

Prognostic Classifier Based on Genome-Wide DNA Methylation Profiling in Well-Differentiated Thyroid Tumors

Mariana Bisarro dos Reis,^{1,2*} Mateus Camargo Barros-Filho,^{1*} Fábio Albuquerque Marchi,¹ Caroline Moraes Beltrami,¹ Hellen Kuasne,¹ Clóvis Antônio Lopes Pinto,³ Srikant Ambatipudi,^{4,5} Zdenko Herceg,⁴ Luiz Paulo Kowalski,⁶ and Silvia Regina Rogatto^{2,7}

¹International Research Center, CIPE, A.C. Camargo Cancer Center and National Institute of Science and Technology in Oncogenomics, São Paulo 01509-010, SP, Brazil; ²Department of Urology, Faculty of Medicine, UNESP, São Paulo State University, Botucatu 18618-970, SP, Brazil; ³Department of Pathology, A.C. Camargo Cancer Center, São Paulo 01509-010, SP, Brazil; ⁴Epigenetics Group; International Agency for Research on Cancer (IARC), Lyon 69372, France; ⁵MRC Integrative Epidemiology Unit, University of Bristol, Bristol BS8 1TH, United Kingdom; ⁶Department of Head and Neck Surgery and Otorhinolaryngology, A.C. Camargo Cancer Center, São Paulo 01509-010, SP, Brazil; and ⁷Department of Clinical Genetics, Vejle Hospital and Institute of Regional Health Research, University of Southern Denmark, Vejle, 7100, Denmark

Context: Even though the majority of well-differentiated thyroid carcinoma (WDTC) is indolent, a number of cases display an aggressive behavior. Cumulative evidence suggests that the deregulation of DNA methylation has the potential to point out molecular markers associated with worse prognosis.

Objective: To identify a prognostic epigenetic signature in thyroid cancer.

Design: Genome-wide DNA methylation assays (450k platform, Illumina) were performed in a cohort of 50 nonneoplastic thyroid tissues (NTs), 17 benign thyroid lesions (BTLs), and 74 thyroid carcinomas (60 papillary, 8 follicular, 2 Hürthle cell, 1 poorly differentiated, and 3 anaplastic). A prognostic classifier for WDTC was developed *via* diagonal linear discriminant analysis. The results were compared with The Cancer Genome Atlas (TCGA) database.

Results: A specific epigenetic profile was detected according to each histological subtype. BTLs and follicular carcinomas showed a greater number of methylated CpG in comparison with NTs, whereas hypomethylation was predominant in papillary and undifferentiated carcinomas. A prognostic classifier based on 21 DNA methylation probes was able to predict poor outcome in patients with WDTC (sensitivity 63%, specificity 92% for internal data; sensitivity 64%, specificity 88% for TCGA data). High-risk score based on the classifier was considered an independent factor of poor outcome (Cox regression, $P < 0.001$).

Conclusions: The methylation profile of thyroid lesions exhibited a specific signature according to the histological subtype. A meaningful algorithm composed of 21 probes was capable of predicting the recurrence in WDTC. (*J Clin Endocrinol Metab* 102: 4089–4099, 2017)

ISSN Print 0021-972X ISSN Online 1945-7197

Printed in USA

This article has been published under the terms of the Creative Commons Attribution License (CC BY; <https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Copyright for this article is retained by the author(s).

Received 13 April 2017. Accepted 11 August 2017.

First Published Online 16 August 2017

*These authors contributed equally.

Abbreviations: ATC, anaplastic thyroid carcinoma; BTL, benign thyroid lesion; DLDA, diagonal linear discriminant analysis; FA, follicular adenoma; FTC, follicular thyroid carcinoma; GP, good prognosis; *Gai*, G-protein α -i; HCC, Hürthle cell carcinoma; LT, lymphocytic thyroiditis; NT, nonneoplastic thyroid tissue; PDTC, poorly differentiated thyroid carcinoma; PP, poor prognosis; PTC, papillary thyroid carcinoma; TC, thyroid cancer; TCGA, The Cancer Genome Atlas; WDTC, well-differentiated thyroid carcinoma.

Thyroid cancer (TC) is the most common endocrine cancer (1) and comprises a group of tumors with remarkably different features. Papillary thyroid carcinoma (PTC) and follicular thyroid carcinoma (FTC) are well-differentiated thyroid carcinomas (WDTCs) comprising 90% of all TC cases, whereas poorly differentiated thyroid carcinoma (PDTC) and anaplastic thyroid carcinoma (ATC) subtypes are extremely aggressive and less frequently diagnosed (2). Despite the indolent biological behavior of the well-differentiated carcinomas, these tumors occasionally give rise to less differentiated and more aggressive TCs (3). Moreover, a significant number of patients with WDTC develop locoregional recurrences or even radioiodine-resistant distant metastases (3).

DNA methylation in TC has been intensively studied, and several markers have been described. To date, the majority of these reports are based on the analysis of candidate genes, focusing mainly in altered methylation in CpG islands and their association with gene expression (4, 5).

Currently, a limited number of genome-wide methylation studies have been reported in TC, and most of them are focused on PTC. Rodríguez-Rodero *et al.* (6), using a low-density platform (Illumina Infinium HM 27K microarray) in eight samples (two PTC, two FTC, two medullary, and two ATC), reported that different methylation profiles were associated with distinct TC subtypes. In 2014, The Cancer Genome Atlas (TCGA) project reported the methylation profiles in a large cohort of PTC cases ($n = 496$) performed in a high-density platform (Illumina Infinium HM 450 array) (7). Blending different genomic approaches, the authors proposed a PTC classification based on the molecular subtypes and recognized potentially useful candidates to contribute to management of the disease.

Aberrant DNA methylation has emerged as an important mechanism involved in cancer progression. In PTC samples, the methylation pattern of *TIMP3*, *SLC5A8*, and *DAPK* was significantly associated with extrathyroidal invasion, lymph node metastasis, multifocality, advanced tumor stages, and *BRAF* mutation (8). In 2014, Wang *et al.* (9) showed that *RUNX3* site-specific hypermethylation significantly increased the risk of PTC recurrence. Using genome-wide methylation profiling (Illumina Infinium HM 27K platform) in well-differentiated thyroid tumors and validating the data in an independent series of cases, Mancikova *et al.* (10) identified *WT1* and *EI24* as markers for the risk of developing tumor recurrence. Overall, these findings indicate that methylation and prognosis have been poorly explored in TCs.

In this study, a comprehensive global methylation (Illumina Infinium HM 450K microarray) analysis was

performed in a large cohort of cases including nonneoplastic thyroid tissues (NTs), benign thyroid lesions (BTLs), and malignant thyroid lesions. Biomarkers and biological pathways epigenetically regulated involved in thyroid carcinogenesis were described. Subsequently, we developed a prognostic classifier capable of predicting the recurrence in patients with WDTC. Performance of the DNA methylation-based algorithm was tested against the TCGA database.

Materials and Methods

Patients

The samples were retrospectively obtained from patients who underwent surgery at A.C. Camargo Cancer Center, São Paulo, Brazil. The study was approved by the Institutional Ethics Committee (protocol no. 475.385). The patients were advised of the procedures and provided written informed consent. All FTC/Hürthle cell carcinoma (HCC) and PDTC/ATC samples available in the institutional tumor biobank were included, whereas the 60 PTC samples were selected from patients treated with total thyroidectomy followed by radioiodine therapy (to standardize the treatment) and with ≥ 5 years of follow-up. Ten FTC/HCC, four PDTC/ATC, eight follicular adenoma (FA), six nodular goiter, three lymphocytic thyroiditis (LT), and 50 surrounding normal tissue samples from patients with PTC were also included in this study. A subset of samples (41 PTCs matched with adjacent NTs) used in this analysis was previously reported (11). Clinical and histopathological data for the malignant lesions are summarized in the Table 1.

DNA extraction, bisulfite conversion, and DNA methylation analysis

Genomic DNA from fresh frozen thyroid tissues was extracted *via* the phenol-chloroform method and quantified by Qubit® dsDNA BR Assay (Qubit® 2.0 Fluorometer, Life Technologies, Carlsbad, CA). The DNA (500 ng) was bisulfite converted with an EZ DNA Methylation-Gold™ Kit (Zymo Research, Irvine, CA) according to the manufacturer's recommendations.

Genome-wide methylation assays were performed with the Human Methylation 450 BeadChip (Illumina, San Diego, CA). The data were captured in Illumina HiScan system, and the β values ranged from 0 (*unmethylated*) to 1 (*methylated*). R language was used for the methylation analysis, as previously described (12). The quality control parameters were followed, probes were filtered and normalized, and the batch effects were assessed. Briefly, cross-reactive probes (≥ 49 bases), single nucleotide polymorphisms (minor allele frequency $> 5\%$), sex-associated probes (13), low-quality probes ($P > 0.05$), and low bead count (< 3) in $\geq 5\%$ of samples were excluded with the WaterRmelon package. Data were normalized *via* the beta-mixture quantile normalization method (14). The *sva* package was used to correct the batch effects (15). The annotation was performed according to the Illumina (hg19) and HUGO Gene Nomenclature Committee.

Unsupervised hierarchical clustering analysis was performed as described in Supplemental Material with BRB Array Tools

Table 1. Clinicopathologic Characteristics of the Patients Enrolled in the Study

Characteristics	PTC ^a		Other Subtypes	
	n	%	n	%
Age				
<55 y	51	85	4	28.6
≥55 y	9	15	10	71.4
Sex				
Female	44	73.3	10	71.4
Male	16	26.7	4	28.6
Histology (excluded PTC)				
FTC	—	—	8	57.1
HCC	—	—	2	14.3
PDTC	—	—	1	7.1
ATC	—	—	3	21.4
Predominant variant (PTC only)				
Classic	47	78.3	—	—
Follicular	10	16.7	—	—
Other	3	5.0	—	—
Tumor dimension, cm				
Median (interval)	1.3 (0.5–5.5)	—	2.3 (0.9–13)	—
Micro-PTC (≤1 cm)	25	41.7	—	—
PTC (>1 cm)	35	58.3	—	—
Extrathyroidal extension				
No	30	52.6	7	53.8
Yes	27	47.4	6	46.2
Ni	3	—	1	—
Invasion				
Vascular	2	3.4	5	38.5
Lymphatic	3	5.1	1	7.7
Both	1	1.7	2	15.4
No	53	89.8	5	38.5
Ni	1	—	1	—
Perineural invasion				
No	42	93.3	10	23.1
Yes	3	6.7	3	76.9
Ni	15	—	1	—
Lymph node metastasis				
No (cN0, pN0)	32	53.3	9	69.2
Yes (pN1)	28	46.7	4	30.8
Ni	—	—	1	—
Clinical evolution				
Free of disease	44	73.3	5	71.4
Confirmed recurrence	16	26.7	2	28.6
Ni ^b	—	—	7	—
Death				
No	59	98.3	10	71.4
Yes	1	1.7	4	28.6
Follow-up (mo) ^c	83.8 (1.1–142.9)	—	63.3 ^d (4.0–139.4)	—

cN0, no clinical evidence of cancer in the regional lymph nodes; Ni, not informed; pN0, no pathological evidence of cancer in regional lymph nodes; pN1, pathological confirmation of cancer in regional lymph nodes.

^aPTC samples were enriched by more aggressive features and recurrence in follow-up as a consequence of the inclusion criteria (only patients submitted to total thyroidectomy followed by radioiodine therapy and patients without recurrence with ≥5 years of follow-up; all relapsed patients were enrolled).

^bThree FTCs had <5 years of follow-up with no recurrence, and 4 undifferentiated carcinomas showed <6 months of follow-up after the diagnosis.

^cData up to August 2014.

^dOnly for FTC samples.

version 4.4.0 software (Biometric Research Branch, National Cancer Institute). Differentially methylated loci were identified by comparing BTL, PTC, FTC/HTC, and PDTC/ATC samples with NT by using the limma package, adopting an adjusted $P < 0.05$

and mean delta beta ($\Delta\beta$) < -0.2 or $> +0.2$. Raw microarray and normalized data are available in Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) under specific accession number (GSE97466).

Pathway enrichment analysis

In silico pathway analysis was conducted with Ingenuity Pathway Analysis (IPA®, QIAGEN, Redwood City, CA, www.qiagen.com/ingenuity) and KOBAS (version 2.0; <http://kobas.cbi.pku.edu.cn/home.do>) software, comprising the list of differentially methylated probes obtained in all comparisons (BTL *vs* NT, PTC *vs* NT, FTC/HCC *vs* NT, and PDTC/ATC *vs* NT). Only probes mapped within the interval until 200 bp and 200 to 1500 bp upstream of the transcription start site and those in CpGs located within the 5'UTR and in the first exon were included.

DNA methylation-based algorithm to predict well-differentiated (PTC and FTC/HCC) TC outcomes

Differentially methylated probes with potential biological significance in more aggressive subtypes of thyroid malignancies were initially selected to design a prognostic classifier. This first step consisted in obtaining probes located in the promoter regions by comparing between PDTC/ATC and NT. WDTC (PTC and FTC/HCC) was grouped in poor prognosis (PP) and good prognosis (GP) cases. Patients with confirmed recurrent loco-regional disease (histological analysis) or distant metastasis (conclusive imaging test) in the follow-up were categorized as having WDTC-PP. The WDTC-GP group included only patients without any suspicion of active disease on clinical follow-up (normal serum thyroglobulin levels and no evidence of disease in the imaging screening) for ≥ 5 years (WDTC-GP: 43 PTC and 5 FTC). Because multifocal thyroid tumors are probably related to an independent clonal origin (16) and to obtain a more robust analysis, WDTC-PP exhibiting multifocal primary tumors was excluded (WDTC-PP: 6 PTC and 2 FTC).

The $\Delta\beta > 0.11$ was used as a criterion for the comparison between WDTC-PP, NT, and BTL, which allowed the filtering of probes exclusively altered in WDTC-PP and ATC/PDTC. The remaining probes were submitted to a diagonal linear discriminant analysis (DLDA) method. The classification performance was calculated with a leave-one-out cross-validation test (BRB Array Tools version 4.4.0). The predictive model used was based on scores stratified as low (NT range), intermediate (below DLDA threshold), and high (above DLDA threshold).

Application of the prognostic classifier in PTC samples from the TCGA database

TCGA database (<http://tcga-data.nci.nih.gov/tcga/>) was used to test the performance of the DNA methylation-based algorithm. Methylation (level 3, 450k Illumina platform; November 2015) and clinical data were obtained from TC samples, which comprised only patients with PTC. Clinical data were thoroughly evaluated under the same inclusion criteria previously defined in our study, resulting in 48 WDTC-GP and 14 WDTC-PP. The *in silico* data were used to test performance of the mathematical model previously established without any further adjustment.

Data processing and statistical analysis

Methylation quality control data and differential analysis were performed in R and Bioconductor packages (methyumi, sva, limma). The Bonferroni method was used to adjust *P* values for multiple comparisons. DNA methylation β -values were graphically represented in a heatmap generated by BRB ArrayTools. Disease-free survival analysis (SPSS version 21.0; Chicago, IL) was

performed *via* Kaplan–Meier estimator with log-rank test and multivariable analysis with Cox (proportional hazards) regression, adopting a two-tailed *P* < 0.05 value as significant.

Results

Global DNA methylation patterns unveils TC subtypes

DNA methylation profiles of 74 tumor samples, 50 surrounding NTs, and 17 BTLs were investigated. Unsupervised hierarchical clustering analysis revealed four clusters with pathological and molecular differences (Fig. 1A): cluster 1 was enriched with FAs, nodular goiters, and minimally invasive FTCs, showing more frequent hypermethylation events; cluster 2 displayed all NT samples; cluster 3 comprised exclusively PTC samples, characterized by an evident methylation loss; and cluster 4 exhibited all ATC/PDTC, lymphocytic chronic thyroiditis, and the remaining FTC/HCC (most extensively invasive). In contrast with the normal-like cluster (cluster 2), cluster 4 was enriched by alterations involved in immune response genes and clusters 1 and 3 by signal transduction-related genes (Supplemental Table 1).

The comparison of BTL, PTC, FTC/HCC, and PDTC/ATC *vs* NT ($|\Delta\beta| > 0.2$ and adjusted *P* < 0.05) revealed 1753 differentially methylated probes in BTL (222 hypomethylated and 1531 hypermethylated), 3015 in PTC (2773 hypomethylated and 242 hypermethylated), 5575 in FTC/HCC (1475 hypomethylated and 4100 hypermethylated), and 35,167 in PDTC/ATC (28,252 hypomethylated and 6195 hypermethylated) (Fig. 1B; Supplemental Table 2).

DNA methylation changes were found in all genomic regions. The follicular lesions presented hypomethylated and hypermethylated CpG sites closer to the transcription start sites in gene body and intergenic regions (Fig. 1C). Hypermethylation of CpG island regions was identified in all comparisons, whereas hypomethylation was more frequently observed outside the CpG islands (Fig. 1D). A total of 620 hypermethylated probes was exclusively found in BTL, 113 in PTC, 2244 in FTC/HCC, and 5411 in PDTC/ATC. Similarly, 95 hypomethylated probes were found in BTL, 1,996 in PTC, 465 in FTC/HCC, and 26,734 in PDTC/ATC (Supplemental Fig. 1).

Molecular pathways associated with differentially methylated promoter regions

In silico analysis using the list of promoter regions hypomethylated and hypermethylated obtained in the comparison between thyroid lesions and NTs revealed the involvement of different pathways according to the tumor subtype (Table 2; Supplemental Table 3). Activation of vitamin D receptors (VDRs) and retinoid X receptors (RXRs) was associated with FTC and PTC (IPA

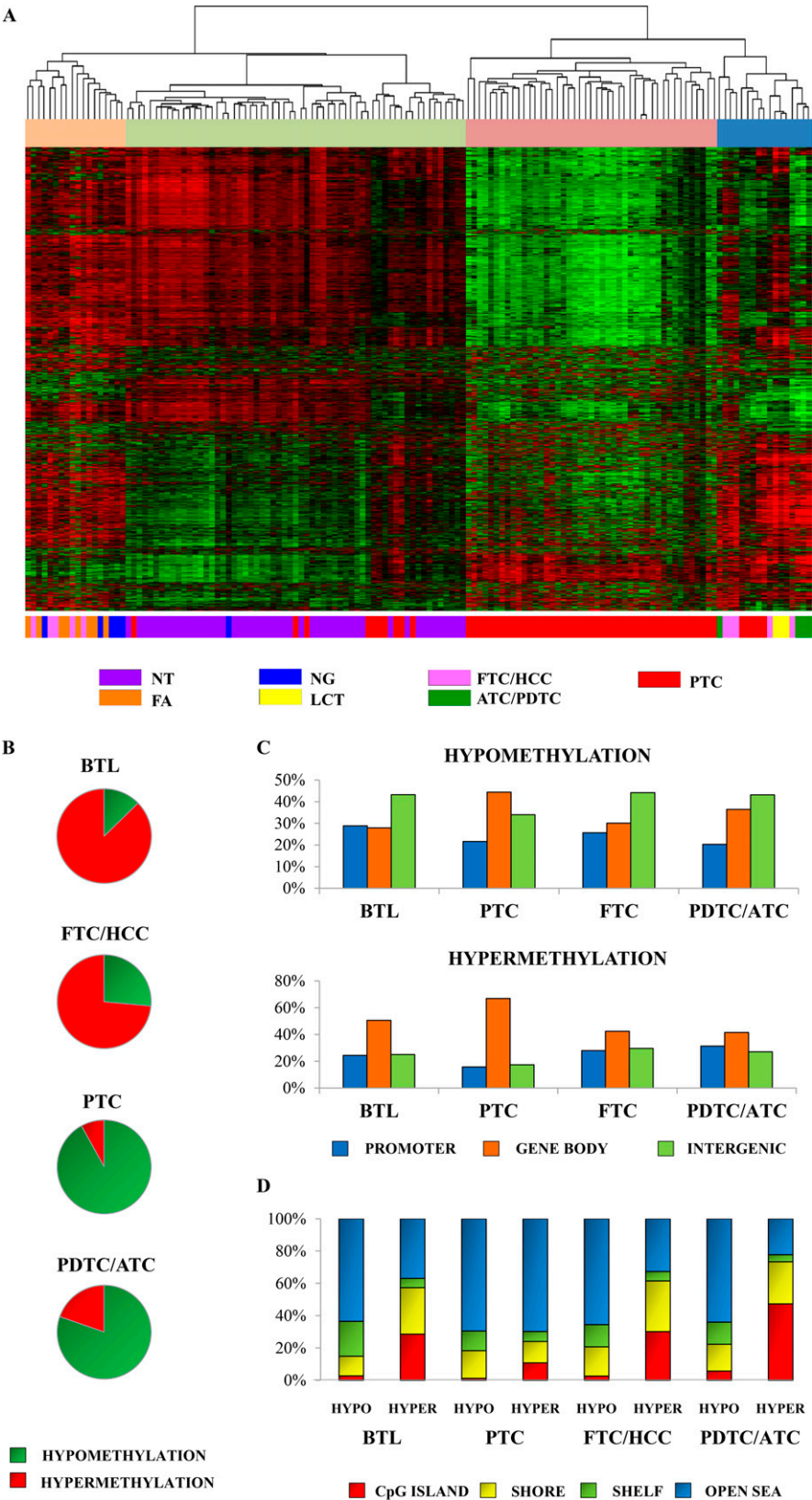


Figure 1. Comparison of CpG site methylation regions between benign lesions, TCs, and surrounding NTs. (A) Unsupervised hierarchical clustering analysis of the methylation profiles of thyroid tissues using the most variable probes within the platform (8016 probes). Four major groups were detected and defined by the most variable probes. The rows indicate the CpG sites, and the columns indicate the samples. (B) Percentage of hypermethylated (red) and hypomethylated (green) probes in BTL, PTC, FTC/HCC, and PDTC/ATC. (C) CpG site proportions by location relative to promoter, gene body, and intergenic region, according to Illumina 450 K annotation. (D) CpG site proportions by location relative to CpG island regions, according to Illumina 450 K annotation.

Table 2. Top 10 Significant Canonical Pathways Identified in the Comparisons by Ingenuity Pathway Analysis Software

Analysis	Pathway (Ingenuity Pathway Analysis)	Prediction State (z Score)	Genes Found/ Database	P (Fisher Exact Test)	P (Benjamini–Hochberg Correction)
BTL vs NT	Reelin signaling in neurons	—	6/79	0.002	0.438
	eNOS signaling	Inhibited (−1.9)	8/142	0.003	0.438
	Leukocyte extravasation signaling	Inhibited (−0.7)	9/198	0.006	0.443
	IL-4 signaling	—	5/76	0.009	0.443
	CCR3 signaling in eosinophils	—	6/117	0.014	0.443
	FAK signaling	—	5/87	0.015	0.443
	L-dopachrome biosynthesis	—	1/1	0.017	0.443
	RhoA signaling	Inhibited (−2.2)	6/122	0.017	0.443
	Amyotrophic lateral sclerosis signaling	—	5/98	0.025	0.443
	IL-12 signaling and production in macrophages	—	6/135	0.026	0.443
FTC/HCC vs NT	Inhibition of matrix metalloproteases	—	6/39	0.014	0.946
	Leukocyte extravasation signaling	Inhibited (−1.4)	18/198	0.015	0.946
	eNOS Signaling	Inhibited (−1.7)	14/142	0.016	0.946
	Tec kinase signaling	Inhibited (−1.9)	15/158	0.017	0.946
	HGF signaling	Inhibited (−2.1)	11/105	0.020	0.946
	Reelin signaling in neurons	—	9/79	0.021	0.946
	FAK signaling	—	9/87	0.036	0.946
	Fatty acid α -oxidation	—	3/16	0.047	0.946
	VDR/RXR activation	Activated (0.4)	8/78	0.049	0.946
	Factors promoting cardiogenesis in vertebrates	—	9/92	0.049	0.946
PTC vs NT	TREM1 signaling	Activated (2.8)	8/74	<0.001	0.044
	IL-10 signaling	—	7/68	<0.001	0.079
	RhoGDI signaling	Inhibited (−2.3)	10/173	0.003	0.192
	Signaling by Rho family GTPases	Activated (2.1)	12/234	0.003	0.192
	The visual cycle	—	3/15	0.003	0.192
	LXR/RXR activation	Inhibited (−2.1)	8/121	0.003	0.192
	VDR/RXR activation	Activated (2.2)	6/78	0.005	0.232
	LPS/IL-1 mediated inhibition of RXR function	Activated (1.0)	11/219	0.005	0.232
	Integrin signaling	Activated (3.0)	10/201	0.008	0.301
	Role of JAK1 and JAK3 in γ c cytokine signaling	—	5/63	0.009	0.301
PDTC/ATC vs NT	Agranulocyte adhesion and diapedesis	—	58/189	<0.001	0.001
	Atherosclerosis signaling	—	41/124	<0.001	0.002
	LXR/RXR activation	Inhibited (−1.0)	39/121	<0.001	0.004
	Transcriptional regulatory network in embryonic stem cells	—	18/40	<0.001	0.004
	FXR/RXR activation	—	39/127	<0.001	0.009
	Granulocyte adhesion and diapedesis	—	50/177	<0.001	0.009
	GABA receptor signaling	—	23/67	<0.001	0.031
	Estrogen biosynthesis	—	15/37	<0.001	0.033
	<i>Gai</i> signaling	Activated (2.8)	35/120	<0.001	0.033
	<i>Gas</i> signaling	Activated (1.0)	32/109	<0.001	0.045

$P < 0.05$), whereas the G-protein alpha-i (*Gai*) signaling pathway was identified as deregulated in ATC/PDTC (Supplemental Fig. 2, Supplemental Table 3) (adjusted $P < 0.05$).

Prognostic classifier performance in well-differentiated TC

An altered DNA methylation pattern in a specific locus has the potential to be a useful biomarker for the identification of aggressive TC subtypes. Based on this

premise, a prognostic classifier was designed with 7700 probes mapped in the promoter regions obtained from the comparison between PDTC/ATC and NT. A total of 21 probes showed $\Delta\beta < -0.1$ or $> +0.1$ comparing WDTC-PP with WDTC-GP cases, with 17 hypomethylated (*OR6K2*, *LAIR2*, *PFKFB2*, *OR4K15*, *NLRP11*, *OR9G4*, *THSD7B*, *OR2M3*, *OR52B2*, *OR2T6*, *FFAR2*, *TN3*, *OR52M1*, *DCD*, *ADGRE2*, *HRH1*, and *CXXC5*) and four hypermethylated (*GPR21*, *MBP*, *YPEL4*, and *ATP6V0C*) in the WDTC-PP group

(Table 3). These 21 probes were submitted to a DLDA to train a prognostic predictor classifier, which was able to distinguish WDTC-PP from WDTC-GP with 63% sensitivity and 92% specificity (63% sensitivity and 90% specificity in leave-one-out cross-validation). Performance according to stratification of the predictive model in low, intermediate, and high scores is shown in Fig. 2A.

Disease-free survival analysis revealed a higher risk of recurrence according to the score categories (Fig. 2B). Multivariable analysis showed that high score is an independent factor for lower disease-free survival ($P < 0.001$; odds ratio 31.0; 95% CI, 4.9 to 195.6) (Supplemental Table 4). The prognostic classifier found in our analysis was tested with the TCGA methylation data in PTC samples, confirming a similar predictive accuracy (64% sensitivity and 88% specificity) (Fig. 2C) and lower disease-free survival in cases with high scores (Fig. 2D). Contrasting the methylation definition of “high risk” with the clinical-pathological “high risk” classification (17), the latter presented lower sensitivity (21% to 38%) and slightly higher specificity (96% to 98%) (Supplemental Fig. 3). The individual clinical pathological data and algorithm scores of internal and external (TCGA) data sets are described in Supplemental Table 5.

Discussion

The thyroid samples derived from follicular cells revealed high methylation variability, as showed in the unsupervised hierarchical clustering analysis. From the four major groups exposed as results of this analysis, LT was

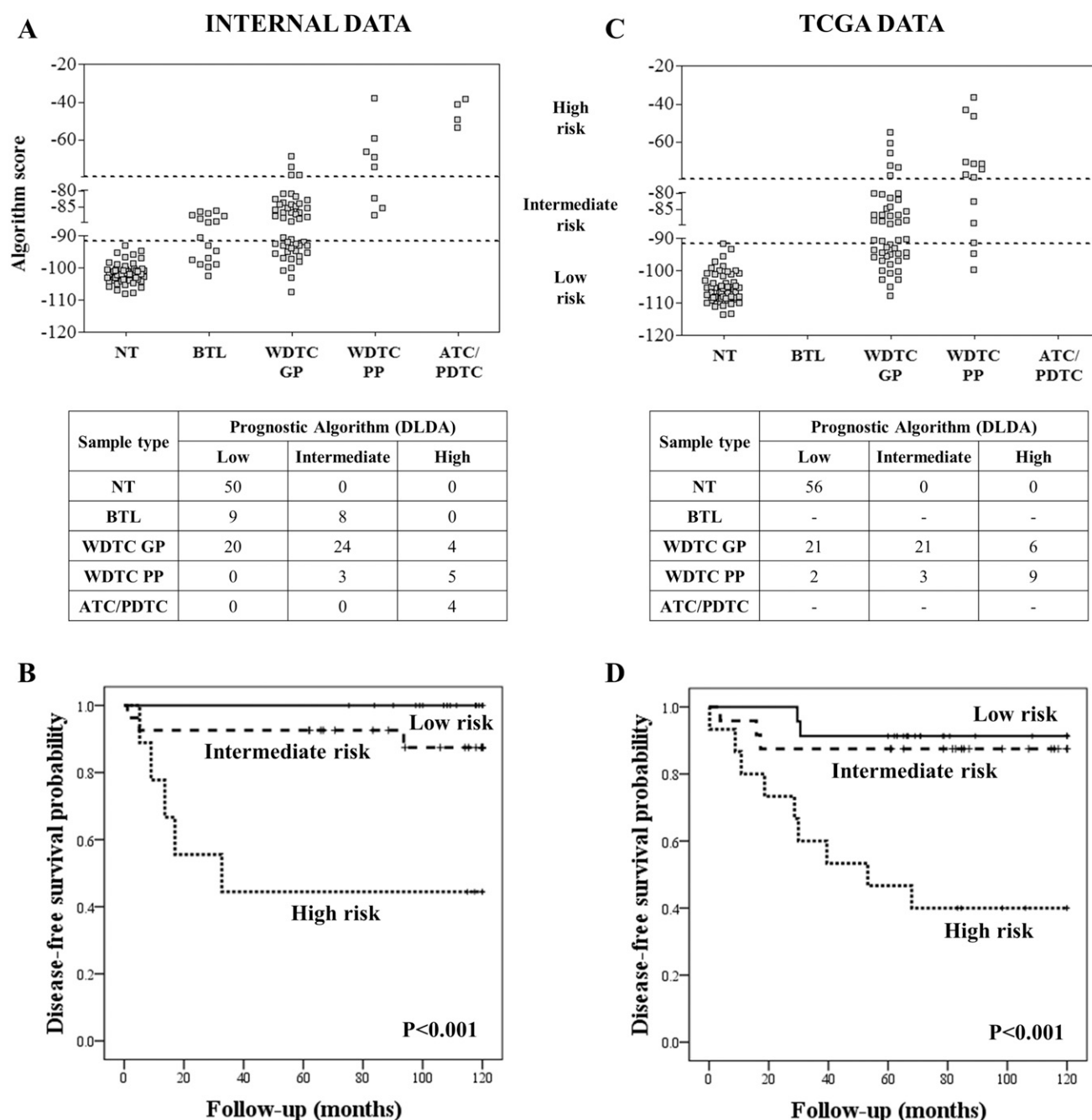
grouped with highly aggressive tumors including undifferentiated or poorly differentiated tumors and extensively invasive FTCs (cluster 4). Lymphocytic thyroiditis is a chronic autoimmune thyroid disease enriched with immune cell infiltration and inflammatory cytokines (18, 19), similar to highly proliferative and invasive thyroid tumors (20). Interestingly, an enrichment of genes related to immunological response was detected in cluster 4. Therefore, an important source of the methylation heterogeneity may be the inflammation response, as previously described in molecular profiling studies of thyroid tumor samples (21, 22).

The methylation profiles of the sample groups compared with NTs revealed that CpG site hypermethylation was more frequently observed in benign lesions and FTCs and hypomethylation in papillary and poorly differentiated or anaplastic carcinomas. These findings point out distinct epigenetic mechanisms related to the development or progression of thyroid lesions according to histological subtype.

Changes in the methylation pattern of benign lesions have also been reported (23), suggesting the relevance of this process at early stages of tumorigenesis. CpGs with methylation gains were more frequently observed in FTCs (4100) than in BTLs (1531) compared with NTs. Mancikova *et al.* (10) reported similar findings comparing gene promoter methylation profiles of FTCs and FAs. We also found aberrant methylation in gene body in both lesions, BTL and FTC. In addition to the morphological and genetic similarities described in adenomas (in our analysis, eight FAs were included in the BTL

Table 3. The Prognostic Classifier Comprising 21 Probes Included 4 Hypermethylated (DLDA Positive Values) and 17 Hypomethylated (DLDA Negative Values) Probes

Probe ID	Gene Symbol	Promoter Region	$\Delta\beta$ ATC/PDTC-NT	$\Delta\beta$ WDTC-PP-WDTC-GP	DLDA Score
cg00119186	OR6K2	1st exon	-0.411	-0.105	-16.72
cg08905487	LAI2	TSS200	-0.406	-0.134	-12.14
cg02710090	PFKFB2	TSS1500	-0.253	-0.109	-11.41
cg07531287	OR4K15	1st exon	-0.280	-0.112	-10.68
cg25730098	NLRP11	5'UTR	-0.321	-0.118	-10.65
cg03178489	OR9G4	1st exon	-0.307	-0.104	-9.33
cg08231096	THSD7B	TSS200	-0.352	-0.101	-9.23
cg07598464	OR2M3	1st exon	-0.388	-0.131	-8.41
cg21966764	OR52B2	TSS1500	-0.263	-0.110	-7.31
cg06846214	OR2T6	TSS1500	-0.337	-0.109	-6.38
cg20199836	FFAR2	TSS200	-0.369	-0.102	-6.06
cg17372806	RTN3	TSS1500	-0.330	-0.103	-5.81
cg03316101	OR52M1	1st exon	-0.309	-0.123	-5.03
cg01687040	DCD	TSS1500	-0.355	-0.103	-4.75
cg04103514	ADGRE2	TSS200	-0.379	-0.122	-4.70
cg06457736	HRH1	TSS200	-0.460	-0.120	-3.48
cg19628988	CXXC5	5'UTR	-0.362	-0.104	-3.28
cg14115756	GPR21	TSS1500	0.457	0.117	4.07
cg03560685	MBP	TSS1500	0.513	0.109	4.98
cg22705929	YPEL4	TSS1500	0.324	0.122	5.66
cg05884711	ATP6VOC	TSS1500	0.341	0.191	9.60



group) and FTCs (24), these entities also shared an elevated number of hypermethylated CpGs (479) (Supplemental Fig. 1).

PTCs displayed fewer aberrantly methylated genes compared with other carcinoma subtypes. In accordance with previous studies using methylation arrays in PTC samples, loss of methylation in both promoter and gene body regions was more frequently observed than hypermethylation (10, 25). Moreover, a high degree of loss of methylation was found in more aggressive

undifferentiated TCs, a similar finding previously described in acute myeloid leukemia and in breast cancer progression (26, 27). All histological subtypes of thyroid lesions exhibited hypermethylation in CpG island promoters as frequent event; hypomethylation in CpGs outside the promoter regions was also detected. These findings are in agreement with global methylation studies performed in TC (6, 10).

In silico pathway analysis *via* probes mapped in the promoter region revealed the activation of the *Gai*

pathway in poorly differentiated and undifferentiated carcinomas. These G proteins stimulate the kinase activity of cSRC protein (28), which is involved in a variety of cell functions, including survival, proliferation, and cell migration. In our study, *Gai*, *SHC*, and *RAS* were hypomethylated, and *RAP1GAP* (negative regulator of *Gai* pathway) was hypermethylated. Using gene expression analysis, we previously reported *RAP1GAP* down-expression in PTC samples (21). In thyroid cell lines, *RAP1GAP* downregulation was associated with migratory and invasive properties (29). Moreover, *RAP1GAP* promoter hypermethylation and down-regulation were reported in the most aggressive forms of TC (30). Our data give additional support to the involvement of the SRC upstream regulators in the prognosis of poorly differentiated and undifferentiated TCs.

Previously, we detected the VDR and RXR activation pathway as potentially deregulated in PTC samples (21), which is probably related to DNA methylation changes. In the current study, we demonstrated the predicted activation of this pathway in FTC samples but not in PDTC and ATC. These results are in agreement with high expression levels of VDR (31) and RXR (32) described in differentiated TC compared with normal tissues. Furthermore, high expression levels of VDR in patients with ATC are considered rare (31). These results reinforce that methylation is an important mechanism of gene regulation related to tumor progression, because of the importance of VDRs and RXRs in growth and cell differentiation.

Studies aiming to identify reliable markers associated with TC aggressiveness at the time of diagnosis can result in individualized treatment. Tumors harboring *BRAF* mutations have been associated with more aggressive phenotypes (3). Recently, *TERT* promoter mutations were associated with a worse prognosis, especially in *BRAF* mutated cases. Approximately 9% of PTCs were positive for *TERT* mutations (particularly C228T and C250T), whereas PDTCs and ATCs showed a higher frequency of mutations (40% and >70%, respectively) (33).

Methylation analysis has been described as a potent strategy to predict outcomes in several cancer subtypes. In PTCs, *RUNX3* hypermethylation (9) and *TSHR* hypomethylation (34) were reported as related to tumor recurrence. We found *RUNX3* hypermethylated in the promoter region (13 probes) in PDTC/ATC, but no association with recurrence was observed. Nonetheless, one probe (cg00117172) mapped in the nonpromoter region of *RUNX3* (body gene) was hypermethylated in both comparisons, PDTC/ATC *vs* NT ($\Delta\beta = 0.313$) and WDTC-PP *vs* WDTC-GP ($\Delta\beta = 0.102$), reinforcing their potential as biomarkers of tumor aggressiveness.

In this study, using rigorous inclusion criteria aiming to minimize confounding factors, we described a DNA

methylation model to predict WDTC outcome, in terms of recurrence. The model initially was designed based on the assumption that promoter regions differentially methylated in anaplastic and poorly differentiated tumors, which by definition present PP, could also be involved in the more aggressive forms of well-differentiated carcinomas. After filtering the initial list by comparing WDTC-PP with WDTC-GP, NT, and BTL, we designed a mathematical model based on 21 probes. Considering that LT and highly aggressive tumors presented similar methylation profiling, which may be explained by the strong influence of inflammation, the comparison of WDTC-PP with BTL was an essential step to avoid the prediction of inflammatory lesions as high-risk cases. In our cases, no significant differences were found considering the LT in the groups predicted by the DNA methylation prognostic model (low, 30%; intermediate, 19%; high, 33%; Fisher exact test $P = 0.541$).

From the probes included in the algorithm, 17 of 21 showed loss of methylation. Interestingly, 7 of 17 probes are mapped in the promoter regions of olfactory receptor genes. Olfactory receptors are G protein-coupled receptors, which have been reported as mediating tumor proliferation, invasiveness, and metastasis (35, 36). Although these genes are included among those described in the algorithm, their role in TC prognosis remains to be clarified.

Interestingly, high expression of *PFKFB2*, a hypomethylated gene found in the classifier, was reported as being associated with advanced TNM stage and also was implied in poor overall survival in hepatocellular carcinoma (37). In TC, *PFKFB2* overexpression was described in older patients (>40 years old) and associated with PP (38).

The methylation signature of 21 probes herein introduced was capable of discriminating 86% of the samples according to the outcome, demonstrating high specificity (90%) but lower sensitivity (63%). Therefore, more than half of WDTC-PP cases would be detected, with no further implications for overtreatment, which is a well-discussed issue in the management of patients with WDTC (39). A similar predictive power was observed when the algorithm was tested in an independent *in silico* data set of PTC samples (TCGA), which correctly classified 82% of the samples (64% sensitivity and 88% specificity).

Performance of the methylation algorithm was compared with *TERT* promoter mutation (C228T and C250T) with the TCGA data (data not shown). Fifty of the 62 PTC samples included in the TCGA cohort presented *TERT* promoter sequences. Four of these 50 cases were positive for the mutation: Two were classified by the methylation algorithm as intermediate (all disease-free)

and two as high risk (one disease-free and one relapsed case). In addition, the mutation assay was unable to identify relapsed patients with TC (8 of 9). Therefore, the methylation algorithm and *TERT* promoter mutation analysis showed no overlapping. The performance of our methylation algorithm was also compared with the American Thyroid Association guidelines (17), revealing a distinct set of high-risk cases. The high risk prediction based on molecular or clinical–pathological methods increased the sensitivity to 75% and specificity to 90% in the internal data set and 71% sensitivity and 88% specificity in the TCGA data set. These findings suggested that the methylation test would be feasible in clinical routines in combination with conventional stratification methods. A multiplexed methylation assay could be used to increase its applicability, similar to other tools described in the literature (40).

Conclusions

Thyroid carcinoma subtypes, which present the same cellular origin, showed different DNA methylation profiling. A prognostic classifier based on the methylation changes in a small number of loci was reported, which could correctly classify a subset of patients with WDTC with poor outcomes. Even though our study and the TCGA data revealed a limited number of cases classified as WDTC-PP, the rigorous criteria used to select the patients and the power to predict recurrence in both analyses reinforce that this classifier has the potential to be applied as a prognostic tool useful for making therapeutic decisions in TC.

Acknowledgments

The authors thank the A.C. Camargo Cancer Center Biobank for providing and processing the samples.

Financial Support: This study was supported by grants from the São Paulo Research Foundation (FAPESP 2015/20748-5) and the National Institute of Science and Technology in Oncogenomics (FAPESP 2008/57887-9 and CNPq 573589/08-9), Brazil. The students were awarded with fellowships from CNPq (CMB: 371497/2013-2, MBR: 140819/2011-8), Pro-Reitoria de Pós-Graduação–UNESP, SP, Brazil (MBR: 433), and the International Agency for Research on Cancer and the EC FP7 Marie Curie Actions–People–COFUND (S.A.). S.A. is supported by grants from Cancer Research UK (grant C18281/A19169) and the Medical Research Council, University of Bristol (grant MC_UU_12013/2). Z.H. is supported by grants from the Institut National du Cancer (France), the Fondation ARC pour la Recherche sur le Cancer (France), and the EC FP7 EurocanPlatform: A European Platform for Translational Cancer Research (grant 260791).

Correspondence and Reprint Requests: Silvia Regina Rogatto, PhD, Professor, Department of Clinical Genetics,

Kabbeltoft 25, 7100, Vejle, Denmark. E-mail: silvia.regina.rogatto@rsyd.dk and rogatto@fmb.unesp.br.

Disclosure Summary: The authors have nothing to disclose.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin*. 2015;65(1):5–29.
2. Elisei R, Pinchera A. Advances in the follow-up of differentiated or medullary thyroid cancer. *Nat Rev Endocrinol*. 2012;8(8):466–475.
3. Xing M. Molecular pathogenesis and mechanisms of thyroid cancer. *Nat Rev Cancer*. 2013;13(3):184–199.
4. Chen J, Liu C, Yin L, Zhang W. The tumor-promoting function of ECRG4 in papillary thyroid carcinoma and its related mechanism. *Tumour Biol*. 2015;36(2):1081–1089.
5. Lee EK, Chung KW, Yang SK, Park MJ, Min HS, Kim SW, Kang HS. DNA methylation of MAPK signal-inhibiting genes in papillary thyroid carcinoma. *Anticancer Res*. 2013;33(11):4833–4839.
6. Rodríguez-Rodero S, Fernández AF, Fernández-Morera JL, Castro-Santos P, Bayon GF, Ferrero C, Urdinguio RG, Gonzalez-Marquez R, Suarez C, Fernández-Vega I, Fresno Forcelledo MF, Martínez-Cambor P, Mancikova V, Castelblanco E, Perez M, Marrón PI, Mendiola M, Hardisson D, Santisteban P, Riesco-Eizaguirre G, Matías-Guiu X, Carnero A, Robledo M, Delgado-Álvarez E, Menéndez-Torre E, Fraga MF. DNA methylation signatures identify biologically distinct thyroid cancer subtypes. *J Clin Endocrinol Metab*. 2013;98(7):2811–2821.
7. Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*. 2014;159(3):676–690.
8. Hu S, Liu D, Tufano RP, Carson KA, Rosenbaum E, Cohen Y, Holt EH, Kiseljak-Vassiliades K, Rhoden KJ, Tolaney S, Condouris S, Tallini G, Westra WH, Umbricht CB, Zeiger MA, Califano JA, Vasko V, Xing M. Association of aberrant methylation of tumor suppressor genes with tumor aggressiveness and BRAF mutation in papillary thyroid cancer. *Int J Cancer*. 2006;119(10):2322–2329.
9. Wang D, Cui W, Wu X, Qu Y, Wang N, Shi B, Hou P. RUNX3 site-specific hypermethylation predicts papillary thyroid cancer recurrence. *Am J Cancer Res*. 2014;4(6):725–737.
10. Mancikova V, Buj R, Castelblanco E, Inglada-Pérez 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. DNA methylation profiling of well-differentiated thyroid cancer uncovers markers of recurrence free survival. *Int J Cancer*. 2014;135(3):598–610.
11. Beltrami CM, Dos Reis MB, Barros-Filho MC, Marchi FA, Kuasne H, Pinto CAL, Ambatipudi S, Herceg Z, Kowalski LP, Rogatto SR. Integrated data analysis reveals potential drivers and pathways disrupted by DNA methylation in papillary thyroid carcinomas. *Clin Epigenetics*. 2017;9:45.
12. Ambatipudi S, Cuenin C, Hernandez-Vargas H, Ghantous A, Le Calvez-Kelm F, Kaaks R, Barrdahl M, Boeing H, Aleksandrova K, Trichopoulos A, Lagiou P, Naska A, Palli D, Krogh V, Polidoro S, Tumino R, Panico S, Bueno-de-Mesquita B, Peeters PH, Quirós JR, Navarro C, Ardanaz E, Dorronsoro M, Key T, Vineis P, Murphy N, Riboli E, Romieu I, Herceg Z. Tobacco smoking-associated genome-wide DNA methylation changes in the EPIC study. *Epigenomics*. 2016;8(5):599–618.
13. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8(2):203–209.
14. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A beta-mixture quantile normalization

- method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*. 2013;29(2):189–196.
15. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28(6):882–883.
 16. Shattuck TM, Westra WH, Ladenson PW, Arnold A. Independent clonal origins of distinct tumor foci in multifocal papillary thyroid carcinoma. *N Engl J Med*. 2005;352(23):2406–2412.
 17. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL, Tuttle RM; American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer [published correction appears in *Thyroid*. 2010;20(8):942]. *Thyroid*. 2009;19(11):1167–1214.
 18. Ajjan RA, Weetman AP. The pathogenesis of Hashimoto's thyroiditis: further developments in our understanding. *Horm Metab Res*. 2015;47(10):702–710.
 19. Figueroa-Vega N, Alfonso-Pérez M, Benedicto I, Sánchez-Madrid F, González-Amaro R, Marazuela M. Increased circulating pro-inflammatory cytokines and Th17 lymphocytes in Hashimoto's thyroiditis. *J Clin Endocrinol Metab*. 2010;95(2):953–962.
 20. Ryder M, Ghossein RA, Ricarte-Filho JC, Knauf JA, Fagin JA. Increased density of tumor-associated macrophages is associated with decreased survival in advanced thyroid cancer. *Endocr Relat Cancer*. 2008;15(4):1069–1074.
 21. Barros-Filho MC, Marchi FA, Pinto CA, Rogatto SR, Kowalski LP. High diagnostic accuracy based on CLDN10, HMGA2, and LAMB3 transcripts in papillary thyroid carcinoma. *J Clin Endocrinol Metab*. 2015;100(6):E890–E899.
 22. Jarzab B, Wiench M, Fajurewicz K, Simek K, Jarzab M, Oczko-Wojciechowska M, Wloch J, Czarniecka A, Chmielik E, Lange D, Pawlaczek A, Szpak S, Gubala E, Swierniak A. Gene expression profile of papillary thyroid cancer: sources of variability and diagnostic implications. *Cancer Res*. 2005;65(4):1587–1597.
 23. Kim JH, Dhanasekaran SM, Prensner JR, Cao X, Robinson D, Kalyana-Sundaram S, Huang C, Shankar S, Jing X, Iyer M, Hu M, Sam L, Grasso C, Maher CA, Palanisamy N, Mehra R, Kominsky HD, Siddiqui J, Yu J, Qin ZS, Chinnaiyan AM. Deep sequencing reveals distinct patterns of DNA methylation in prostate cancer. *Genome Res*. 2011;21(7):1028–1041.
 24. Caria P, Vanni R. Cytogenetic and molecular events in adenoma and well-differentiated thyroid follicular-cell neoplasia. *Cancer Genet Cytogenet*. 2010;203(1):21–29.
 25. Ellis RJ, Wang Y, Stevenson HS, Bouffraquech M, Patel D, Nilubol N, Davis S, Edelman DC, Merino MJ, He M, Zhang L, Meltzer PS, Kebebew E. Genome-wide methylation patterns in papillary thyroid cancer are distinct based on histological subtype and tumor genotype. *J Clin Endocrinol Metab*. 2014;99(2):E329–E337.
 26. Barrio-Real L, Benedetti LG, Engel N, Tu Y, Cho S, Sukumar S, Kazanietz MG. Subtype-specific overexpression of the Rac-GEF P-REX1 in breast cancer is associated with promoter hypomethylation. *Breast Cancer Res*. 2014;16(5):441.
 27. Qian J, Chen Q, Yao DM, Yang L, Yang J, Wen XM, Zhang YY, Chai HY, Ma JC, Deng ZQ, Lin J. MOK overexpression is associated with promoter hypomethylation in patients with acute myeloid leukemia. *Int J Clin Exp Pathol*. 2015;8(1):127–136.
 28. Ma YC, Huang J, Ali S, Lowry W, Huang XY. Src tyrosine kinase is a novel direct effector of G proteins. *Cell*. 2000;102(5):635–646.
 29. Tsygankova OM, Prendergast GV, Puttaswamy K, Wang Y, Feldman MD, Wang H, Brose MS, Meinkoth JL. Downregulation of Rap1GAP contributes to Ras transformation. *Mol Cell Biol*. 2007;27(19):6647–6658.
 30. Zuo H, Gandhi M, Edreira MM, Hochbaum D, Nimgaonkar VL, Zhang P, Dipaola J, Evdokimova V, Altschuler DL, Nikiforov YE. Downregulation of Rap1GAP through epigenetic silencing and loss of heterozygosity promotes invasion and progression of thyroid tumors. *Cancer Res*. 2010;70(4):1389–1397.
 31. Clinckspoor I, Hauben E, Verlinden L, Van den Bruel A, Vanwallegheem L, Vander Poorten V, Delaere P, Mathieu C, Verstuyf A, Decallonne B. Altered expression of key players in vitamin D metabolism and signaling in malignant and benign thyroid tumors. *J Histochem Cytochem*. 2012;60(7):502–511.
 32. Haugen BR, Larson LL, Pugazhenth U, Hays WR, Klopper JP, Kramer CA, Sharma V. Retinoic acid and retinoid X receptors are differentially expressed in thyroid cancer and thyroid carcinoma cell lines and predict response to treatment with retinoids. *J Clin Endocrinol Metab*. 2004;89(1):272–280.
 33. Fagin JA, Wells SA, Jr. Biologic and clinical perspectives on thyroid cancer. *N Engl J Med*. 2016;375(11):1054–1067.
 34. Smith JA, Fan CY, Zou C, Bodenner D, Kokoska MS. Methylation status of genes in papillary thyroid carcinoma. *Arch Otolaryngol Head Neck Surg*. 2007;133(10):1006–1011.
 35. Sanz G, Leray I, Dewaele A, Sobilo J, Lerondel S, Bouet S, Grébert D, Monnerie R, Pajot-Augy E, Mir LM. Promotion of cancer cell invasiveness and metastasis emergence caused by olfactory receptor stimulation. *PLoS One*. 2014;9(1):e85110.
 36. Weng J, Wang J, Cai Y, Stafford LJ, Mitchell D, Ittmann M, Liu M. Increased expression of prostate-specific G-protein-coupled receptor in human prostate intraepithelial neoplasia and prostate cancers. *Int J Cancer*. 2005;113(5):811–818.
 37. Ji D, Lu ZT, Li YQ, Liang ZY, Zhang PF, Li C, Zhang JL, Zheng X, Yao YM. MACC1 expression correlates with PFKFB2 and survival in hepatocellular carcinoma. *Asian Pac J Cancer Prev*. 2014;15(2):999–1003.
 38. Vriens MR, Moses W, Weng J, Peng M, Griffin A, Bleyer A, Pollock BH, Indelicato DJ, Hwang J, Kebebew E. Clinical and molecular features of papillary thyroid cancer in adolescents and young adults. *Cancer*. 2011;117(2):259–267.
 39. Luster M, Weber T, Verburg FA. Differentiated thyroid cancer—personalized therapies to prevent overtreatment. *Nat Rev Endocrinol*. 2014;10(9):563–574.
 40. Moran S, Martínez-Cardús A, Sayols S, Musulén E, Balañá C, Estival-Gonzalez A, Moutinho C, Heyn H, Diaz-Lagares A, de Moura MC, Stella GM, Comoglio PM, Ruiz-Miró M, Matias-Guiu X, Pazo-Cid R, Antón A, Lopez-Lopez R, Soler G, Longo F, Guerra I, Fernandez S, Assenov Y, Plass C, Morales R, Carles J, Bowtell D, Mileskin L, Sia D, Tothill R, Tabernero J, Llovet JM, Esteller M. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol*. 2016;17(10):1386–1395.