOH+	LOH-	LOH-
	N21	Male	46	LOH-L	2	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-	LOH-	LOH+
	N22	Male	47	LOH-L	2	LOH+	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-	LOH-
	N23	Female	59	LOH-L	1	LOH-	LOH-	LOH-	LOH-	LOH+	LOH-	LOH-	LOH-
	N24	Male	44	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-
	N25	Female	74	LOH-L	0	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-	LOH-

Tissue	ID	Result of methylation-specific PCR[‡]													
		Type	MAGEA2	TFF2	DDX53	MSLN	MASPIN	BGLAP	MUC8	KIAA1752	CDKN2A	ESR2	PPARG	MLH1	CDH1	VDR	RUNX3
Stomach	S1	Normal	Level 5	Level 2	Level 5	Level 4	Level 3	Level 3	Level 4	Level 2	Level 2	Level 1	Level 2	Level 1	Level 1	Level 1	Level 2
		Cancer	Level 5	Level 2	Level 4	Level 5	Level 4	Level 3	Level 5	Level 2	Level 1	Level 1	Level 1	Level 1	Level 1	Level 2	Level 3
	S2	Normal	Level 5	Level 3	Level 4	Level 4	Level 2	Level 3	Level 4	Level 2	Level 2	Level 2	Level 2	Level 2	Level 1	Level 1	Level 2
		Cancer	Level 2	Level 2	Level 3	Level 3	Level 2	Level 3	Level 3	Level 3	Level 2	Level 1	Level 1	Level 2	Level 1	Level 1	Level 4
	S3	Normal	Level 5	Level 2	Level 4	Level 4	Level 4	Level 4	Level 5	Level 2	Level 2	Level 1	Level 1	Level 2	Level 2	Level 2	Level 2
		Cancer	Level 4	Level 3	Level 4	Level 2	Level 2	Level 4	Level 3	Level 3	Level 1	Level 1	Level 1	Level 2	Level 1	Level 1	Level 4

Supplementary Table 3 *(Continued from the previous page)* Results of LOH and MSP analyses on 100 patients

Tissue	ID	Type	MAGEA2	TFF2	DDX53	MSLN	MASP-IN	BGLAP	MUC8	KIAA1752	CDKN-2A	ESR2	PPARG	MLH1	CDH1	VDR	RUNX3
	S4	Normal	Level 5	Level 2	Level 5	Level 3	Level 2	Level 3	Level 4	Level 2	Level 2	Level 2	Level 2	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 4	Level 2	Level 4	Level 3	Level 3	Level 2	Level 4	Level 3	Level 2	Level 2	Level 1	Level 2	Level 2	Level 1	Level 4
	S5	Normal	Level 5	Level 2	Level 4	Level 5	Level 2	Level 3	Level 4	Level 2	Level 2	Level 2	Level 1	Level 2	Level 2	Level 1	Level 1
		Cancer	Level 2	Level 2	Level 2	Level 3	Level 2	Level 2	Level 4	Level 1	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
	S6	Normal	Level 5	Level 3	Level 4	Level 4	Level 2	Level 3	Level 5	Level 1	Level 2	Level 1	Level 2	Level 2	Level 2	Level 1	Level 2
		Cancer	Level 3	Level 3	Level 3	Level 3	Level 2	Level 4	Level 4	Level 1	Level 2	Level 1	Level 2	Level 2	Level 2	Level 1	Level 3
	S7	Normal	Level 5	Level 2	Level 4	Level 4	Level 4	Level 3	Level 4	Level 2	Level 2	Level 1	Level 2	Level 4	Level 1	Level 2	Level 1
		Cancer	Level 5	Level 2	Level 4	Level 4	Level 2	Level 2	Level 4	Level 2	Level 2	Level 1	Level 1	Level 4	Level 1	Level 2	Level 2
	S8	Normal	Level 5	Level 5	Level 4	Level 4	Level 4	Level 4	Level 5	Level 2	Level 2	Level 1	Level 2	Level 3	Level 2	Level 2	Level 2
		Cancer	Level 2	Level 4	Level 4	Level 5	Level 2	Level 3	Level 5	Level 2	Level 1	Level 1	Level 2	Level 3	Level 1	Level 1	Level 2
	S9	Normal	Level 4	Level 2	Level 4	Level 4	Level 2	Level 3	Level 5	Level 2	Level 2	Level 3	Level 2	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 1	Level 2	Level 5	Level 4	Level 2	Level 3	Level 2	Level 2	Level 1	Level 2	Level 1	Level 2	Level 2	Level 1	Level 4
	S10	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 3	Level 4	Level 2	Level 2	Level 1	Level 1	Level 4	Level 2	Level 1	Level 2
		Cancer	Level 5	Level 1	Level 5	Level 3	Level 1	Level 3	Level 4	Level 2	Level 2	Level 1	Level 1	Level 3	Level 2	Level 1	Level 5
	S11	Normal	Level 5	Level 3	Level 5	Level 4	Level 2	Level 3	Level 5	Level 3	Level 2	Level 1	Level 2	Level 2	Level 2	Level 1	Level 2
		Cancer	Level 3	Level 2	Level 3	Level 3	Level 2	Level 4	Level 4	Level 3	Level 3	Level 1	Level 1	Level 2	Level 2	Level 2	Level 3
	S12	Normal	Level 5	Level 3	Level 4	Level 4	Level 2	Level 3	Level 5	Level 2	Level 2	Level 2	Level 2	Level 2	Level 3	Level 2	Level 2
		Cancer	Level 5	Level 2	Level 4	Level 3	Level 2	Level 3	Level 5	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 1	Level 3
	S13	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 4	Level 4	Level 2	Level 3	Level 2	Level 2	Level 4	Level 1	Level 1	Level 2
		Cancer	Level 5	Level 3	Level 4	Level 4	Level 3	Level 4	Level 3	Level 2	Level 2	Level 2	Level 2	Level 1	Level 1	Level 1	Level 2
	S14	Normal	Level 5	Level 2	Level 4	Level 5	Level 2	Level 3	Level 5	Level 2	Level 1	Level 1	Level 2	Level 3	Level 2	Level 1	Level 4
		Cancer	Level 5	Level 3	Level 4	Level 4	Level 3	Level 3	Level 5	Level 1	Level 2	Level 2	Level 3	Level 3	Level 2	Level 1	Level 4
	S15	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 3	Level 4	Level 2	Level 3	Level 1	Level 1	Level 1	Level 1	Level 2	Level 2
		Cancer	Level 5	Level 1	Level 5	Level 5	Level 3	Level 3	Level 4	Level 3	Level 5	Level 1	Level 1	Level 1	Level 2	Level 1	Level 4
	S16	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 3	Level 5	Level 2	Level 2	Level 1	Level 2	Level 2	Level 2	Level 1	Level 2
		Cancer	Level 5	Level 2	Level 2	Level 5	Level 2	Level 3	Level 4	Level 3	Level 2	Level 1	Level 2	Level 1	Level 2	Level 1	Level 3
	S17	Normal	Level 5	Level 2	Level 5	Level 3	Level 2	Level 4	Level 4	Level 2	Level 2	Level 1	Level 1	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 5	Level 2	Level 4	Level 3	Level 2	Level 4	Level 4	Level 2	Level 3	Level 2	Level 2	Level 2	Level 1	Level 3	Level 4
	S18	Normal	Level 5	Level 2	Level 4	Level 4	Level 2	Level 3	Level 4	Level 2	Level 2	Level 1	Level 2	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 5	Level 2	Level 4	Level 4	Level 2	Level 4	Level 5	Level 2	Level 1	Level 2	Level 3	Level 3	Level 1	Level 2	Level 2
	S19	Normal	Level 4	Level 2	Level 4	Level 2	Level 1	Level 3	Level 4	Level 2	Level 1	Level 2	Level 2	Level 2	Level 2	Level 1	Level 4
		Cancer	Level 2	Level 2	Level 5	Level 2	Level 2	Level 3	Level 4	Level 2	Level 2	Level 2	Level 2	Level 4	Level 2	Level 2	Level 5
	S20	Normal	Level 5	Level 2	Level 4	Level 4	Level 2	Level 3	Level 5	Level 1	Level 2	Level 2	Level 1	Level 2	Level 1	Level 2	Level 4
		Cancer	Level 4	Level 3	Level 2	Level 4	Level 2	Level 4	Level 5	Level 1	Level 2	Level 2	Level 1	Level 2	Level 2	Level 2	Level 4
	S21	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 3	Level 4	Level 3	Level 2	Level 1	Level 1	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 5	Level 3	Level 2	Level 5	Level 2	Level 3	Level 5	Level 4	Level 5	Level 2	Level 1	Level 4	Level 1	Level 3	Level 4
	S22	Normal	Level 5	Level 2	Level 5	Level 2	Level 2	Level 3	Level 4	Level 2	Level 2	Level 1	Level 1	Level 3	Level 2	Level 2	Level 4
		Cancer	Level 5	Level 2	Level 5	Level 2	Level 2	Level 2	Level 4	Level 2	Level 2	Level 1	Level 1	Level 4	Level 2	Level 1	Level 5
	S23	Normal	Level 5	Level 3	Level 5	Level 3	Level 3	Level 3	Level 4	Level 2	Level 2	Level 2	Level 1	Level 4	Level 1	Level 1	Level 2
		Cancer	Level 5	Level 2	Level 5	Level 3	Level 3	BGLAP	MUC8	Level 2	Level 2	ESR2	PPARG	Level 4	CDH1	Level 1	VDR Level 2
	S24	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 3	Level 5	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 3	Level 2
		Cancer	Level 5	Level 3	Level 4	Level 5	Level 2	Level 4	Level 4	Level 2	Level 4	Level 2	Level 3	Level 4	Level 1	Level 1	Level 4
	S25	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 4	Level 4	Level 1	Level 2	Level 3	Level 1	Level 4	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 5	Level 5	Level 1	Level 4	Level 5	Level 1	Level 2	Level 2	Level 2	Level 1	Level 5	Level 1	Level 4
Colon	C1	Normal	Level 5	Level 1	Level 4	Level 5	Level 2	Level 4	Level 5	Level 3	Level 1	Level 2	Level 1	Level 4	Level 1	Level 3	Level 2
		Cancer	Level 2	Level 3	Level 4	Level 5	Level 1	Level 4	Level 3	Level 2	Level 1	Level 2	Level 1	Level 2	Level 1	Level 2	Level 4
	C2	Normal	Level 5	Level 2	Level 5	Level 4	Level 1	Level 4	Level 5	Level 2	Level 2	Level 3	Level 1	Level 3	Level 1	Level 2	Level 2
		Cancer	Level 4	Level 2	Level 5	Level 2	Level 1	Level 4	Level 5	Level 3	Level 2	Level 2	Level 1	Level 4	Level 1	Level 2	Level 3
	C3	Normal	Level 5	Level 2	Level 5	Level 5	Level 2	Level 4	Level 4	Level 2	Level 1	Level 2	Level 2	Level 4	Level 1	Level 3	Level 2
		Cancer	Level 3	Level 2	Level 5	Level 3	Level 2	Level 3	Level 4	Level 2	Level 1	Level 2	Level 2	Level 4	Level 1	Level 1	Level 2
	C4	Normal	Level 4	Level 2	Level 4	Level 4	Level 2	Level 4	Level 4	Level 1	Level 2	Level 2	Level 2	Level 2	Level 1	Level 2	Level 1
		Cancer	Level 4	Level 2	Level 3	Level 4	Level 1	Level 4	Level 4	Level 1	Level 1	Level 3	Level 2	Level 1	Level 1	Level 4	Level 2
	C5	Normal	Level 5	Level 3	Level 4	Level 5	Level 1	Level 4	Level 4	Level 2	Level 2	Level 2	Level 1	Level 3	Level 2	Level 2	Level 1
		Cancer	Level 2	Level 2	Level 3	Level 5	Level 2	Level 4	Level 5	Level 2	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 1
	C6	Normal	Level 5	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 2	Level 1	Level 2	Level 2	Level 3	Level 1	Level 2	Level 1

Supplementary Table 3 *(Continued from the previous page)* Results of LOH and MSP analyses on 100 patients

Tissue	ID	Type	MAGEA2	TFF2	DDX53	MSLN	MASP-IN	BGLAP	MUC8	KIAA1752	CDKN-2A	ESR2	PPARG	MLH1	CDH1	VDR	RUNX3
		Cancer	Level 3	Level 2	Level 4	Level 5	Level 2	Level 3	Level 3	Level 2	Level 1	Level 2	Level 2	Level 3	Level 1	Level 2	Level 2
	C7	Normal	Level 4	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 2	Level 1	Level 1	Level 3	Level 3	Level 1	Level 1	Level 2
		Cancer	Level 4	Level 2	Level 4	Level 5	Level 1	Level 4	Level 4	Level 2	Level 1	Level 1	Level 1	Level 3	Level 1	Level 1	Level 2
	C8	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 4	Level 4	Level 2	Level 1	Level 1	Level 1	Level 2	Level 1	Level 1	Level 2
		Cancer	Level 5	Level 2	Level 5	Level 5	Level 2	Level 5	Level 5	Level 3	Level 1	Level 1	Level 1	Level 4	Level 1	Level 2	Level 1
	C9	Normal	Level 5	Level 2	Level 4	Level 5	Level 3	Level 4	Level 4	Level 2	Level 2	Level 1	Level 1	Level 3	Level 1	Level 1	Level 2
		Cancer	Level 2	Level 1	Level 2	Level 4	Level 3	Level 4	Level 4	Level 3	Level 3	Level 1	Level 2	Level 4	Level 1	Level 2	Level 4
	C10	Normal	Level 5	Level 2	Level 4	Level 5	Level 2	Level 4	Level 4	Level 2	Level 1	Level 1	Level 2	Level 4	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 3	Level 4	Level 3	Level 2	Level 4	Level 3	Level 2	Level 1	Level 2	Level 2	Level 4	Level 1	Level 2	Level 3
	C11	Normal	Level 5	Level 1	Level 5	Level 4	Level 2	Level 4	Level 5	Level 2	Level 2	Level 1	Level 2	Level 4	Level 1	Level 2	Level 2
		Cancer	Level 5	Level 1	Level 5	Level 2	Level 1	Level 3	Level 4	Level 1	Level 1	Level 1	Level 2	Level 4	Level 1	Level 2	Level 2
	C12	Normal	Level 5	Level 3	Level 4	Level 4	Level 2	Level 3	Level 4	Level 2	Level 2	Level 1	Level 1	Level 4	Level 1	Level 1	Level 2
		Cancer	Level 5	Level 3	Level 4	Level 5	Level 1	Level 4	Level 5	Level 2	Level 4	Level 1	Level 1	Level 4	Level 1	Level 2	Level 3
	C13	Normal	Level 5	Level 4	Level 4	Level 4	Level 2	Level 3	Level 4	Level 2	Level 2	Level 1	Level 2	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 4	Level 3	Level 4	Level 4	Level 2	Level 3	Level 4	Level 2	Level 2	Level 2	Level 2	Level 2	Level 1	Level 2	Level 3
	C14	Normal	Level 5	Level 2	Level 3	Level 4	Level 2	Level 4	Level 5	Level 2	Level 1	Level 2	Level 1	Level 3	Level 1	Level 2	Level 2
		Cancer	Level 5	Level 1	Level 2	Level 5	Level 2	Level 3	Level 4	Level 2	Level 2	Level 2	Level 2	Level 3	Level 1	Level 1	Level 4
	C15	Normal	Level 5	Level 2	Level 5	Level 4	Level 1	Level 4	Level 4	Level 2	Level 2	Level 2	Level 1	Level 4	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 5	Level 3	Level 1	Level 4	Level 4	Level 2	Level 2	Level 2	Level 1	Level 4	Level 1	Level 2	Level 3
	C16	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 3	Level 4	Level 2	Level 1	Level 1	Level 1	Level 3	Level 1	Level 1	Level 2
		Cancer	Level 3	Level 2	Level 2	Level 4	Level 2	Level 3	Level 5	Level 2	Level 5	Level 1	Level 2	Level 3	Level 1	Level 4	Level 2
	C17	Normal	Level 5	Level 3	Level 4	Level 4	Level 1	Level 4	Level 4	Level 1	Level 2	Level 1	Level 2	Level 4	Level 1	Level 1	Level 2
		Cancer	Level 4	Level 3	Level 3	Level 3	Level 2	Level 4	Level 4	Level 1	Level 2	Level 2	Level 2	Level 2	Level 2	Level 2	Level 2
	C18	Normal	Level 4	Level 3	Level 5	Level 4	Level 1	Level 4	Level 4	Level 2	Level 1	Level 1	Level 1	Level 4	Level 1	Level 2	Level 2
		Cancer	Level 4	Level 2	Level 4	Level 5	Level 2	Level 3	Level 5	Level 2	Level 1	Level 1	Level 1	Level 2	Level 1	Level 2	Level 2
	C19	Normal	Level 5	Level 4	Level 5	Level 4	Level 2	Level 4	Level 4	Level 2	Level 1	Level 1	Level 2	Level 3	Level 1	Level 2	Level 1
		Cancer	Level 5	Level 4	Level 4	Level 4	Level 2	Level 3	Level 4	Level 2	Level 2	Level 2	Level 2	Level 3	Level 1	Level 3	Level 2
	C20	Normal	Level 5	Level 3	Level 5	Level 5	Level 2	Level 4	Level 4	Level 1	Level 1	Level 1	Level 1	Level 2	Level 2	Level 1	Level 2
		Cancer	Level 5	Level 3	Level 5	Level 3	Level 1	Level 4	Level 4	Level 2	Level 2	Level 1	Level 1	Level 2	Level 2	Level 1	Level 2
	C21	Normal	Level 5	Level 2	Level 4	Level 4	Level 2	Level 4	Level 4	Level 2	Level 1	Level 1	Level 2	Level 5	Level 2	Level 1	Level 2
		Cancer	Level 5	Level 1	Level 4	Level 2	Level 2	Level 3	Level 4	Level 2	Level 2	Level 3	Level 1	Level 4	Level 2	Level 5	Level 2
	C22	Normal	Level 4	Level 4	Level 4	Level 5	Level 2	Level 4	Level 4	Level 2	Level 2	Level 1	Level 2	Level 4	Level 1	Level 1	Level 2
		Cancer	Level 4	Level 3	Level 4	Level 5	Level 2	Level 5	Level 3	Level 2	Level 2	Level 1	Level 3	Level 4	Level 1	Level 2	Level 3
	C23	Normal	Level 5	Level 3	Level 5	Level 5	Level 1	Level 3	Level 4	Level 2	Level 1	Level 2	Level 3	Level 4	Level 1	Level 2	Level 2
		Cancer	Level 4	Level 2	Level 5	Level 5	Level 2	Level 3	Level 4	Level 2	Level 2	Level 2	Level 2	Level 4	Level 1	Level 2	Level 3
	C24	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 4	Level 4	Level 2	Level 1	Level 3	Level 1	Level 4	Level 1	Level 1	Level 2
		Cancer	Level 5	Level 3	Level 4	Level 4	Level 2	Level 4	Level 5	Level 2	Level 2	Level 1	Level 1	Level 5	Level 1	Level 1	Level 3
	C25	Normal	Level 5	Level 3	Level 4	Level 4	Level 1	Level 4	Level 4	Level 2	Level 2	Level 1	Level 2	Level 4	Level 2	Level 2	Level 2
		Cancer	Level 5	Level 4	Level 4	Level 4	Level 1	Level 4	Level 4	Level 2	Level 2	Level 1	Level 2	Level 2	Level 2	Level 2	Level 2
Breast	B1	Normal	Level 5	Level 3	Level 5	Level 4	Level 1	Level 3	Level 4	Level 2	Level 3	Level 1	Level 2	Level 1	Level 1	Level 1	Level 4
		Cancer	Level 5	Level 2	Level 4	Level 3	Level 1	Level 3	Level 4	Level 2	Level 2	Level 1	Level 1	Level 4	Level 1	Level 1	Level 4
	B2	Normal	Level 5	Level 3	Level 5	Level 4	Level 1	Level 2	Level 5	Level 2	Level 2	Level 2	Level 2	Level 1	Level 1	Level 1	Level 5
		Cancer	Level 5	Level 2	Level 5	Level 3	Level 1	Level 2	Level 4	Level 2	Level 1	Level 2	Level 2	Level 2	Level 1	Level 1	Level 5
	B3	Normal	Level 5	Level 3	Level 4	Level 3	Level 1	Level 3	Level 4	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 4
		Cancer	Level 5	Level 5	Level 4	Level 3	Level 1	Level 3	Level 4	Level 2	Level 1	Level 2	Level 2	Level 3	Level 1	Level 1	Level 4
	B4	Normal	Level 5	Level 3	Level 4	Level 4	Level 2	Level 3	Level 4	Level 2	Level 3	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 3	Level 3	Level 3	Level 2	Level 4	Level 4	Level 2	Level 3	Level 2	Level 1	Level 2	Level 2	Level 2	Level 3
	B5	Normal	Level 5	Level 3	Level 5	Level 3	Level 2	Level 3	Level 5	Level 2	Level 2	Level 2	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 3	Level 5	Level 3	Level 2	Level 3	Level 5	Level 1	Level 2	Level 2	Level 2	Level 2	Level 2	Level 1	Level 3
	B6	Normal	Level 5	Level 3	Level 5	Level 4	Level 1	Level 4	Level 5	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 3	Level 3	Level 3	Level 5	Level 2	Level 4	Level 4	Level 2	Level 3	Level 1	Level 3	Level 3	Level 1	Level 2	Level 3
	B7	Normal	Level 5	Level 2	Level 3	Level 4	Level 2	Level 4	Level 4	Level 2	Level 3	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
		Cancer	Level 4	Level 2	Level 3	Level 5	Level 2	Level 4	Level 5	Level 3	Level 3	Level 1	Level 2	Level 3	Level 1	Level 3	Level 4
	B8	Normal	Level 5	Level 4	Level 4	Level 3	Level 2	Level 4	Level 5	Level 2	Level 3	Level 1	Level 3	Level 2	Level 1	Level 2	Level 3
		Cancer	Level 5	Level 4	Level 4	Level 3	Level 2	Level 4	Level 5	Level 2	Level 1	Level 1	Level 4	Level 2	Level 1	Level 2	Level 4

Supplementary Table 3 *(Continued from the previous page)* Results of LOH and MSP analyses on 100 patients

Tissue	ID	Type						Result of methylation-specific PCR[‡]									
			MAGEA2	TFF2	DDX53	MSLN	MASP-IN	BGLAP	MUC8	KIAA1752	CDKN-2A	ESR2	PPARG	MLH1	CDH1	VDR	RUNX3
	B9	Normal	Level 5	Level 2	Level 4	Level 4	Level 1	Level 2	Level 5	Level 2	Level 3	Level 2	Level 2	Level 1	Level 1	Level 1	Level 4
		Cancer	Level 5	Level 2	Level 4	Level 4	Level 1	Level 2	Level 5	Level 2	Level 3	Level 1	Level 3	Level 1	Level 1	Level 1	Level 5
	B10	Normal	Level 5	Level 3	Level 5	Level 4	Level 1	Level 4	Level 4	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 2	Level 3
		Cancer	Level 3	Level 3	Level 3	Level 5	Level 2	Level 4	Level 4	Level 2	Level 3	Level 1	Level 3	Level 3	Level 1	Level 2	Level 4
	B11	Normal	Level 5	Level 2	Level 4	Level 3	Level 1	Level 3	Level 4	Level 2	Level 3	Level 2	Level 2	Level 2	Level 1	Level 2	Level 3
		Cancer	Level 4	Level 3	Level 3	Level 3	Level 1	Level 4	Level 4	Level 2	Level 2	Level 1	Level 3	Level 3	Level 1	Level 2	Level 3
	B12	Normal	Level 5	Level 3	Level 3	Level 4	Level 1	Level 4	Level 4	Level 2	Level 3	Level 1	Level 1	Level 2	Level 1	Level 2	Level 3
		Cancer	Level 5	Level 2	Level 4	Level 5	Level 1	Level 4	Level 5	Level 2	Level 2	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
	B13	Normal	Level 5	Level 3	Level 4	Level 3	Level 2	Level 4	Level 4	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 4	Level 5	Level 2	Level 3	Level 4	Level 2	Level 3	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
	B14	Normal	Level 5	Level 3	Level 4	Level 4	Level 1	Level 4	Level 4	Level 2	Level 3	Level 1	Level 2	Level 3	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 3	Level 4	Level 3	Level 1	Level 4	Level 3	Level 2	Level 1	Level 2	Level 2	Level 4	Level 1	Level 2	Level 3
	B15	Normal	Level 5	Level 3	Level 4	Level 3	Level 1	Level 3	Level 5	Level 2	Level 3	Level 1	Level 2	Level 3	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 2	Level 4	Level 5	Level 1	Level 3	Level 4	Level 3	Level 3	Level 2	Level 2	Level 3	Level 1	Level 1	Level 4
	B16	Normal	Level 4	Level 3	Level 4	Level 3	Level 1	Level 4	Level 3	Level 2	Level 3	Level 1	Level 2	Level 3	Level 1	Level 1	Level 3
		Cancer	Level 3	Level 3	Level 4	Level 4	Level 2	Level 4	Level 3	Level 3	Level 2	Level 1	Level 3	Level 2	Level 1	Level 1	Level 4
	B17	Normal	Level 5	Level 3	Level 4	Level 5	Level 1	Level 3	Level 3	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 4	Level 4	Level 4	Level 1	Level 3	Level 3	Level 2	Level 3	Level 1	Level 2	Level 5	Level 1	Level 1	Level 3
	B18	Normal	Level 5	Level 2	Level 4	Level 3	Level 2	Level 4	Level 4	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 4	Level 3	Level 2	Level 4	Level 3	Level 2	Level 3	Level 2	Level 2	Level 4	Level 1	Level 1	Level 3
	B19	Normal	Level 3	Level 3	Level 4	Level 4	Level 2	Level 3	Level 4	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 3	Level 4	Level 3	Level 3	Level 1	Level 3	Level 4	Level 2	Level 3	Level 1	Level 1	Level 3	Level 1	Level 2	Level 3
	B20	Normal	Level 5	Level 3	Level 4	Level 3	Level 1	Level 2	Level 5	Level 2	Level 4	Level 1	Level 2	Level 1	Level 1	Level 1	Level 4
		Cancer	Level 5	Level 4	Level 4	Level 4	Level 1	Level 2	Level 5	Level 2	Level 4	Level 2	Level 2	Level 2	Level 1	Level 1	Level 4
	B21	Normal	Level 5	Level 2	Level 5	Level 4	Level 1	Level 3	Level 5	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 3	Level 5	Level 4	Level 1	Level 3	Level 5	Level 2	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
	B22	Normal	Level 4	Level 4	Level 4	Level 3	Level 1	Level 3	Level 4	Level 1	Level 3	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 3	Level 3	Level 3	Level 5	Level 1	Level 4	Level 5	Level 2	Level 2	Level 1	Level 3	Level 2	Level 1	Level 1	Level 4
	B23	Normal	Level 4	Level 4	Level 4	Level 4	Level 2	Level 4	Level 3	Level 2	Level 3	Level 1	Level 1	Level 2	Level 1	Level 1	Level 4
		Cancer	Level 4	Level 5	Level 3	Level 4	Level 2	Level 4	Level 4	Level 2	Level 2	Level 1	Level 2	Level 2	Level 1	Level 1	Level 4
	B24	Normal	Level 4	Level 3	Level 4	Level 3	Level 2	Level 1	Level 5	Level 2	Level 4	Level 1	Level 2	Level 1	Level 1	Level 1	Level 4
		Cancer	Level 4	Level 4	Level 4	Level 4	Level 2	Level 2	Level 5	Level 2	Level 4	Level 2	Level 2	Level 1	Level 1	Level 2	Level 5
	B25	Normal	Level 4	Level 4	Level 4	Level 4	Level 1	Level 4	Level 4	Level 2	Level 3	Level 1	Level 2	Level 3	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 4	Level 4	Level 5	Level 1	Level 4	Level 4	Level 2	Level 2	Level 1	Level 2	Level 1	Level 2	Level 2	Level 3
Naso-pharynx	N1	Normal	Level 5	Level 3	Level 5	Level 4	Level 1	Level 4	Level 4	Level 2	Level 2	Level 2	Level 2	Level 3	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 2	Level 4	Level 4	Level 1	Level 3	Level 5	Level 1	Level 1	Level 2	Level 2	Level 1	Level 1	Level 1	Level 5
	N2	Normal	Level 5	Level 2	Level 5	Level 3	Level 1	Level 1	Level 5	Level 2	Level 3	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 5	Level 5	Level 1	Level 3	Level 5	Level 2	Level 2	Level 1	Level 1	Level 1	Level 1	Level 2	Level 5
	N3	Normal	Level 5	Level 2	Level 4	Level 3	Level 1	Level 2	Level 5	Level 2	Level 3	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 4	Level 2	Level 1	Level 2	Level 3	Level 2	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
	N4	Normal	Level 5	Level 2	Level 5	Level 4	Level 2	Level 3	Level 2	Level 2	Level 2	Level 2	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 5	Level 4	Level 2	Level 3	Level 1	Level 1	Level 2	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
	N5	Normal	Level 5	Level 2	Level 5	Level 4	Level 1	Level 1	Level 4	Level 2	Level 2	Level 2	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 2	Level 5	Level 4	Level 1	Level 2	Level 4	Level 2	Level 3	Level 1	Level 1	Level 3	Level 1	Level 2	Level 2
	N6	Normal	Level 5	Level 4	Level 5	Level 3	Level 1	Level 2	Level 4	Level 2	Level 2	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 3	Level 5	Level 4	Level 1	Level 1	Level 5	Level 2	Level 1	Level 1	Level 1	Level 1	Level 1	Level 1	Level 5
	N7	Normal	Level 5	Level 3	Level 5	Level 3	Level 1	Level 1	Level 4	Level 2	Level 2	Level 1	Level 2	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 5	Level 4	Level 3	Level 4	Level 1	Level 3	Level 5	Level 2	Level 1	Level 2	Level 2	Level 1	Level 1	Level 2	Level 2
	N8	Normal	Level 5	Level 4	Level 5	Level 3	Level 1	Level 2	Level 4	Level 2	Level 2	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 3	Level 3	Level 3	Level 1	Level 1	Level 4	Level 2	Level 1	Level 1	Level 2	Level 2	Level 1	Level 1	Level 5
	N9	Normal	Level 4	Level 4	Level 5	Level 3	Level 1	Level 3	Level 4	Level 2	Level 1	Level 2	Level 2	Level 1	Level 1	Level 1	Level 2
		Cancer	Level 4	Level 3	Level 4	Level 3	Level 1	Level 3	Level 4	Level 1	Level 1	Level 1	Level 3	Level 1	Level 1	Level 1	Level 2
	N10	Normal	Level 5	Level 3	Level 5	Level 3	Level 1	Level 4	Level 4	Level 2	Level 3	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 5	Level 5	Level 1	Level 4	Level 4	Level 2	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 4
	N11	Normal	Level 4	Level 1	Level 4	Level 4	Level 1	Level 1	Level 4	Level 2	Level 2	Level 1	Level 1	Level 3	Level 1	Level 1	Level 3

(Continued to the next page)

Supplementary Table 3 *(Continued from the previous page)* Results of LOH and MSP analyses on 100 patients

| Tissue	ID	Type	Result of methylation specific PCR[‡]														
			MAGEA2	TFF2	DDX53	MSLN	MASP-IN	BGLAP	MUC8	KIAA1752	CDKN-2A	ESR2	PPARG	MLH1	CDH1	VDR	RUNX3
		Cancer	Level 4	Level 2	Level 4	Level 4	Level 1	Level 2	Level 4	Level 2	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 5
	N12	Normal	Level 5	Level 2	Level 4	Level 3	Level 1	Level 4	Level 4	Level 2	Level 3	Level 2	Level 1	Level 1	Level 1	Level 2	Level 4
		Cancer	Level 5	Level 2	Level 5	Level 4	Level 1	Level 3	Level 3	Level 3	Level 2	Level 1	Level 1	Level 1	Level 1	Level 2	Level 5
	N13	Normal	Level 5	Level 3	Level 5	Level 3	Level 1	Level 2	Level 5	Level 2	Level 1	Level 1	Level 1	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 3	Level 5	Level 3	Level 1	Level 2	Level 4	Level 2	Level 1	Level 1	Level 1	Level 1	Level 1	Level 1	Level 4
	N14	Normal	Level 5	Level 1	Level 4	Level 3	Level 1	Level 1	Level 4	Level 2	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 1	Level 4	Level 5	Level 1	Level 2	Level 4	Level 1	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 5
	N15	Normal	Level 5	Level 3	Level 4	Level 4	Level 1	Level 2	Level 4	Level 2	Level 3	Level 1	Level 3	Level 1	Level 1	Level 1	Level 2
		Cancer	Level 5	Level 2	Level 4	Level 4	Level 2	Level 1	Level 4	Level 2	Level 2	Level 1	Level 3	Level 2	Level 2	Level 1	Level 2
	N16	Normal	Level 4	Level 2	Level 4	Level 3	Level 1	Level 1	Level 5	Level 2	Level 2	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 4	Level 5	Level 5	Level 1	Level 2	Level 5	Level 2	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 5
	N17	Normal	Level 5	Level 2	Level 4	Level 4	Level 1	Level 2	Level 5	Level 2	Level 1	Level 1	Level 2	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 5	Level 3	Level 5	Level 4	Level 2	Level 1	Level 4	Level 3	Level 2	Level 1	Level 3	Level 2	Level 1	Level 2	Level 2
	N18	Normal	Level 5	Level 3	Level 5	Level 4	Level 1	Level 2	Level 4	Level 2	Level 2	Level 1	Level 1	Level 2	Level 1	Level 2	Level 2
		Cancer	Level 4	Level 3	Level 4	Level 4	Level 1	Level 2	Level 4	Level 2	Level 2	Level 1	Level 2	Level 2	Level 2	Level 2	Level 4
	N19	Normal	Level 5	Level 2	Level 4	Level 4	Level 2	Level 2	Level 5	Level 2	Level 3	Level 2	Level 3	Level 2	Level 1	Level 1	Level 2
		Cancer	Level 5	Level 2	Level 5	Level 5	Level 2	Level 1	Level 4	Level 2	Level 1	Level 1	Level 2	Level 2	Level 1	Level 1	Level 3
	N20	Normal	Level 5	Level 2	Level 5	Level 4	Level 1	Level 2	Level 5	Level 2	Level 1	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 5	Level 2	Level 5	Level 4	Level 2	Level 1	Level 5	Level 2	Level 1	Level 1	Level 2	Level 2	Level 2	Level 2	Level 4
	N21	Normal	Level 4	Level 2	Level 5	Level 2	Level 1	Level 2	Level 4	Level 2	Level 2	Level 1	Level 3	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 1	Level 4	Level 3	Level 2	Level 1	Level 5	Level 3	Level 2	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
	N22	Normal	Level 4	Level 2	Level 4	Level 3	Level 2	Level 2	Level 4	Level 2	Level 2	Level 2	Level 2	Level 2	Level 2	Level 1	Level 2
		Cancer	Level 4	Level 3	Level 4	Level 4	Level 2	Level 1	Level 4	Level 1	Level 2	Level 2	Level 2	Level 2	Level 2	Level 2	Level 2
	N23	Normal	Level 4	Level 3	Level 5	Level 3	Level 1	Level 1	Level 4	Level 2	Level 3	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 3	Level 5	Level 5	Level 1	Level 2	Level 4	Level 2	Level 2	Level 2	Level 3	Level 1	Level 1	Level 1	Level 5
	N24	Normal	Level 4	Level 2	Level 4	Level 4	Level 1	Level 2	Level 4	Level 2	Level 2	Level 1	Level 2	Level 1	Level 1	Level 1	Level 3
		Cancer	Level 4	Level 2	Level 4	Level 4	Level 1	Level 2	Level 4	Level 2	Level 1	Level 2	Level 2	Level 1	Level 1	Level 1	Level 5
	N25	Normal	Level 4	Level 5	Level 5	Level 3	Level 1	Level 4	Level 5	Level 2	Level 1	Level 1	Level 2	Level 2	Level 1	Level 1	Level 2
		Cancer	Level 4	Level 5	Level 5	Level 3	Level 1	Level 4	Level 5	Level 2	Level 1	Level 1	Level 2	Level 2	Level 1	Level 2	Level 2

*, LOH-H=high-level chromosomal losses, LOH-L=low-level chromosomal losses. [†], Loss of single allele in the heterozygous case without MSI was interpreted as a loss of heterozygosity (LOH+). [‡], Proportion of the methylated CpGs was divided into 5 levels; 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).
LOH, loss of heterozygosity; MSP, methylayion-specific PCR.

Supplementary Table 4 List of SAGE and EST libraries used in analysis of transcript population

Tissue	No.	Library name	Age	Sex	Information
Embryo	1	SAGE_Embryonic_stem_cell_HES3_normal_p16_CL_ SHE10		F	Embryonic stem cell line - passage 16
	2	SAGE_Embryonic_stem_cell_HES4_normal_p36_CL_ SHE11		M	Embryonic stem cell line - passage 36
Placenta	1	SAGE_Placenta_first_trimester_normal_B_1			First trimester placenta
	2	SAGE_Placenta_normal_B_1		F	Full-term placenta
Normal tissues					
Stomach	1	SAGE_Stomach_normal_MD_13S		M/F	Male and Female, Microdissection
	2	SAGE_Stomach_normal_epithelium_B_body1			Normal gastric epithelium, bulk
	3	SAGE_Stomach_normal_B_antrum			A pool of normal gastric epithelial tissues obtained from the antrum
	4	SAGE_Stomach_normal_MD_14S		M/F	Male and Female, Microdissection
Colon	1	SAGE_Colon_normal_B_NC1			Normal colonic epithelium
	2	SAGE_Colon_normal_B_NC2			Normal colonic epithelium
Breast	1	SAGE_Breast_normal_epithelium_AP_Br_N	35	F	Normal human luminal mammary epithelial cells purified with BER-EP4 Ab.
	2	SAGE_Breast_normal_epithelium_AP_1	43	F	Normal human luminal mammary epithelial cells purified with BER-EP4 Ab.
	3	SAGE_Breast_normal_myoepithelium_AP_IDC7	47	F	Myoepithelial cells purified CD10 Ab.
	4	SAGE_Breast_normal_stroma_AP_1	31	F	Normal breast tissue
	5	SAGE_Breast_normal_myoepithelium_AP_myoepithelial	27	F	Myoepithelial cells purified CD10 Ab.
	6	SAGE_Breast_normal_stroma_B_IDC8	44	F	Normal breast tissue
Naso-	1	NCI_CGAP_HN11			Tongue, microdissected
pharynx	2	NCI_CGAP_HN19			Nasopharyngeal epithelium
	3	NCI_CGAP_HN9			Retromolar trigone, microdissected
Cancer tissues					
Stomach	1	SAGE_Stomach_adenocarcinoma_MD_HG7	71	F	Adenocarcinoma, Microdissection
	2	SAGE_Stomach_adenocarcinoma_MD_G329	60		Adenocarcinoma, Microdissection
	3	SAGE_Stomach_adenocarcinoma_MD_HS29	66	F	Adenocarcinoma, Microdissection
	4	SAGE_Stomach_adenocarcinoma_B_G234	57	M	Adenocarcinoma,,gastroesophageal junction, T4N0M0
	5	SAGE_Stomach_carcinoma_B_xenograph_X43	78	F	Poorly differentiated carcinoma, T4N1M0
	6	SAGE_Stomach_carcinoma_B_G189			Poorly differentiated, T4N0M0
Colon	1	SAGE_Colon_adenocarcinoma_B_Tu102			Colon, primary tumor
	2	SAGE_Colon_adenocarcinoma_B_Tu98			Colon, primary tumor
Breast	1	SAGE_Breast_carcinoma_B_95-259	56.5	F	Invasive ductal carcinoma, ER-, PR-
	2	SAGE_Breast_carcinoma_B_95-347	52.8	F	Invasive ductal carcinoma, ER+, PR+
	3	SAGE_Breast_carcinoma_B_BWHT18	44	F	ER+, PR+, ErbB2 -
	4	SAGE_Breast_carcinoma_B_IDC-4		F	Invasive ductal carcinoma, ER-, ErbB2-, p53+
	5	SAGE_Breast_carcinoma_B_IDC-5		F	Invasive ductal carcinoma, ER+, ErbB2-, p53-
	6	SAGE_Breast_metastatic_carcinoma_B_95-260	56.5	F	Metastatic tumor. Paired with breast tumor No. 2
	7	SAGE_Breast_metastatic_carcinoma_B_95-348	52.8	F	Metastatic tumor. Paired with breast tumor No. 3
	8	SAGE_Breast_carcinoma_MD_LCIS	34	F	Extensive LCIS
Naso-	1	NCI_CGAP_HN12			Tongue, microdissected
pharynx	2	NCI_CGAP_HN21			Nasopharyngeal carcinoma
	3	NCI_CGAP_HN16			Retromolar trigone, microdissected

SAGE and EST libraries were downloaded from http://cgap.nci.nih.gov/.
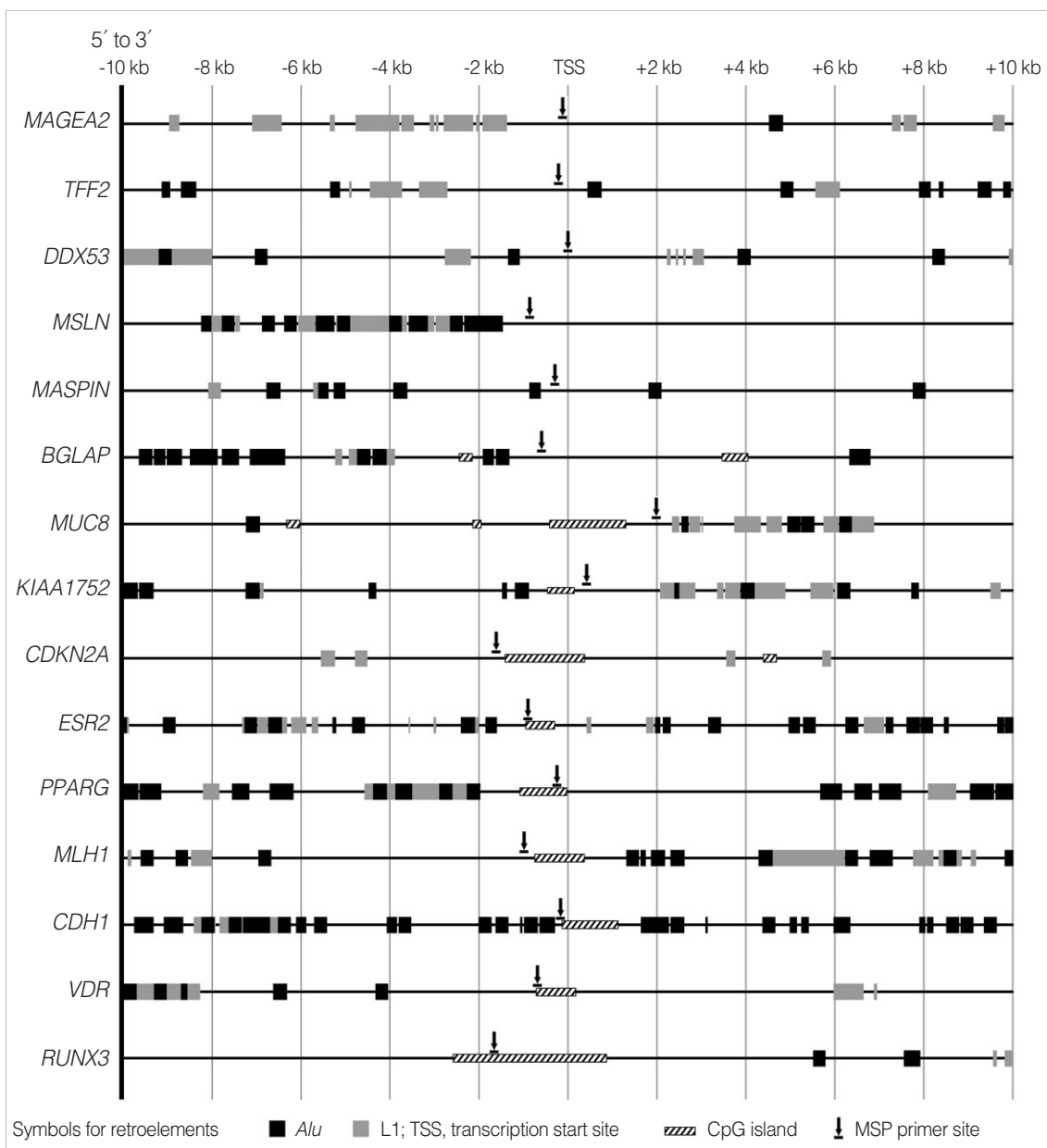
Supplementary Fig. 1. Schematic diagram of CpG islands and nearby retroelement distributions in the 5´-end regions of 15 selected genes. Of the six genes lacking CpG islands, the *MAGEA2* and *TFF2* genes have the L1 elements in a nearby 3-kb window, the *DDX53* and *MSLN* genes have the *Alu* and L1 elements, and the *MASPIN* and *BGLAP* genes have the *Alu* elements. The remaining nine genes contain CpG islands at the transcriptional start sites. The *MUC8*, *KIAA1752*, and *CDKN2A* genes were examined at the extragenic sites of the CpG-island margins close to the L1 elements, the *ESR2* and *PPARG2* genes were examined at those close to the *Alu* and L1 elements, and the *CDH1* gene was examined at those close to the *Alu* elements. The CpG islands of the *MLH1*, *VDR*, and *RUNX3* genes were bordered by the retroelements at a long distance. The *MUC8* and *KIAA1752* genes were examined at the intragenic sites of the CpG island margins close to the L1 elements.
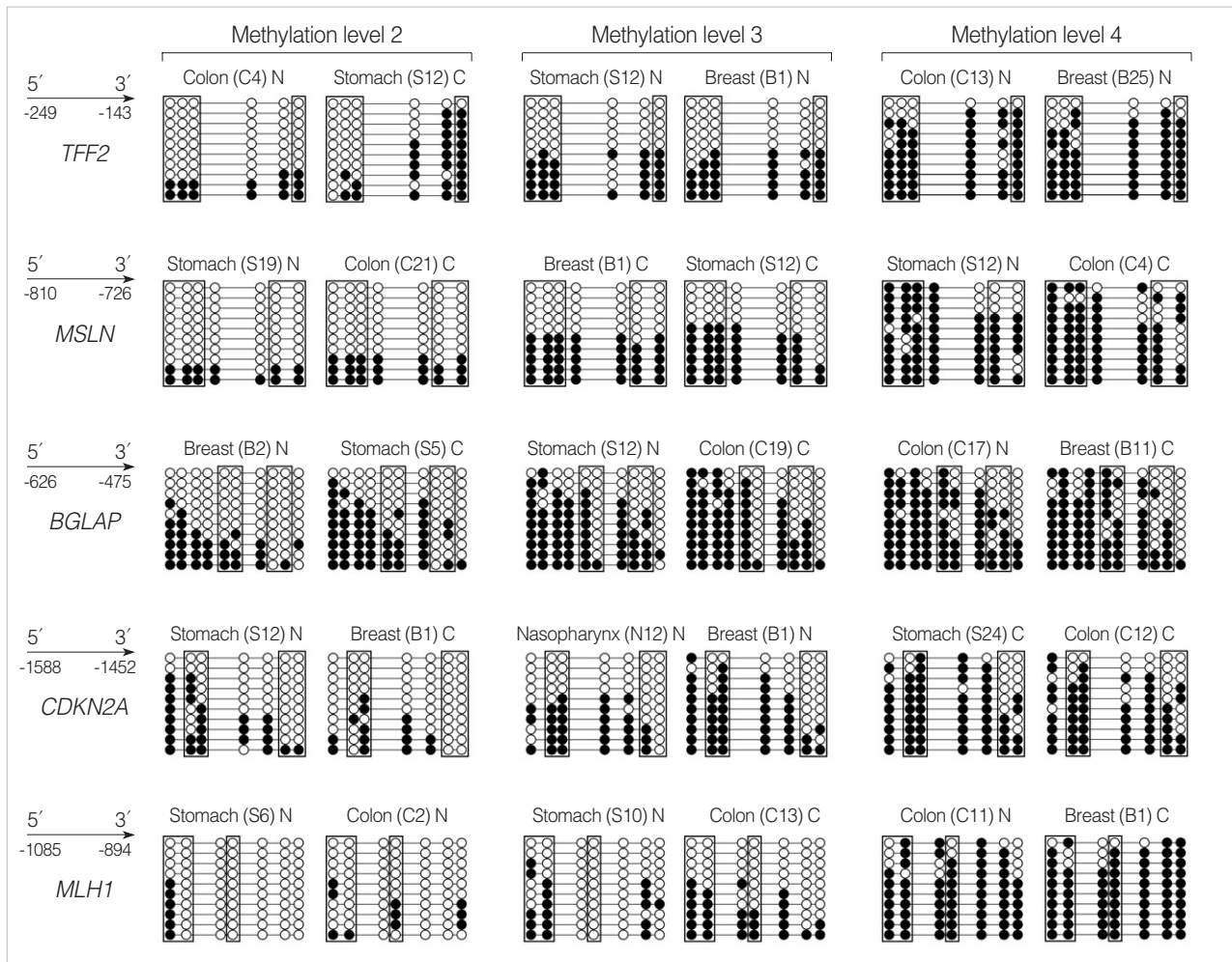
Supplementary Fig. 2. The methylation-variable CpGs of the *TFF2*, *MSLN*, *BGLAP*, *CDKN2A*, and *MLH1* genes were analyzed by cloning and sequencing the common PCR DNA. The CpG-island-negative genes (*TFF2*, *MSLN*, and *BGLAP*) showed both hypomethylated and hypermethylated DNA strands in the transitional area or a methylation gradient between the hypomethylated proximal CpGs and the hypermethylated distal CpGs. The CpG-island-positive genes (*CDKN2A* and *MLH1*) were unmethylated at the proximal CpG site and methylated at the distal CpG site. The proportion of hypermethylated DNA strand was consistent with the level of methylation estimated by the intensity of methylation-specific PCR band.