THYROID
Volume 30, Number 2, 2020

DOI: 10.1089/thy.2019.0011

Comprehensive DNA Methylation Profiling Identifies Novel Diagnostic Biomarkers for Thyroid Cancer

Jong-Lyul Park,[1–3] Sora Jeon,[4,5] Eun-Hye Seo,[1,3] Dong Hyuck Bae,[1,3] Young Mun Jeong,[4] Yourha Kim,[4,5] Ja Seong Bae,[5,6] Seon-Kyu Kim,[3] Chan Kwon Jung,[5,7] and Yong Sung Kim[1,2]

Background: There are no reliable biomarkers to accurately differentiate indolent thyroid tumors from more aggressive thyroid cancers. This study aimed to develop new DNA methylation markers for diagnosis and recurrence risk stratification of papillary thyroid carcinoma (PTC).
Methods: Thyroid tumor-specific DNA methylation profiling was investigated in 34 fresh frozen tissues, which included nontumor ($n=7$), noninvasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP, $n=6$) and PTC ($n=21$), using the Illumina HumanMethylation EPIC array. We performed a genome-wide assessment of thyroid tumor-specific differentially methylated CpG sites in the discovery set, then validated the top candidate markers in an independent set of 293 paraffin tissue samples comprised of follicular adenoma (FA, $n=61$), Hürthle cell adenoma (HA, $n=24$), NIFTP ($n=56$), PTC ($n=120$), follicular thyroid carcinoma ($n=27$), and Hürthle cell carcinoma ($n=5$), by pyrosequencing.
Results: Three selected markers (cg10705422, cg17707274, and cg26849382) differentiated nonmalignant (FA, HA, and NIFTP) tumors from differentiated thyroid cancers with area under the receiver operating characteristic curve of 0.83, 0.83, and 0.80, respectively. Low DNA methylation levels for three markers were significantly associated with recurrent or persistent disease (odds ratio (OR) = 3.860 [95% confidence interval (CI) 1.194–12.475]) and distant metastasis (OR = 4.009 [CI 1.098–14.632]) in patients with differentiated thyroid cancer. A subgroup analysis for the validation set showed that PTC patients with low DNA methylation levels more frequently had aggressive histology, extrathyroidal extension, lymph node metastasis, $BRAF^{V600E}$ mutations, and recurrent or persistent disease than those with high levels of methylation markers. All PTC patients who developed disease recurrence had low DNA methylation levels for three markers.
Conclusions: DNA methylation levels of three markers can be useful for differentiating differentiated thyroid cancer from nonmalignant follicular thyroid lesions, and may serve as prognostic biomarkers for predicting recurrent or persistent disease after surgery for differentiated thyroid cancer.

Keywords: DNA methylation, thyroid neoplasms, methylation markers, NIFTP, papillary thyroid cancer, recurrence

Introduction

Papillary thyroid carcinoma (PTC) is the most common thyroid cancer, and its incidence has rapidly increased over the last three decades (1). There are >10 histologic variants of PTC that are characterized by disparate molecular and clinical features (2). The follicular variant is the second most common type of PTC, and consists of infiltrative and encapsulated forms (2). According to the 2017 World Health Organization classification of thyroid tumors, most cases of noninvasive encapsulated follicular variant (EFV) of PTC have been reclassified as noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP). NIFTP tends to have an indolent clinical behavior with uncertain malignant potential and can be considered to reflect a carcinoma *in situ* (2,3). Aggressive variants of PTC include tall cell, columnar cell, and hobnail variants, in which risk of recurrence and disease mortality is increased compared with other PTC variants (4).

[1]Genome Editing Research Center; [2]Personalized Genomic Medicine Research Center; Korea Research Institute of Bioscience and Biotechnology, Daejeon, Republic of Korea.
[3]Department of Functional Genomics, University of Science and Technology, Daejeon, Republic of Korea.
Departments of [4]Biomedicine and Health Sciences, [6]Surgery, and [7]Hospital Pathology; [5]Cancer Research Institute; College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea.

Encapsulated thyroid tumors with a follicular growth pattern include follicular adenoma (FA), NIFTP, follicular thyroid carcinoma (FTC), and invasive EFVPTC (2). Although NIFTP and invasive EFVPTC have different nuclear features from those of FA and FTC, cytologic differential diagnosis by preoperative fine needle aspiration cytology in these tumors is challenging. It is impossible to cytologically distinguish NIFTP from invasive EFVPTC and FA from FTC in thyroid fine needle aspiration specimens (5–9). Furthermore, the follicular patterned tumors have similar molecular alterations, with a dominant prevalence of *RAS* mutations (2). Until now, there has been no accurate method for the preoperative diagnosis of these thyroid tumors. Their final diagnosis is rendered by pathologic examination on surgical specimen. Therefore, these follicular patterned tumors are considered to be a surgical disease requiring diagnostic lobectomy (2,3).

DNA methylation is the most well-known epigenetic modification, occurring from the addition of a methyl (CH_3) group to the 5′-position of the cytosine of cytosine-guanine dinucleotides (CpG). Inactivation of tumor suppressor genes can occur through hypermethylation at the promoter region of genes, and oncogenes can be activated by promoter hypomethylation (10,11). In addition, recent studies have revealed that DNA methylation at enhancer or superenhancer regions plays important roles in cancer progression through regulation of target gene expression (12,13). Most previous studies and The Cancer Genome Atlas (TCGA) project used the Infinium HumanMethylation450K BeadChip methylation microarray platform (450K; Illumina, San Diago, CA), due to its low cost, small amount of input DNA, simple workflow, and fast sample processing time. However, the Infinium HumanMethylation 450K microarray focuses on the coding RNAs loci and lacks coverage of the enhancer regions (14). In this regard, the previous studies were limited by the low genome coverage of the method. EPIC BeadChip (Illumina) is a recently developed platform that covers >850,000 CpG methylation sites. The EPIC BeadChip microarray covers 66%, 72%, and 83% for the ENCODE open chromatin, ENCODE transcription factor binding sites (TFBSs) in open chromatin, and FANTOM5 enhancer database, respectively (15). To date, a DNA methylation study covering 850,000 CpG methylation sites in thyroid tumors has not yet been reported.

In this study, we performed genome-wide DNA methylation profiling using the EPIC BeadChip microarray to identify disease-specific DNA methylation markers in thyroid tumors. We subsequently selected potential DNA methylation markers to differentiate NIFTP from the other entities, and verified their clinicopathologic utility in an independent cohort.

Methods

Study subjects

This study was approved by the Institutional Review Board of Seoul St. Mary's Hospital of the Catholic University of Korea (KC16SISI0709). Informed consent was obtained. Thyroid tumor tissue and matched adjacent normal thyroid samples were obtained from the Biobank of Seoul St. Mary's Hospital.

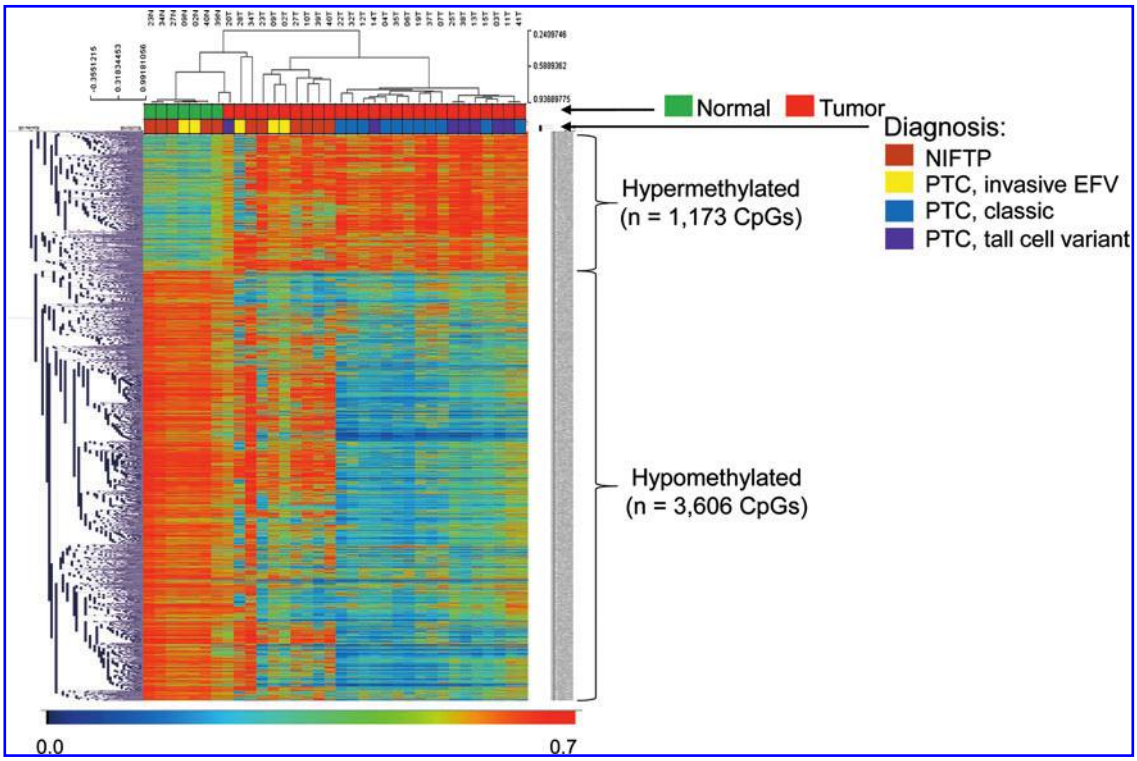

FIG. 1. Unsupervised hierarchical clustering of 34 thyroid normal and tumor tissue samples, using thyroid cancer-specific DMC sites. Thyroid cancer-specific DMCs were selected based on the *p*-value (<0.005) and methylation differences (>0.2 or <0.2) between thyroid normal and tumor tissue samples. The columns represent the cases, and the lines represent the CpG sites. The red and blue colors indicate high and low methylation levels, respectively. DMC, differentially methylated CpG; EFV, encapsulated follicular variant; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features; PTC, papillary thyroid carcinoma.

For profiling of DNA methylation, we used 34 fresh frozen tissue samples, including matched normal ($n=7$), NIFTP ($n=6$), invasive EFVPTC ($n=3$), classic PTC ($n=11$), and tall cell variants (TCVs) PTC ($n=7$). In addition, we used 293 formalin-fixed paraffin-embedded tissue samples to validate our selected DNA methylation markers by pyrosequencing. Pathology slides of all cases were reviewed and classified according to the 2017 World Health Organization classification of tumors of endocrine organs (2). The validation cohort consisted of FA ($n=61$), Hürthle cell adenoma (HA, $n=24$), NIFTP ($n=56$), PTC ($n=120$), FTC ($n=27$), and Hürthle cell carcinoma (HCC, $n=5$). Tumor stages were categorized according to the eighth edition of the American Joint Committee on Cancer (AJCC) staging manual. Recurrence risk was evaluated using the American Thyroid Association (ATA) classification for risk of recurrence (4). Detailed demographic and baseline characteristics of patients are described in Supplementary Table S1.

DNA isolation and BRAF mutational analysis

Genomic DNA was isolated from fresh frozen tissues and 10 μm thick paraffin-embedded tissue sections using RecoverAll™ Total Nucleic Acid Isolation Kit (Life Technologies, Carlsbad, CA), according to the manufacturer's instructions. The quality and quantity of the extracted genomic DNA were analyzed with an ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). After polymerase chain reaction (PCR) amplification of the extracted DNA, sequences of *BRAF* exon 15 were analyzed by direct sequencing of amplicons, as described previously (16,17).

DNA methylation microarray experiment and data analysis

EPIC BeadChip (Illumina) was used for methylation array experiments, per instructions of the manufacturer. In brief, 500 ng of genomic DNA collected from thyroid normal and tumor tissues was treated with 20 μL sodium bisulfite solution included in the EZ DNA Methylation-Gold Kit (Zymo Research, Orange, CA). Bisulfite-converted DNA (4 μL) was amplified using the Infinium Methylation Assay kit (Illumina). Amplified DNA was hybridized to an EPIC BeadChip and scanned with the Illumina iSCAN system. CpG methylation values were calculated as average-β values using the minfi package (version 1.26.2) (18) of R software, and we used the

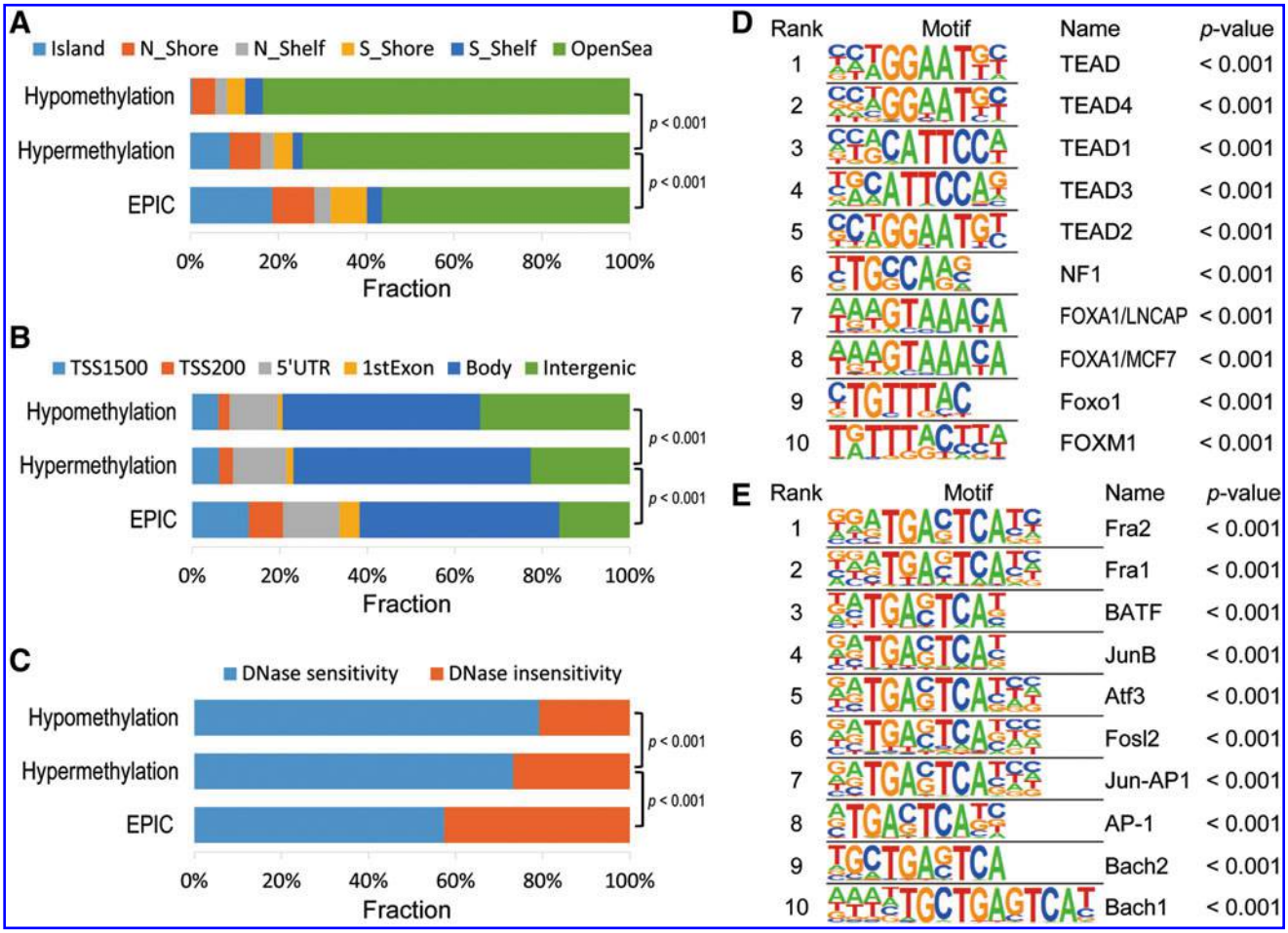

FIG. 2. Distribution features of hypo- and hypermethylated DMC, and known motif analysis. (**A**) CpG relationship, (**B**) genomic distribution, and association with (**C**) DNase sensitivity of DMC loci. The sequences within +100 bp or −100 bp flanking each of the hypo- or hypermethylated CpG sites were used for known motif analyses. The 10 most significant motifs for (**D**) hypermethylated or (**E**) hypomethylated DMCs were subsequently compared with known transcription factor-binding sites. Shore, ~0–2 kb from the CGI; shelf, ~2–4 kb from the CGI; open sea, >4 kb from the CGI. CGI, CpG island.

functional normalization method to remove technical variations (19). Measurements with detection p-values <0.05 were considered to have a significant signal above background. All primary methylation array data were deposited in the GEO database under accession number GSE121377.

Public RNA-sequencing and DNA methylome data collection

Public RNA-sequencing data of thyroid normal and PTC samples were obtained from the TCGA dataset (https://portal.gdc.cancer.gov/) to estimate the correlation between DNA methylation and gene expression in PTC. In addition, we collected the Infinium HumanMethylation 450K data from TCGA dataset to confirm the methylation pattern of our top 10 candidate DNA methylation markers.

Motif and gene ontology analysis

Analysis of sequence motifs was performed using a HOMER package (version 4.10) with the default parameter settings, using thyroid cancer-specific hypo- or hypermethylated sites located at DNase-sensitive regions (20). Regions for motif analyses were defined as 100 bp upstream to 100 bp downstream of the differentially methylated CpGs (DMCs). We performed gene ontology analysis using DMC linked to 5′-regulatory regions (http://david.abcc.ncifcrf.gov/) to predict the functions of DMC-linked genes.

DNA methylation analysis by pyrosequencing

A pyrosequencing technique was employed to validate the selected DNA methylation markers in an independent cohort. In brief, 500 ng of total DNA from each of the paraffin-embedded tissue sections was used for bisulfite conversion using the EZ DNA Methylation-Gold kit (Zymo Research). Each sample was eluted with 20 μL elution buffer from the kit. Next, 1 μL of the bisulfite-converted DNA was used in a 20 μL PCR mixture containing primer sets and 2x Master Mix (Doctor Protein, Seoul, Korea), and amplified using a GeneAmp PCR system 9700 (Applied Biosystems, Waltham, MA). For pyrosequencing, forward, reverse, and sequencing primers were designed using PSQ Assay Design v2.0.1.15 (Biotage, Kungsgatan, Sweden). Standard pyrosequencing was then performed. In brief, 20 μL of PCR product was immobilized on 3 μL of Streptavidin Sepharose High Performance (GE Healthcare Bio-Sciences, Uppsala, Sweden) and annealed with sequencing primer for 10 minutes at 80°C. Finally, the generated pyrograms were analyzed using PyroMark analysis software (Biotage). Sequences for the primer sets (Bioneer, Daejeon, Korea) are shown in Supplementary Table S2.

Statistical analysis

The methylation status analysis was performed blinded to clinicopathologic data, and, conversely, clinicopathologic utility of DNA methylation markers was determined blinded to methylation status. We used Student's t-test or analysis of variance (ANOVA) to evaluate the significance of differences in gene expression and DNA methylation levels between normal and thyroid cancer tissues, or NIFTP and other subtypes in thyroid tumors. The relationship between clinicopathologic features and the levels of methylation was analyzed using parametric (chi-squared test) and nonparametric (Fisher's exact) assessments, where appropriate. Logistic regression was performed to assess the associations of clinicopathologic variables and DNA methylation levels with the

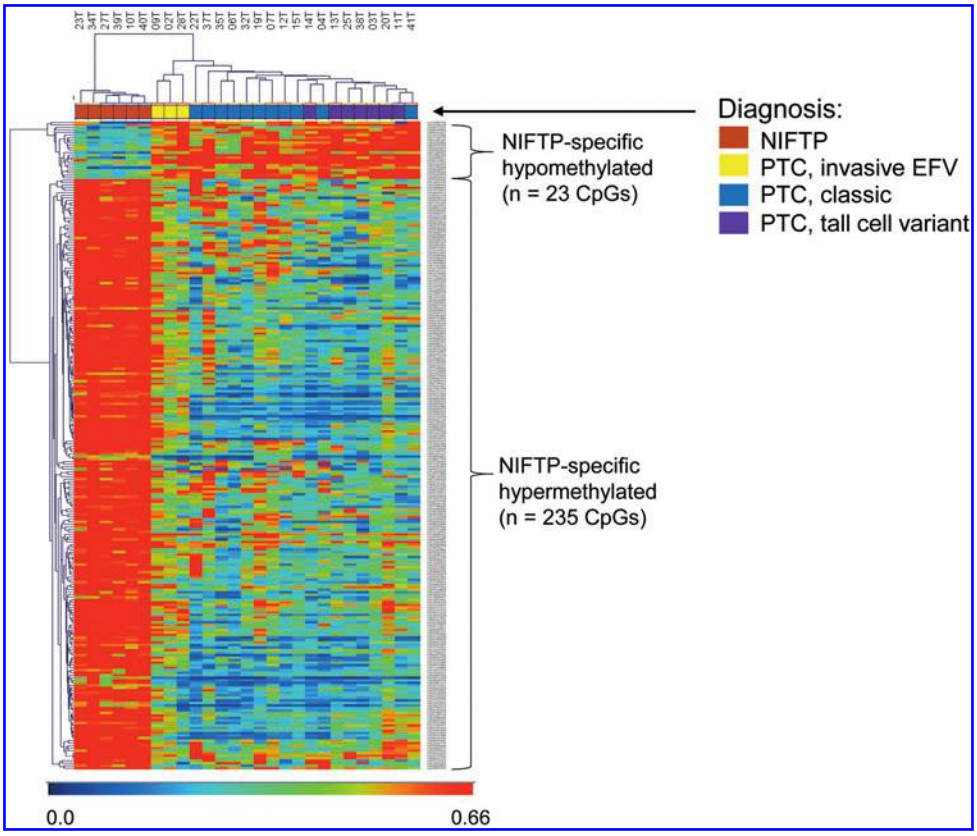

FIG. 3. Heatmap of NIFTP-specific hypo- and hypermethylated CpG sites. NIFTP-specific DMCs were selected based on the p-value and methylation differences. (1) p-Value <0.005 and methylation differences >0.3 or <–0.3 (average-β scale) between NIFTP and PTC and (2) p-value <0.005 and DNA methylation differences >0.2 or <0.2 (average-β scale) between NIFTP and invasive EFV of PTC. The columns represent the cases, and the lines represent the CpG sites. The red and blue colors indicate high and low methylation levels, respectively.

TABLE 1. TOP 10 CANDIDATE DNA METHYLATION MARKERS TO DISCRIMINATE NONINVASIVE FOLLICULAR THYROID NEOPLASM WITH PAPILLARY-LIKE NUCLEAR FEATURES AND PAPILLARY THYROID CARCINOMA

Illumina ID[a]	*Delta-β (NIFTP vs. other)*	*Delta-β (NIFTP vs. IEFVPTC)*	p *(NIFTP vs. other)*	p *(NIFTP vs. IEFVPTC)*	*AUC*	*Chromosome*	*MapInfo*[b]	*Symbol*	*CpG island*[c]	*Gene feature group*
cg10705422	0.59	0.38	1.60E-10	9.47E-04	1	chr11	12188825	*MICAL2*	OpenSea	Body
cg15441605	0.5	0.42	1.45E-11	4.68E-05	1	chr9	12814643	*LURAP1L-AS1*	OpenSea	TSS1500
cg24327132	0.59	0.34	1.23E-10	1.94E-03	1	chr15	72520632	*PKM2*	N_Shore	5′UTR
cg16336556	0.51	0.4	1.14E-10	3.43E-04	1	chr2	33295138	*LTBP1*	OpenSea	Body
cg17707274	0.53	0.37	1.73E-11	1.49E-03	1	chr11	1.02E+08	*MMP7*	OpenSea	TSS200
cg00567113	0.48	0.42	1.30E-11	4.17E-03	1	chr3	87382813		OpenSea	
cg06034194	0.46	0.43	7.49E-11	7.96E-04	1	chr9	12814626	*LURAP1L-AS1*	OpenSea	TSS1500
cg21341586	0.51	0.38	2.38E-06	3.58E-03	1	chr4	99851281	*EIF4E*	S_Shore	5′UTR
cg26849382	0.52	0.37	8.00E-08	4.81E-03	1	chr5	1.41E+08	*DIAPH1*	OpenSea	Body
cg05763918	0.48	0.4	1.83E-09	4.01E-03	1	chr4	1.29E+08	*LOC100507487*	OpenSea	Body

[a]Illumina ID is the unique identification number in the HumanEPIC BeadChip.
[b]MapInfo indicates the genomic location in human reference genome 37 (GRCh37/hg19), released by the Genome Reference Consortium in March 3, 2009 (www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/).
[c]Shore and shelf are adjacent to the CpG island (2- and 4-kb regions flanking the CpG island, respectively). N and S mean upstream and downstream of the CpG island, respectively.
AUC, area under the receiver operating characteristic curve; IEFV, invasive encapsulated follicular variant; IEFVPTC, IEFV of papillary thyroid carcinoma; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features.

adverse clinical outcomes. Hierarchical clustering was performed with Multiple Experiment Viewer (MEV) software (version 4.8.1), using Pearson's correlation method. The receiver operating characteristic (ROC) and the respective areas under the ROC curve (AUC) were calculated for each DNA methylation marker, using the ROCR package of the R software (version 3.4.0). ROC curve analysis estimated the optimal cutoff values maximizing sensitivity and specificity between low and high levels of methylation. Results with p-values <0.05 were considered significant.

Results

Identification of thyroid tumor-specific DMC sites

To identify thyroid tumor-specific DMCs, we calculated the p-value and methylation differences between thyroid normal ($n=7$) and tumor ($n=27$) samples. Thus, we applied the following two criteria: (1) p-value <0.005 and (2) methylation differences >0.2 or <–0.2 (average-β scale) between thyroid normal and tumor samples. As a result, we selected 3606 hypo- and 1173 hypermethylated CpG sites. Of the selected DMCs, more than half of the DMCs include novel loci (70.93% and 56.90% for hypo- and hypermethylated DMCs), compared with the Infinium HumanMethylation 450K bead array (Supplementary Fig. S1). Unsupervised hierarchical clustering of the DNA methylation of the DMC candidates is shown in Figure 1.

To estimate the correlation between DNA methylation and gene expression, we collected RNA-sequencing data of thyroid normal and tumor tissues from the TCGA dataset. Of the DMCs, 318 hypo- and 114 hypermethylated DMCs were located in the 5′-regulatory regions (promoter regions) of 266 and 88 genes, respectively. Next, we evaluated the DNA methylation levels and mRNA expression levels of the 88 and 266 genes in the thyroid normal and tumor tissues. There was a negative association between DNA methylation and gene expression levels (Supplementary Fig. S2).

Feature of thyroid cancer-specific DMC loci

We analyzed distribution features of the thyroid cancer-specific DMCs based on the CpG island relation, gene structure, and DNase sensitivity. CpG island relations were classified into four categories: islands, shore (up to 2 kb from CpG island), shelves (2–4 kb from CpG island), and open sea (>4 kb from CpG island). We observed that hypo- or hypermethylated DMCs in thyroid cancer were enriched in the open sea regions compared with the reference distribution on the EPIC bead array (Fig. 2A). Gene structures were classified into six categories: TSS1500, TSS200, 1stExon, Body, and intergenic. Our results showed that hypo- and hypermethylated DMCs were predominantly located in the intergenic loci, compared with the EPIC bead array (Fig. 2B). When DMCs were separated into two categories based on DNase sensitivity, both hypo- and hypermethylated DMCs were enriched on the DNase-sensitive loci (Fig. 2C).

TFs are proteins with DNA binding activity that are involved in the regulation of transcription. Generally, TFs modulate gene expression by binding to gene promoter regions or to distal regions called enhancers. The distance between a TFBS and a transcription start site (TSS) of a gene regulated by the TF can be up to several megabases, and

depends on the chromatin structure of the region (21). To examine whether thyroid cancer-specific DMCs were associated with TFBSs, we performed an enrichment analysis of known binding motifs. Sequences within 100 bp upstream or 100 bp downstream flanking each of the hypo- and hypermethylated CpG sites located at the DNase-sensitive loci were used for the enrichment analysis of known binding motifs. Our results reveal that hypermethylated CpG sites are enriched in the TEAD TF family, and hypomethylated DMCs were enriched in the Fra2, FraI, BATF, and JunB TF. The top 10 binding motifs for the hypo- and hypermethylated CpG sites are shown in Figure 2D and E, respectively.

Identification of NIFTP-specific DNA methylation markers

To select the NIFTP-specific DNA methylation markers, we calculated the *p*-values and methylation differences. Thus, we applied the following two criteria: (1) *p*-value <0.005 and methylation differences >0.3 or <–0.3 (average-β scale) between NIFTP ($n=6$) and PTC ($n=21$), and (2) *p*-value <0.005 and DNA methylation differences >0.2 or <0.2 (average-β scale) between NIFTP ($n=6$) and invasive EFVPTC ($n=3$). As a result, we selected 23 hypo- and 235 hypermethylated CpG sites, respectively. We performed unsupervised hierarchical clustering with these DMC candidates (Fig. 3).

For evaluation in an independent cohort, we selected the top 10 DNA methylation candidates using DNA methylation differences and AUC values to discriminate NIFTP from PTCs (classic PTC, invasive EFVPTC, TCVPTC). The top 10 candidate DNA methylation markers are summarized in Table 1. Of 10 candidate DNA methylation markers, four DNA methylation markers were included in the Infinium HumanMethylation 450K platform. Thus, we could evaluate the methylation patterns of candidate DNA methylation markers in the TCGA cohort, and observed that four candidate DNA methylation markers were more heavily hypomethylated in tumor tissue regardless of PTC variants than adjacent nontumor thyroid (Fig. 4).

Validation of potential DNA methylation markers in an independent cohort by pyrosequencing

The NIFTP-specific DNA methylation markers from the microarray data were evaluated with a bisulfite modification-based pyrosequencing assay of 293 paraffin tissue samples comprised of FA ($n=61$), HA ($n=24$), NIFTPs ($n=56$), PTC ($n=120$), FTC ($n=27$), and HCC ($n=5$). Among the top 10 candidate DNA methylation markers, we succeeded in pyrosequencing design for cg10705422, cg17707274, and cg26849382, and failed pyrosequencing conditions for the other candidate CpG sites. Each representative pyrogram is shown in Supplementary Figure S3.

Three candidates were more strongly hypomethylated in PTC than other thyroid tumors (Fig. 5). The three selected DNA methylation markers were further evaluated for their capacity to differentiate nonmalignant (FA, HA, and NIFTP) from malignant (PTC, FTC, and HCC) thyroid tumors. To

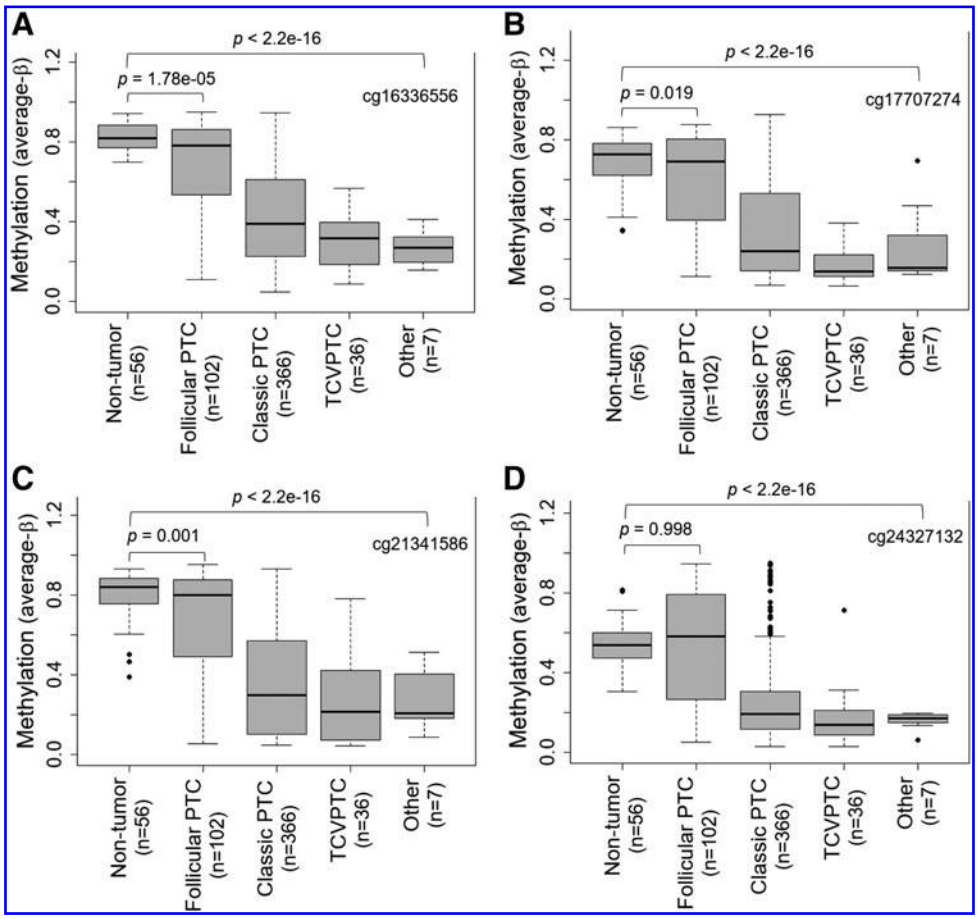

FIG. 4. Quantification of methylation levels of candidate DNA methylation markers according to the histologic types of PTC in The Cancer Genome Atlas (TCGA) cohort. DNA methylation levels of cg16336556 (**A**), cg17707274 (**B**), cg21341586 (**C**), and cg24327132 (**D**) were evaluated in nontumor tissue ($n=56$) and PTC (classic type, $n=366$; follicular variant, $n=102$; TCV, $n=36$), other variant, $n=7$). Significance of differences was determined by ANOVA. ANOVA, analysis of variance; TCV, tall cell variant.

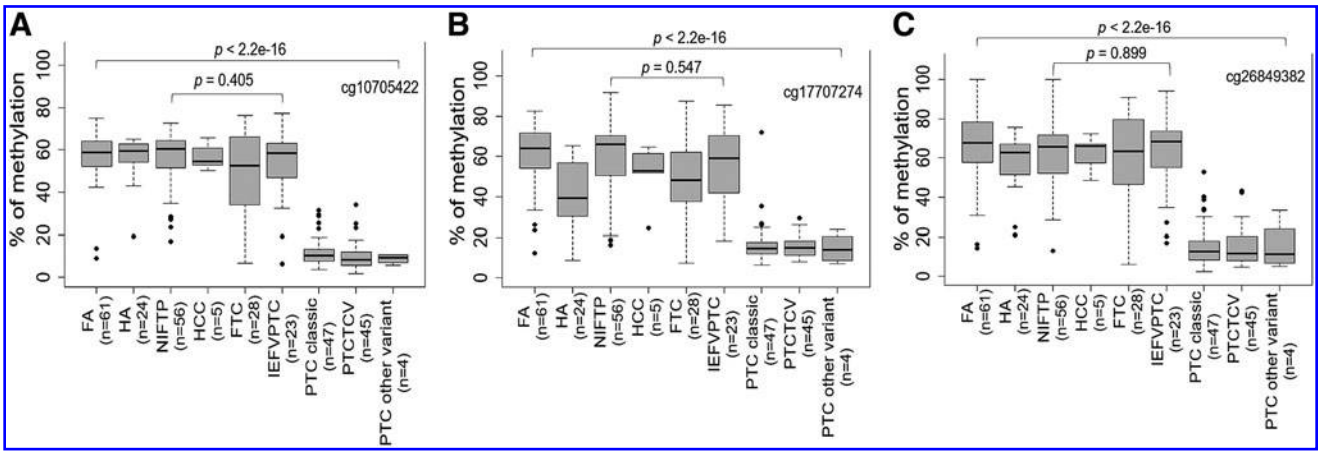

FIG. 5. Evaluation of candidate methylation markers in an independent cohort by pyrosequencing. A total of 293 formalin-fixed paraffin-embedded tissue samples were used for quantification of DNA methylation levels of cg10705422 (**A**), cg17707274 (**B**), and cg26849382 (**C**) by the pyrosequencing assay. Significance of differences was determined by ANOVA. FA, follicular adenoma; FTC, follicular thyroid carcinoma; HA, Hürthle cell adenoma; HCC, Hürthle cell carcinoma; IEFV, invasive encapsulated follicular variant.

estimate the AUC value, we performed ROC analysis, and observed good sensitivity and specificity. The AUC values of cg10705422, cg17707274, and cg26849382 were 0.83, 0.83, and 0.80, respectively (Fig. 6A–C). Their optimal cutoff values are described in Figure 6.

Clinicopathologic utility of the three DNA methylation markers

All patients were classified into two subgroups according to different levels of methylation of three DNA methylation markers and categorized into four groups based on the number of markers showing low DNA methylation levels: all high (group 1, $n=116$), one low (group 2, $n=39$), two low (group 3, $n=21$), or all low (group 4, $n=117$). PTC was mostly enriched in group 4 (Fig. 7A). Patients classified as high risk according to the ATA recurrence risk stratification system were most frequently observed in group 4 (Fig. 7B). Recurrent or persistent diseases were mostly found in group 4 (Fig. 7C). Patients with stage IV thyroid cancer at thyroid surgery were only observed in group 4 (Fig. 7D). Multivariate logistic regression analysis showed that hypomethylation of three DNA methylation markers was significant predictor for recurrent or persistent disease (odds ratio (OR) = 3.860 [95% confidence interval (CI) 1.194–12.475]) and distant metastasis (OR = 4.009 [CI 1.098–14.632]) in 152 patients with differentiated thyroid cancer (Table 2).

A subgroup analysis was conducted in 120 patients with PTC, as shown in Table 3. Low levels of methylation at cg10705422, cg17707274, and cg26849382 were associated with histologic variants ($p<0.001$), extrathyroidal extension ($p<0.001$), multifocality ($p<0.001$), lymph node metastasis ($p<0.001$), $BRAF^{V600E}$ mutations ($p<0.001$), recurrent or persistent disease ($p=0.040$), and increased ATA recurrence risk ($p<0.001$).

Discussion

Thyroid cancer is a prime example for which widely available imaging modalities have resulted in an increased

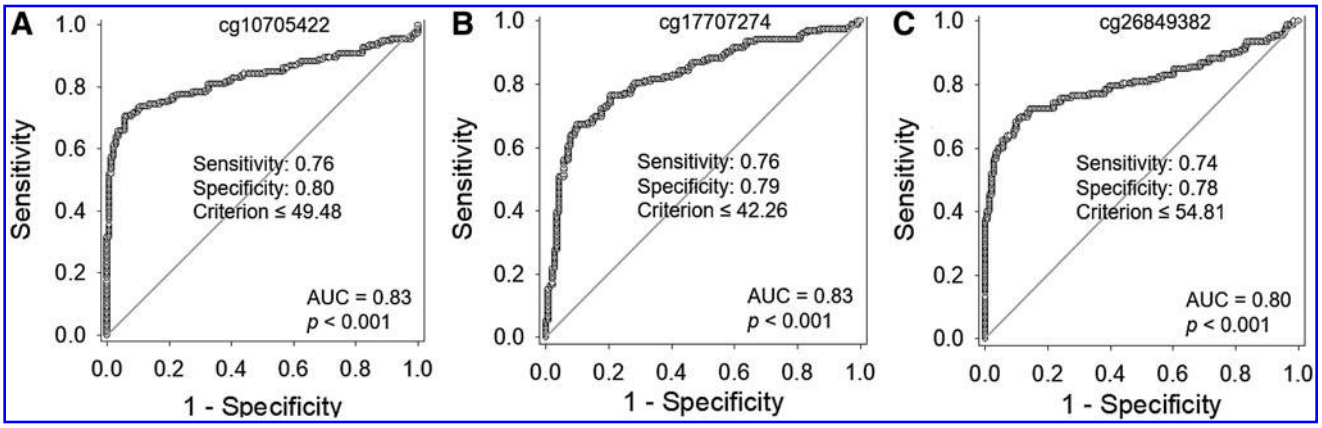

FIG. 6. Receiver operating characteristic curve analyses of three selected DNA methylation markers for discrimination of nonmalignant tumors (FA, HA, and NIFTP) from PTC, FTC, and HCC. The AUC indicates the probability that the classifier ranks a randomly chosen positive instance higher than a randomly chosen negative instance. The gray curves indicate 95% confidence bounds. The criteria for low DNA methylation levels of cg10705422 (**A**), cg17707274 (**B**), and cg26849382 (**C**) were defined using the AUC. AUC, area under the receiver operating characteristic curve.

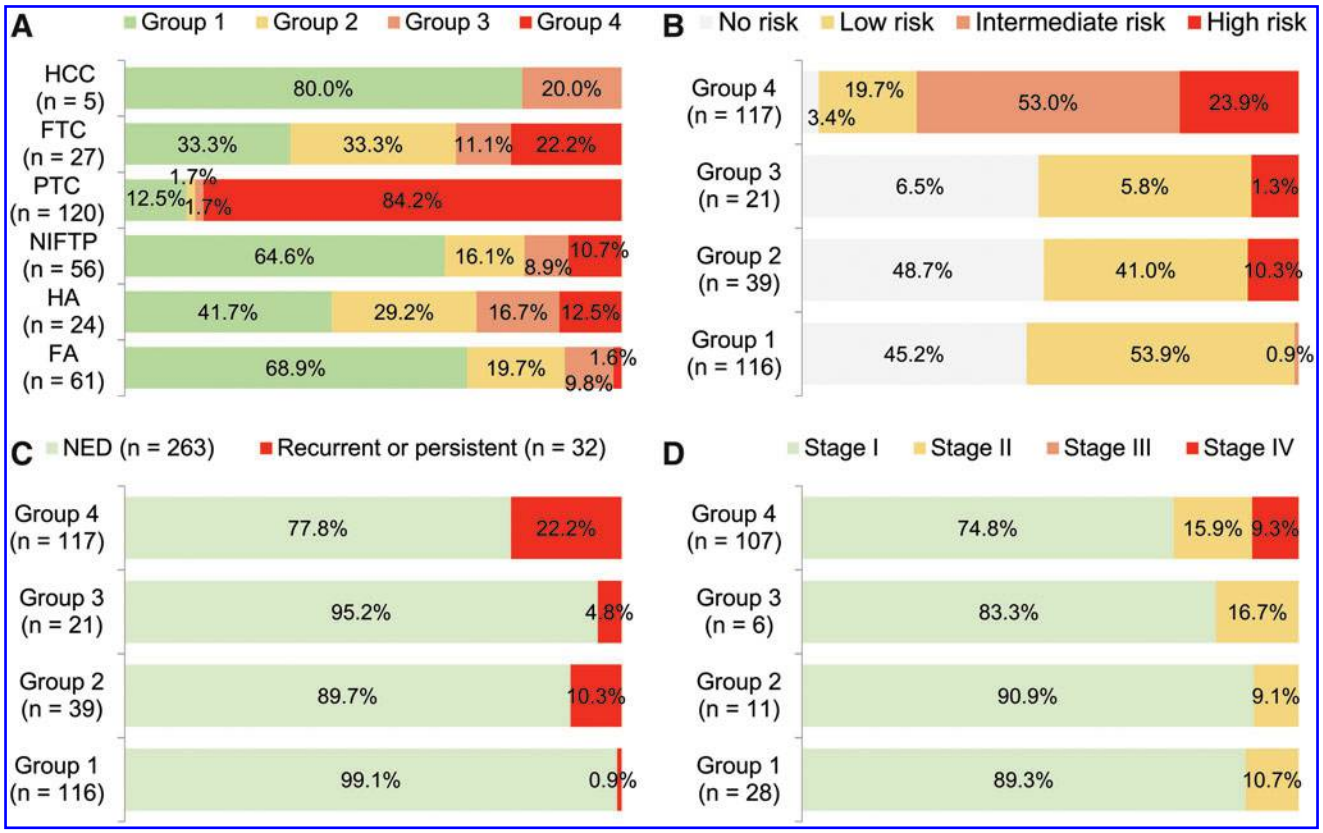

FIG. 7. Diagnostic performance from combination of three DNA methylation markers. Thyroid tumors were divided into all high (group 1), one low (group 2), two low (group 3), and all low (group 4) levels of DNA methylation markers. The distribution of histologic subtypes of tumor (**A**), disease recurrence risk (**B**), and recurrent or persistent disease (**C**) in all 293 thyroid tumors, and tumor stage (**D**) in 152 thyroid cancers were stratified based on these four groups.

incidence of early cancers with indolent behavior, a phenomenon commonly described as cancer ''overdiagnosis,'' which often results in subsequent ''overtreatment'' (1). Therefore, identifying biomarkers for risk stratification of thyroid tumors may provide tools to reduce medical overtreatment. In this study, we show that a diagnostic classifier based on three DNA methylation markers could differentiate nonmalignant tumors and differentiated thyroid cancers, and serve as a prognostic marker to stratify differentiated thyroid cancer patients by their risk for recurrent or persistent disease.

Alterations in DNA methylation have been shown to play a role in tumorigenesis and disease progression in many malignancies, including thyroid cancer. In this study, we quantitatively profiled the genome-wide DNA methylation of thyroid tumors using the Illumina HumanMethylation EPIC bead array. The results of our gene ontology analyses were largely consistent with previous results. For example, DMCs between thyroid tumor and normal tissues were enriched in thyroid cancer-associated processes, such as the inflammatory response (22), cellular response to tumor necrosis factor

Table 2. Multivariate Logistic Regression Analyses for Risk Factors Associated with Adverse Outcomes in 152 Patients with Differentiated Thyroid Cancer

	Recurrent or persistent disease[a]		*Distant metastasis*[b]	
	Adjusted OR [CI]	p	*Adjusted OR [CI]*	p
Age ≥55 years	2.005 [0.851–4.725]	0.112	2.975 [1.193–7.417]	0.019
Male sex	0.485 [0.183–1.289]	0.147	0.315 [0.106–0.942]	0.039
PTC	0.473 [0.113–1.977]	0.305	0.341 [0.072–1.608]	0.174
Aggressive histology[c]	3.421 [1.150–10.180]	0.027	4.624 [1.307–16.356]	0.018
Multifocal tumor	1.278 [0.505–3.233]	0.605	1.185 [0.428–3.280]	0.744
$BRAF^{V600E}$ mutation	3.421 [1.460–16.718]	0.010	0.212 [0.059–0.763]	0.018
Hypomethylation of three DNA methylation markers	3.860 [1.194–12.475]	0.024	4.009 [1.098–14.632]	0.036

[a]Includes biochemical ($n=2$) and structural locoregional recurrence ($n=3$), and distant metastasis ($n=27$).
[b]Includes synchronous ($n=20$) and metachronous ($n=7$) distant metastasis.
[c]Includes TCV ($n=45$), columnar cell variant ($n=2$), and hobnail variant ($n=1$) of PTC, encapsulated angioinvasive FTC with extensive angioinvasion ($n=3$), and widely invasive FTC ($n=1$).
CI, 95% confidence interval; FTC, follicular thyroid carcinoma; OR, odds ratio; TCV, tall cell variant.

Table 3. Methylation Levels of Three Selected DNA Methylation Markers in 120 Papillary Thyroid Carcinoma Samples, with Respect to Clinical Pathologic Characteristics

	cg10705422			*cg17707274*			*cg26849382*		
Characteristic	*Low methylation*	*High methylation*	p	*Low methylation*	*High methylation*	p	*Low methylation*	*High methylation*	p
Age (years)			0.319			0.667			0.319
<55	73 (88.0%)	10 (12.0%)		72 (86.7%)	11 (13.3%)		73 (88.0%)	10 (12.0%)	
≥55	30 (81.1%)	7 (18.9%)		31 (83.8%)	6 (16.2%)		30 (81.1%)	7 (81.1%)	
Sex			0.277			0.969			0.277
Male	51 (89.5%)	6 (10.5%)		49 (86.0%)	8 (14.0%)		51 (89.5%)	6 (10.5%)	
Female	52 (82.5%)	11 (17.5%)		54 (85.7%)	9 (14.3%)		52 (82.5%)	11 (17.5%)	
Tumor size (cm)	1.9±1.1	2.3±1.4	0.203	1.9±1.1	2.4±1.4	0.099	1.9±1.1	2.3±1.4	0.203
Histologic subtypes			<0.001			<0.001			<0.001
Classic	48 (100%)	0		47 (97.9%)	1 (2.1%)		48 (100%)	0	
IEFV	6 (26.1%)	17 (73.9%)		7 (30.4%)	16 (69.6%)		6 (26.1%)	17 (73.9%)	
TCV	45 (100%)	0		45 (100%)	0		45 (100%)	0	
Other	4 (100%)	0		4 (100%)	0				
Histologic aggressiveness			<0.001			<0.001			<0.001
Nonaggressive variant	55 (76.4%)	17 (23.6%)		55 (76.4%)	17 (23.6%)		55 (76.4%)	17 (23.6%)	
Aggressive variant	48 (100%)	0		48 (100%)	0		48 (100%)	0	
Extrathyroidal extension			<0.001			<0.001			<0.001
Absent	24 (60.0%)	16 (40.0%)		24 (60.0%)	16 (40.0%)		24 (60.0%)	16 (40.0%)	
Microscopic	61 (98.4%)	1 (1.6%)		61 (98.4%)	1 (1.6%)		61 (98.4%)	1 (1.6%)	
Gross	18 (100%)	0		18 (100%)	0		18 (100%)	0	
Multifocality			<0.001			0.001			0.001
Absent	46 (75.4%)	15 (24.6%)		46 (75.4%)	15 (24.6%)		46 (75.4%)	15 (24.6%)	
Present	56 (96.6%)	2 (3.4%)		56 (96.6%)	2 (3.4%)		56 (96.6%)	2 (3.4%)	
Lymph node metastasis			<0.001			<0.001			<0.001
Absent	33 (67.3%)	16 (32.7%)		33 (67.3%)	16 (32.7%)		33 (67.3%)	16 (32.7%)	
Present	70 (98.6%)	1 (1.4%)		70 (98.6%)	1 (1.4%)		70 (98.6%)	1 (1.4%)	
pT stage			0.978			0.747			0.978
pT1	62 (87.3%)	9 (12.7%)		63 (88.7%)	8 (11.3%)		62 (87.3%)	9 (12.7%)	
pT2	21 (80.8%)	5 (19.2%)		20 (76.9%)	6 (23.1%)		21 (80.8%)	5 (19.2%)	
pT3	15 (83.3%)	3 (16.7%)		15 (83.3%)	3 (16.7%)		15 (83.3%)	3 (16.7%)	
pT4	5 (100%)	0		5 (100%)	0		5 (100%)	0	
Distant metastasis[a]			0.042			0.042			0.042
Absent	84 (83.2%)	17 (16.8%)		84 (83.2%)	17 (16.8%)		84 (83.2%)	17 (16.8%)	
Present[a]	19 (100%)	0		19 (100%)	0		19 (100%)	0	
BRAF^{V600E} mutation			<0.001			<0.001			<0.001
Negative	22 (56.4%)	17 (43.6%)		23 (59.0%)	16 (41.0%)		22 (56.4%)	17 (43.6%)	
Positive	81 (100%)	0		80 (98.8%)	1 (1.2%)		81 (100%)	0	

(*continued*)

TABLE 3. (CONTINUED)

Characteristic	cg10705422		p	cg17707274		p	cg26849382		p
	Low methylation	*High methylation*		*Low methylation*	*High methylation*		*Low methylation*	*High methylation*	
Recurrent or persistent disease			0.022			0.022			0.022
Absent	79 (82.3%)	17 (17.7%)		79 (82.3%)	17 (17.7%)		79 (82.3%)	17 (17.7%)	
Present	24 (100%)	0		24 (100%)	0		24 (100%)	0	
ATA recurrence risk			<0.001			<0.001			<0.001
Low	16 (50.0%)	16 (50.0%)		16 (50.0%)	16 (50.0%)		16 (50.0%)	16 (50.0%)	
Intermediate	65 (98.5%)	1 (1.5%)		65 (98.5%)	1 (1.5%)		65 (98.5%)	1 (1.5%)	
High	22 (100%)	0		22 (100%)	0		22 (100%)	0	
AJCC stage			0.099			0.099			0.099
I	78 (83.0%)	16 (17.0%)		78 (83.0%)	16 (17.0%)		78 (83.0%)	16 (17.0%)	
II	16 (94.1%)	1 (5.9%)		16 (94.1%)	1 (5.9%)		16 (94.1%)	1 (5.9%)	
III	0	0		0	0		0	0	
IV	9 (100%)	0		9 (100%)	0		9 (100%)	0	

[a]Includes 17 synchronous and 2 metachronous metastases.
AJCC, American Joint Committee on Cancer; ATA, American Thyroid Association.

(23), positive regulation of angiogenesis (24), and immune response (25) (Supplementary Table S3). These results suggest that DNA methylation changes at the 5′-regulatory regions may be tightly associated with thyroid carcinogenesis through regulation of gene expression.

Recently, methylation changes at the outside of promoter, such as the enhancer regions, were shown to be important for regulation of gene expression (12,13), as well as promoter hypo- or hypermethylation. Our motif analyses showed that hypermethylated CpG sites were enriched in the TEAD TF family. The TEAD TF family is involved in the Hippo pathway (26), and tightly associated with cancer progression (27). In addition, hypomethylated DMCs were enriched in Fra2, Fra1, BATF, and JunB TF. Fra expression levels in thyroid tumor cells have been shown to modulate transcription of several tumor progression markers (28). Our results suggest that DNA methylation changes at TF binding sites in open chromatin may play an important role in the thyroid carcinogenesis, as well as DNA methylation at promoter regions.

Three DNA methylation markers (cg10705422, cg17707274, and cg26849382), selected for their clinical utility in thyroid tumors, were located in the gene body of *MICAL2*, the promoter of *MMP7*, and the gene body of *DIAPH1* gene, respectively. The *MICAL2* gene is a tumor-promoting factor, which can accelerate tumor progression through regulation of cell proliferation and epithelial–mesenchymal transition (29). The *MMP7* gene is one of the smallest members of the *MMP* family and is a highly potent metalloprotease. *MMP7* expression is required to mediate cell invasion and tumor formation (30), and it is regulated by DNA methylation (31). The expression levels of the *DIAPH1* gene associated with cg26849382 were positively correlated with metastasis of colorectal cancer (32). These previous results suggest that our selected DNA methylation markers may be associated with carcinogenesis or cancer progression. It is well known that *MMP7* gene expression is regulated by promoter DNA methylation (31,33,34), but DNA methylation regulation mechanisms in the gene body of *MICAL2* and *DIAPH1* have not yet been reported. Further studies are required to estimate the association between DNA methylation and gene expression of these two genes.

The data demonstrate a discrepancy in methylation patterns of three DNA methylation markers between the Illumina HumanMethylation EPIC bead array data and pyrosequencing validation in independent invasive EFVPTC and NIFTP cohorts (Fig. 5). The inconsistent results may be due to the small sample size in the genome-wide DNA methylation profiling step. Therefore, further studies will be needed to identify NIFTP-specific DNA methylation markers.

The three selected DNA methylation markers showed good sensitivity and specificity to discriminate nonmalignant thyroid tumors from differentiated thyroid cancers. Low methylation levels of three markers were independently associated with recurrent or persistent disease and distant metastasis in patients with differentiated thyroid cancer. In patients with PTC, the hypomethylation of DNA methylation markers was associated with the $BRAF^{V6000E}$ mutations, high recurrence risk defined by ATA, recurrent or persistent disease, and aggressive clinicopathologic features, including lymph node metastasis, extrathyroidal extension, aggressive histology, and distant metastasis. These results are similar to

previously published results (35,36). Mancikova *et al.* (35) reported that distinct methylation patterns were tightly associated with the specific mutation in thyroid cancer. Hou *et al.* (36) reported that thyroid cancer cells in which *BRAF* has been knocked down showed hypermethylation of an important proportion of genes.

In conclusion, we report that hypomethylation of three DNA methylation markers, cg10705422, cg17707274, and cg26849382, may serve as novel diagnostic and prognostic biomarker for differentiated thyroid cancer. These DNA methylation markers may be clinically useful for efficiently stratifying thyroid tumors.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This research was supported by a grant (HI16C2013) from the Korean Health Technology R&D Project (Cheongju-si, Chungcheongbuk-do, Republic of Korea), Ministry of Health and Welfare, Republic of Korea. This study was also supported by a grant (2017R1D1A1B03029597) from the Basic Science Research Program through the National Research Foundation of Korea (Daejeon, Republic of Korea) funded by the Ministry of Science and ICT.

Supplementary Material

Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3
Supplementary Table S1
Supplementary Table S2
Supplementary Table S3

References

1. Vaccarella S, Franceschi S, Bray F, Wild CP, Plummer M, Dal Maso L 2016 Worldwide thyroid-cancer epidemic? The increasing impact of overdiagnosis. N Engl J Med **375:**614–617.
2. Lloyd RV, Osamura RY, Klöppel G, Rosai J 2017 WHO Classification of Tumours of Endocrine Organs. Vol 10. Fourth edition. WHO Press, Geneva, Switzerland.
3. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, Barletta JA, Wenig BM, Al Ghuzlan A, Kakudo K, Giordano TJ, Alves VA, Khanafshar E, Asa SL, El-Naggar AK, Gooding WE, Hodak SP, Lloyd RV, Maytal G, Mete O, Nikiforova MN, Nose V, Papotti M, Poller DN, Sadow PM, Tischler AS, Tuttle RM, Wall KB, LiVolsi VA, Randolph GW, Ghossein RA 2016 Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of indolent tumors. JAMA Oncol **2:**1023–1029.
4. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L 2016 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. Thyroid **26:**1–133.
5. Bychkov A, Keelawat S, Agarwal S, Jain D, Jung CK, Hong S, Lai CR, Satoh S, Kakudo K 2018 Impact of non-invasive follicular thyroid neoplasm with papillary-like nuclear features on the Bethesda system for reporting thyroid cytopathology: a multi-institutional study in five Asian countries. Pathology **50:**411–417.
6. Kim M, Park HJ, Min HS, Kwon HJ, Jung CK, Chae SW, Yoo HJ, Choi YD, Lee MJ, Kwak JJ, Song DE, Kim DH, Lee HK, Kim JY, Hong SH, Sohn JS, Lee HS, Park SY, Hong SW, Shin MK 2017 The Use of the Bethesda System for Reporting Thyroid Cytopathology in Korea: a Nationwide Multicenter Survey by the Korean Society of Endocrine Pathologists. J Pathol Transl Med **51:**410–417.
7. Mahajan S, Agarwal S, Kocheri N, Jain D, Mathur SR, Iyer VK 2018 Cytopathology of non-invasive follicular thyroid neoplasm with papillary-like nuclear features: a comparative study with similar patterned papillary thyroid carcinoma variants. Cytopathology **29:**233–240.
8. Kim M, Kim JE, Kim HJ, Chung YR, Kwak Y, Park SY 2018 Cytologic diagnosis of noninvasive follicular thyroid neoplasm with papillary-like nuclear features and its impact on the risk of malignancy in the Bethesda system for reporting thyroid cytopathology: an institutional experience. J Pathol Transl Med **52:**171–178.
9. Han K, Ha H-J, Kong JS, Kim J-S, Myung JK, Koh JS, Park S, Shin M-S, Song W-T, Seol HS, Lee S-S 2018 Cytological features that differentiate follicular neoplasm from mimicking lesions. J Pathol Transl Med **52:**110–120.
10. Ehrlich M 2009 DNA hypomethylation in cancer cells. Epigenomics **1:**239–259.
11. Esteller M 2002 CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. Oncogene **21:**5427–5440.
12. Bell RE, Golan T, Sheinboim D, Malcov H, Amar D, Salamon A, Liron T, Gelfman S, Gabet Y, Shamir R, Levy C 2016 Enhancer methylation dynamics contribute to cancer plasticity and patient mortality. Genome Res **26:** 601–611.
13. Heyn H, Vidal E, Ferreira HJ, Vizoso M, Sayols S, Gomez A, Moran S, Boque-Sastre R, Guil S, Martinez-Cardus A, Lin CY, Royo R, Sanchez-Mut JV, Martinez R, Gut M, Torrents D, Orozco M, Gut I, Young RA, Esteller M 2016 Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. Genome Biol **17:**11.
14. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, Esteller M 2011 Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics **6:**692–702.
15. Moran S, Arribas C, Esteller M 2016 Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics **8:** 389–399.
16. Cho U, Mete O, Kim MH, Bae JS, Jung CK 2017 Molecular correlates and rate of lymph node metastasis of non-invasive follicular thyroid neoplasm with papillary-like nuclear features and invasive follicular variant papillary thyroid carcinoma: the impact of rigid criteria to distinguish non-invasive follicular thyroid neoplasm with papillary-like nuclear features. Mod Pathol **30:**810–825.
17. Jung CK, Kim Y, Jeon S, Jo K, Lee S, Bae JS 2018 Clinical utility of EZH1 mutations in the diagnosis of follicular-patterned thyroid tumors. Hum Pathol **81:**9–17.
18. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA 2014 Minfi: a flex-

ible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. Bioinformatics **30:**1363–1369.
19. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen KD 2014 Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol **15:**503.
20. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK 2010 Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell **38:**576–589.
21. Dekker J, Heard E 2015 Structural and functional diversity of topologically associating domains. FEBS Lett **589:** 2877–2884.
22. Guarino V, Castellone MD, Avilla E, Melillo RM 2010 Thyroid cancer and inflammation. Mol Cell Endocrinol **321:**94–102.
23. Lv N, Gao Y, Guan H, Wu D, Ding S, Teng W, Shan Z 2015 Inflammatory mediators, tumor necrosis factor-alpha and interferon-gamma, induce EMT in human PTC cell lines. Oncol Lett **10:**2591–2597.
24. Stabenow E, Tavares MR, Ab'Saber AM, Parra-Cuentas ER, de Matos LL, Eher EM, Capelozzi VL, Ferraz AR 2005 Angiogenesis as an indicator of metastatic potential in papillary thyroid carcinoma. Clinics (Sao Paulo) **60:**233–240.
25. Zhang S, Wang Y, Chen M, Sun L, Han J, Elena VK, Qiao H 2017 CXCL12 methylation-mediated epigenetic regulation of gene expression in papillary thyroid carcinoma. Sci Rep **7:**44033.
26. Yu FX, Guan KL 2013 The Hippo pathway: regulators and regulations. Genes Dev **27:**355–371.
27. Pan D 2010 The Hippo signaling pathway in development and cancer. Dev Cell **19:**491–505.
28. Tulchinsky E 2000 Fos family members: regulation, structure and role in oncogenic transformation. Histol Histopathol **15:**921–928.
29. Cai Y, Lu J, Tang F 2018 Overexpression of MICAL2, a novel tumor-promoting factor, accelerates tumor progression through regulating cell proliferation and EMT. J Cancer **9:**521–527.
30. Basu S, Thorat R, Dalal SN 2015 MMP7 is required to mediate cell invasion and tumor formation upon Plakophilin3 loss. PLoS One **10:**e0123979.
31. Sato N, Maehara N, Su GH, Goggins M 2003 Effects of 5-aza-2′-deoxycytidine on matrix metalloproteinase expression and pancreatic cancer cell invasiveness. J Natl Cancer Inst **95:**327–330.
32. Lin YN, Izbicki JR, Konig A, Habermann JK, Blechner C, Lange T, Schumacher U, Windhorst S 2014 Expression of DIAPH1 is up-regulated in colorectal cancer and its down-regulation strongly reduces the metastatic capacity of colon carcinoma cells. Int J Cancer **134:**1571–1582.
33. Sizemore ST, Sizemore GM, Booth CN, Thompson CL, Silverman P, Bebek G, Abdul-Karim FW, Avril S, Keri RA 2014 Hypomethylation of the MMP7 promoter and increased expression of MMP7 distinguishes the basal-like breast cancer subtype from other triple-negative tumors. Breast Cancer Res Treat **146:**25–40.
34. Chernov AV, Strongin AY 2011 Epigenetic regulation of matrix metalloproteinases and their collagen substrates in cancer. Biomol Concepts **2:**135–147.
35. Mancikova V, Buj R, Castelblanco E, Inglada-Perez L, Diez A, de Cubas AA, Curras-Freixes M, Maravall FX, Mauricio D, Matias-Guiu X, Puig-Domingo M, Capel I, Bella MR, Lerma E, Castella E, Reverter JL, Peinado MA, Jorda M, Robledo M 2014 DNA methylation profiling of well-differentiated thyroid cancer uncovers markers of recurrence free survival. Int J Cancer **135:**598–610.
36. Hou P, Liu D, Xing M 2011 Genome-wide alterations in gene methylation by the BRAF V600E mutation in papillary thyroid cancer cells. Endocr Relat Cancer **18:**687–697.

Address correspondence to:
Chan Kwon Jung, MD, PhD
Department of Hospital Pathology
College of Medicine
The Catholic University of Korea
Seoul St. Mary's Hospital
222 Banpo-daero, Seocho-gu
Seoul 06591
Republic of Korea

E-mail: ckjung@catholic.ac.kr

Yong Sung Kim, PhD
Genome Editing Research Center
Korea Research Institute of Bioscience and Biotechnology
Daejeon 34141
Republic of Korea

E-mail: yongsung@kribb.re.kr