PVALB, a New Hürthle Adenoma Diagnostic Marker Identified through Gene Expression

Janete M. Cerutti, Gisele Oler, Rosana Delcelo, Rene Gerardt, Pedro Michaluart, Jr., Sandro J. de Souza, Pedro A. F. Galante, Peng Huang, and Gregory J. Riggins

Genetic Bases of Thyroid Tumor Laboratory (J.M.C., G.O.) and Department of Pathology (R.D.), Universidade Federal de São Paulo, and Department of Pathology (R.G.) and Department of Surgery (P.M.), Universidade de São Paulo, São Paulo, SP 04039-032 Brazil; Laboratory of Computational Biology (S.J.S., P.A.F.G.), Ludwig Institute for Cancer Research, São Paulo Branch, São Paulo, SP 01323-903 Brazil; and Oncology Biostatistics (P.H.) and Department of Neurosurgery (G.J.R.), Johns Hopkins University School of Medicine, Baltimore, Maryland 21231

Context: A better means to accurately identify malignant thyroid nodules and to distinguish them from benign tumors is needed. We previously identified markers for detecting thyroid malignancy, with sensitivity estimated at or close to 100%. One lingering problem with these markers was that false positives occurred with Hürthle cell adenomas (HCA) which lowered test specificity.

Methods: To locate accurate diagnostic markers, we profiled in depth the transcripts of a HCA and a Hürthle cell carcinoma (HCC). From 1146 differentially expressed genes, 18 transcripts specifically expressed in HCA were tested by quantitative PCR in a wide range of thyroid tumors (n = 76). Sensibility and specificity were calculated using receiver operating characteristic (ROC). Selected markers were further validated in an independent set of thyroid tumors (n = 82) by immunohistochemistry. To define the panel that would yield best diagnostic accuracy, these markers were tested in combination with our previous identified markers.

Results: Seventeen of the 18 genes showed statistical significance based on a mean relative level of expression ($P < 0.05$). KLK1 (sensitivity = 0.97) and PVALB (sensitivity = 0.94) were the best candidate markers. The combination of PVALB and C1orf24 increased specificity to >97% and maintained sensitivity for detection of carcinoma.

Conclusion: We identified tumor markers that can be used in combination for a more accurate preoperative diagnosis of thyroid nodules and for postoperative diagnosis of thyroid carcinoma in tumor sections. This improved test would help physicians rapidly focus treatment on true malignancies and avoid unnecessary treatment of benign tumors, simultaneously improving medical care and reducing costs. **(*J Clin Endocrinol Metab* 96: E151–E160, 2011)**

The detection of a new thyroid nodule is a frequent clinical occurrence. The reported prevalence of thyroid nodules depends on the population and screening method (1). By palpation the estimated prevalence is 3–7% of the general population. With the increased use of ultrasound, the number of thyroid nodules diagnosed is on the rise and are found in up to 67% of the population (2). These nonpalpable nodules cannot be overlooked because they might represent a thyroid malignancy. The increasing rate by which thyroid nodules are found in the clinic has brought further attention to the current inadequate means for efficient and accurate diagnosis of thyroid nodules.

Currently, fine-needle aspiration (FNA) cytology is the most reliable, widely used, and cost-effective preoperative

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

doi: 10.1210/jc.2010-1318 Received June 10, 2010. Accepted September 7, 2010.
First Published Online October 6, 2010

Abbreviations: FNA, Fine-needle aspiration; FTA, follicular thyroid adenoma; FTC, follicular thyroid carcinoma; FVPTC, follicular variant of papillary thyroid carcinomas; HCA, Hürthle cell adenoma; HCC, Hürthle cell carcinoma; IHC, immunohistochemical; qPCR, quantitative PCR; RE, relative expression; ROC, receiver operating characteristic; SAGE, Serial Analysis of Gene Expression.

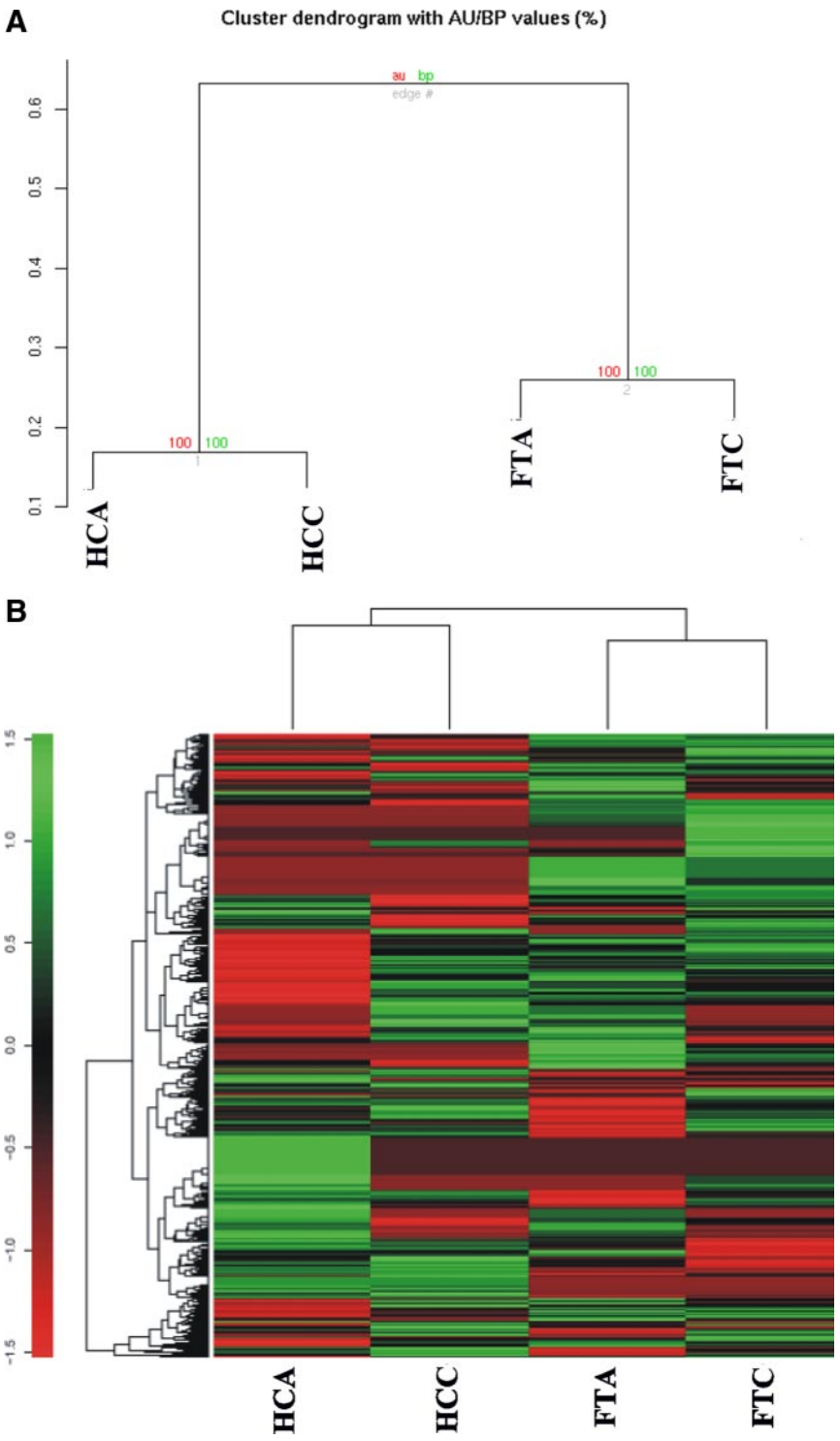

FIG. 1. Thyroid SAGE data analysis. A, Dendogram of four short thyroid libraries. All transcripts presenting reliable (not ambiguous) tags have been used to construct the dendrogram. Genes from thyroid libraries contain less than 2% of ambiguous tags. B, The clustering algorithm delineates patterns in the expression of all tags among all four libraries. Each *row* represents a tag, whereas each *column* corresponds to a SAGE library sample. The absolute abundance of the SAGE tag in the library (SAGE tag number) correlates with the intensity of the color (*green*, not present; *intense red*, highly abundant). This analysis reflects the relationship between the libraries based on their gene expression levels. AU, Approximately unbiased; BP, bootstrap probability.

test for initial evaluation of thyroid nodules (1). Although FNA leads to a better selection of patients for surgery, its accuracy is closely related to the histologic type. For follicular lesions FNA is not accurate, because the benign follicular thyroid adenoma (FTA) and the malignant follicular thyroid carcinoma (FTC) have similar cytological appearance. A definitive diagnosis of FTC currently requires demonstration of capsular or vascular invasion on final histology. In addition, FNA cannot accurately distinguish between benign and malignant Hürthle tumors (3). The current WHO classification describes Hürthle neoplasm as variant of follicular neoplasm and therefore the criteria for malignancy are solely dependent upon histological demonstration of capsular and/or vascular invasion or presence of lymph node or distant metastasis. This has several clinical implications because type and the extent of the surgical procedure as well as follow-up therapy cannot be planned based on FNA diagnosis.

Using immunohistochemistry, we previously demonstrated that the combination of four antibody markers distinguished a wider variety of thyroid tumors that are commonly classified as indeterminate on FNA with an estimated sensitivity of 100% and specificity of 85% for detecting malignancy (4, 5). Although this panel of markers was able to achieve a 100% detection of malignant samples, we encountered false positives from benign Hürthle cell adenoma (HCA), which lowered the test specificity. After we identified C1orf24 as one of the thyroid carcinoma markers with false positive in HCAs, others also reported rare and weakly *C1orf24*-positive thyroid cells within benign lesions, most of which were Hürthle cells (6). Because HCAs stained positive with our carcinoma markers, identifying these tumors in a combined test would improve our overall test accuracy.

In the present study we determined the gene expression profile of an HCA

and a Hürthle cell carcinoma (HCC) in depth using Serial Analysis of Gene Expression (SAGE). We report a two-gene test that can distinguish benign from malignant thyroid tumors with high accuracy and that might form the basis of a simple clinical test. An additional implication of our findings is that the pattern of transcription that distinguishes Hürthle tumors may provide biologic insights regarding their origin and open new perspective for cancer target therapy.

Patients and Methods

Generation of SAGE libraries

HCA (n = 1) and HCC (n = 1) samples were analyzed using SAGE. The libraries were constructed using a longSAGE procedure (7) and were sequenced through the SAGE portion of the Cancer Genome Anatomy Project (8). Tags were extracted from sequence text files and processed to remove duplicate ditags, linker sequences, and repetitive tags using SAGE 2002 software version 4.12 (available at http://www.sagenet.org).

TABLE 1. Candidate HCA biomarkers identified by SAGE and selected for validation by qPCR

Tag sequence	HCA[a]	HCC[a]	FTC[a]	Transcript description[b]	GenBank accession no.	Locus	Gene ontology[c]
CTGGACAAGGACAAAAG	89	0	0	*PVALB*, Parvalbumin	NM_002854	22q12-q13.1	Calcium ion binding
GGGCTACGTCCCTTGTG	51	0	0	*KLK1*, Kallikrein 1	BC005313	19q13.3	Tissue kallikrein activity
CTGAGCGCCGAGTACCC	32	0	0	*OBSCN*, Obscurin, cytoskeletal calmodulin	NM_052843	1q42.13	Protein-tyrosine kinase activity
TCTGCATCTTGGACGCC	23	0	0	*ALDH1L1*, Aldehyde dehydrogenase 1 family member L1	BC027241	3q21.2	Methyltransferase activity
TTCTGCGCCGGCGGAGG	23	0	0	*KLK4*, Kallikrein 4 (prostase enamel matrix prostate)	NM_004917	19q13.41	Peptidase activity
TTGGGTATCCTACTACT	20	0	0	*GFPT1*, Glutamine-fructose-6-phosphate transaminase 1	NM_002056	2p13	Transferase activity
TACCCTCTCAAGCCCAG	18	0	0	*ATP1A3*, ATPase Na+/K^+ transporting α 3 polypeptide	NM_152296	19q13.31	ATP binding
AACTCTTTTCCAGTGTC	18	0	0	*NUDT22*, Nudix (nucleoside diphosphate linked moiety X-type) motif 22	NM_005528	11q13	Heat shock protein binding
CCCTTGAATTCCTTCCT	15	0	1	*TSPAN33*, Tetraspanin 33	NM_005724	15q24.3	Development, activation, growth & motility
CTAAGCTTCTCCAGAGA	15	0	1	*SMARCD3*, SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily d member 3	NM_003078	7q35-q36	Transcription regulator activity
GGCGCCAGCGGGTCCAA	12	0	0	*TUSC2*, Tumor suppressor candidate 2	NM_007275	3p21.3	Negative regulation of cell cycle
TCCCAGTGTTTTTTTTG	12	0	0	*TBC1D1*, (tre-2/USP6 BUB2 cdc16) domain family member 1	NM_015173	4p14	GTPase activator activity
GGATTTGGAGTTAGGTG	12	0	0	*DVL3*, Dishevelled dsh homolog 3 (Drosophila)	NM_004423	3q27	Signal transducer activity
CTAACTTCGTTTGTGCG	12	0	0	*ENAH*, Enabled homolog (Drosophila)	NM_018212	1q42.12	Actin binding
TGCCAACTTTTCTGTTT	9	0	0	*LOC389833*, Similar to hypothetical protein MGC27019	NM_001033515	Unknown	Microtubule cytoskeleton organization and biogenesis
TCACTCCTGGCGACCCT	9	0	0	*ATF5*, Activating transcription factor 5	BC005174	19q13.3	Transcription factor activity
TCTTGTCATACAAATTT	9	0	0	*CHUK*, Conserved helix-loop-helix ubiquitous kinase	NM_001278	10q24-q25	Protein Serine/threonine kinase activity
TATATCTATTGTATGCT	9	0	0	*SH3BGRL*, SH3 domain binding glutamic acid-rich protein like	NM_003022	Xq13.3	SH3 domain binding

[a] SAGE tags counts shown in each column refer to the abundance of SAGE tags in the libraries after normalization to 200,000 total tags. Sage libraries are posted at http://cgap.nci.nih.gov/SAGE.

[b] Transcript description refers to the gene name to which tag was attributed, according to the HUGO/GDB nomenclature committee approved symbols.

[c] Gene classification was by molecular function at http://cgap.nci.nih.gov/Genes/AllAboutGO.

Tag to gene mapping

To provide tag to gene links, we developed Perl scripts to map all experimental tags against the nonredundant tag list downloaded from ACTG web servers and against all virtual tags (a computer prediction of tags produced by SAGE experiment) extracted from human genome (9). These databases were selected because the nonredundant tag list contains virtual tags from all most important cDNAs available from the public domain and the virtual tags from the human genome may contain tags from transcripts not yet identified.

Selection of differentially expressed tags

We used a Monte Carlo analysis to identify differentially expressed tags between HCA and HCC libraries. Tag numbers were normalized to 200,000 tags per library. The null hypothesis was that the abundance level of transcripts was the same among the SAGE libraries. For each tag, we performed 10,000 simulations to determine the likelihood, due to chance alone (*P* value), of obtaining a difference in the expression equal to or greater than the observed difference, given the null hypothesis. The *P* value threshold used for all the analyses was 0.001. A similar method was used by Zhang *et al.* (10).

Conversion of longSAGE tag into shortSAGE tag

To be able to access the level of expression of new HCA candidate markers in the previous described FTA and FTC libraries (available at http://cgap.nci.nih.gov/SAGE) we essentially converted the longSAGE tag into a shortSAGE tag by using only the first ten nucleotides adjacent to the *Nla*III site. If after the conversion two different long SAGE tags link to the same short SAGE, the respective frequencies were summed.

Hierarchical clustering

SAGE is characterized by the presence of a large amount of the data. Grouping SAGE tags into expression clusters would allow visualizing basic relationships and, consequently, to detect similarities between the libraries generated from Hürthle and follicular tumor subtypes. Therefore, the data from two longSAGE libraries converted into shortSAGE (HCA and HCC) and two libraries for shortSAGE (FTA and FTC) were used to perform a hierarchical clustering analysis with Pvclus (11), a package for the statistical software R. The clustering method was based on the pair-wise gene expression distance between the four shortSAGE libraries and the distance was defined as the average linkage. All genes presenting reliable (not ambiguous) tags have been used to hierarchical clustering analysis.

Tissue samples for validation of HCA candidate markers

For the validation by quantitative PCR analysis, specimens were obtained from patients, with their informed consent, undergoing thyroid surgery at Hospital São Paulo, Universidade Federal de São Paulo, and Hospital das Clínicas, Universidade de São Paulo. Tissue specimens were frozen immediately after surgical biopsy and stored at −80 C until use. Final histological classifications were obtained from paraffin-embedded sections and comprised 32 benign lesions, 34 malignant tumors, and 10 normal thyroid tissues. The benign lesions included 13 HCA and 19 FTA. The malignant lesions comprised 10 HCC, 14 FTC, and 10 follicular variant of papillary thyroid carcinomas (FVPTC). The study of patient materials was conducted according to principles detailed on the running Declaration of Helsinki at the time when tissues were obtained.

RNA isolation, cDNA synthesis, and quantitative PCR (qPCR)

Total RNA was isolated using Trizol reagents (Invitrogen Corp., Carlsbad, CA) according to the manufacturer's recommendations. Total RNA (1 μg) was treated with Dnase (Ambion, Austin, TX) and was reverse-transcribed to cDNA using SuperScript II Reverse Transcriptase kit with oligo $(dT)_{12-18}$ primer and 10 U of RNase inhibitor (Invitrogen Corp). An aliquot of cDNA (2 μL) was used in 20-μl PCR reactions containing TaqMan universal PCR master mix, 10 μM of each specific primer and FAM-labeled probes for the target genes or VIC-labeled probe for reference gene (*RS8*) (TaqManGene Assays on Demand; Applied Biosystems, Foster City, CA). qPCR reactions were performed in triplicate, the threshold

TABLE 2. Cut-off point and estimates for sensitivity and specificity of RE values determined by qPCR

Gene	Area under the ROC curve	Cut-off[a]	Number of true carcinoma	Sensitivity	CI sensitivity	Number of true HCA	Specificity	CI specificity
1. *ALDH1L1*	0.941	0.6199	30	0.88	(0.69-0.96)	12	0.92	(0.62-1.00)
2. *TUSC2*	0.787	0.7099	23	0.68	(0.43–0.83)	12	0.92	(0.62–1.00)
3. *ATF5*	0.903	0.6499	23	0.68	(0.43–0.83)	12	0.92	(0.62–1.00)
4. *OBSCN*	0.781	0.8799	25	0.74	(0.54–0.90)	12	0.92	(0.62–1.00)
5. *KLK4*	0.889	0.1999	23	0.68	(0.43–0.83)	12	0.92	(0.62–1.00)
6. *KLK1*	0.971	0.9999	33	0.97	(0.84–1.00)	12	0.92	(0.62–1.00)
7. *CHUK*	0.753	0.3699	21	0.62	(0.38–0.82)	12	0.92	(0.62–1.00)
8. *GFTP1*	0.787	0.6099	20	0.59	(0.36–0.81)	12	0.92	(0.62–1.00)
9. *DVL3*	0.833	0.6899	18	0.53	(0.30–0.78)	11	0.85	(0.49–0.98)
10. *PVALB*	0.894	1.3899	32	0.94	(0.80–0.0.99)	12	0.92	(0.62–1.00)
11. *SH3BGRL*	0.842	0.6999	24	0.71	(0.49–0.87)	12	0.92	(0.62–1.00)
12. *SMARCD3*	0.810	0.6099	19	0.56	(0.33–0.79)	12	0.92	(0.62–1.00)
13. *TBC1D1*	0.873	0.3199	18	0.53	(0.30–0.78)	12	0.92	(0.62–1.00)
14. *ATP1A3*	0.790	0.0319	8	0.24	(0.03–0.65)	11	0.85	(0.49–0.98)
15. *NUDT22*	0.882	1.0199	27	0.79	(0.58–0.92)	12	0.92	(0.62–1.00)
16. *TSPAN3*	0.869	1.5199	25	0.74	(0.54–0.90)	12	0.92	(0.62–1.00)
17. *ENAH*	0.735	0.3299	13	0.38	(0.14–0.68)	12	0.92	(0.62–1.00)

LOC389833 was not included in this step ($P > 0.05$). CI, Confidence interval. Comparison of malignant (CCH, FTC, and FVPTC; $n = 34$) and HCA ($n = 13$).

[a] Under this value, it was classified as malignant.

cycle (Ct) was obtained using Applied Biosystem software and were averaged (SD ≤ 1). Relative expression (RE) level was calculated formula as previously described (12, 13).

Immunohistochemical (IHC) analysis

To validate the results obtained from qPCR analysis and to define the final panel of markers, an independent set of 82 thyroid paraffin-embedded sections was tested. The samples were selected from the archives of the Department of Pathology, Universidade Federal São Paulo and Hospital das Clínicas, Universidade de São Paulo. The slides were incubated with primary antibodies followed by incubation with the labeled polymer Dako EnVision System (Dako laboratories, Carpinteria, CA). Monoclonal anti-Parvalbumin (Sigma-Aldrich, Munich, Germany) was used at a dilution of 1:1000. Anti-KLK1 antibody (Proteintech Group, Inc. Chicago, IL) was used at a dilution of 1:100. Control sections were processed in parallel with rabbit or mouse IgG used at the same concentration as the first antibody. Cells were positive (+) when more than 10% of cells were immunoreactive. All slides were scored in a blinded fashion and evaluated independently by three investigators.

Statistical analysis methods

For the qPCR data, we determined whether the expression values for 18 transcripts differed between HCA (n = 13) and malignant (n = 34) thyroid samples. Because neither qPCR values nor their log transformation were normally distributed according to Kolmogorov-Smirnov test, we used nonparametric Mann-Whitney-Wilcoxon test for comparison between groups. All results with P values of <0.05 were reported. Sensibility and specificity were calculated using a cut-off for the relative gene expression using receiver operating characteristic (ROC). For IHC analysis, we first used classification trees to identify potential markers for FTA, HCA, and malignant tumors, respectively. The combinations of some biomarkers subset were predefined while others were not. Because classification tree generally suffers from large variation, we used logistic regression to look for biomarker combinations that give larger area under the ROC. We partitioned the high dimensional sampling space of these markers and identify the subspaces with 100% sensitivity of malignant tumor diagnosis or high sensitivity and specificity for FTA tumor and HCA tumor diagnosis. Results from regression models and classification trees are then combined to obtain the final robust diagnosis tests.

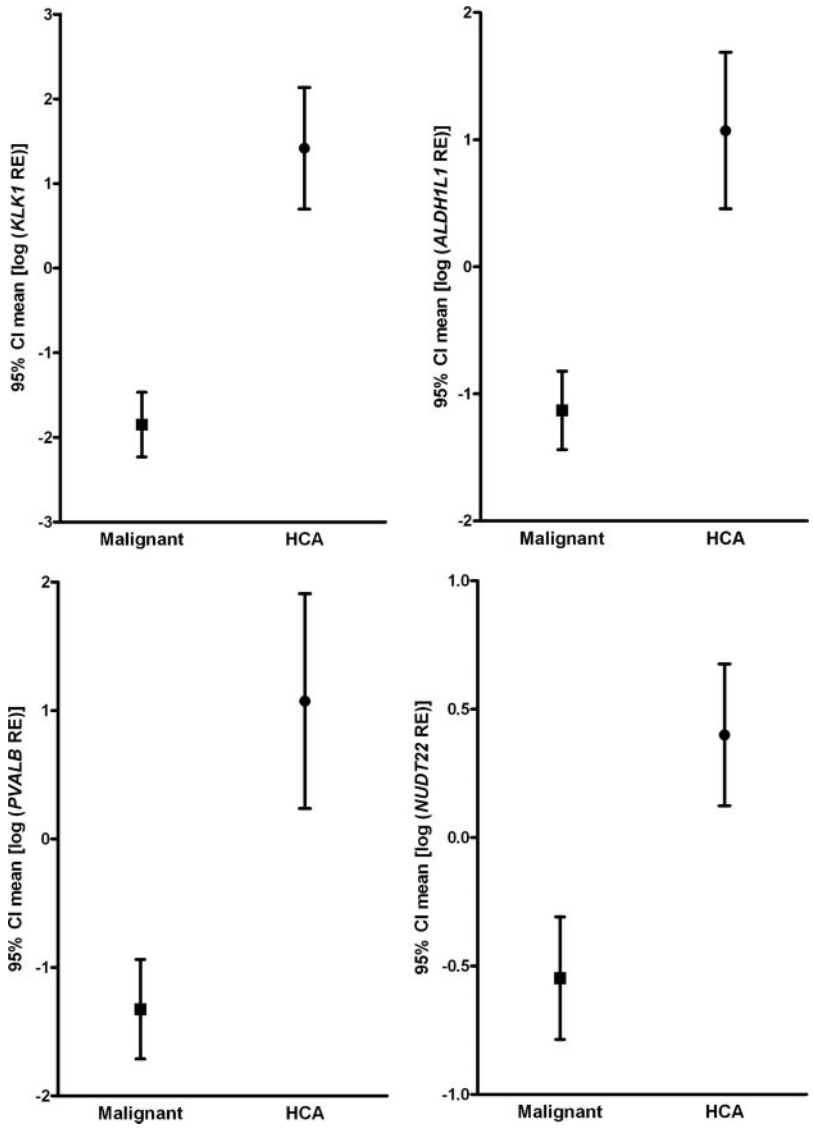

FIG. 2. Graphs showing results of qPCR for transcripts evaluated in the HCA *vs.* malignant groups. The results are expressed as mean of log transformed RE values [95% confidence interval (CI)]. The HCA group included 13 adenomas. The malignant group included 10 HCC, 14 FTC, and 10 follicular variant of thyroid carcinomas. As expected, transcripts are expressed in HCA while their expression is reduced (or absent) in most of malignant tumors. Of these transcripts, *PVALB* and *KLK1* distinguish HCA and malignant class with higher specificity and sensitivity, 0.94 and 0.97, respectively (as detailed in Table 2).

Results

Comparative analysis from SAGE libraries

A total of 143,203 longSAGE tags were obtained. The number of SAGE tags per library were 70,338 (HCA) and 72,865 (HCC). Monte Carlo simulations yielded 1,146 tags overexpressed in HCA and 1,401 were over-expressed in HCC libraries (with P values of ≤ 0.001).

Hierarchical clustering

Data were also analyzed using a clustering algorithm to delineate patterns of gene expression among all four libraries (HCA, HCC, FTA, and FTC SAGE libraries). A visual representation of hierarchical clustering of thyroid libraries was based on their pair-wise distance ($d = 1 - r$), using the hierarchical clustering algorithm from R environment (in

particular with the function pvclust). Values at branches are approximately unbiased *P* values (*red*) and bootstrap probability *P* value (*green*) (Fig. 1A). All genes that had reliable (not ambiguous) tags have been used to construct the dendrogram. Ambiguous tags were excluded as they may represent a sum of the expression of two or more genes. Genes from thyroid libraries contain less than 2% of these ambiguous tags. The pattern of expression is distinct among the follicular tumors and Hürthle tumors as they appear in a separate branch in the dendrogram. The SAGE data were also analyzed by a clustering algorithm to delineate patterns in the expression of all tags among all four libraries (Fig. 1B). Each row represents a tag, whereas each column corresponds to a SAGE library sample. The absolute abundance of the SAGE tag in the library (SAGE tag number) correlates with the intensity of the color (*green*, not present; *intense red*, highly abundant). Expression patterns across different libraries suggest a distinct gene expression profiling.

Selection of candidate markers for HCA

To identify potential HCA markers of 1,146 overexpressed in HCA, those with the greatest fold-induction in HCA and not expressed in HCC were assessed in FTA and FTC shortSAGE libraries.

As predicted from the literature and from our SAGE analysis, very few tags were expressed in both FTA and HCA and absent in malignant tumors (Fig. 1). Therefore, we focus on those transcripts highly expressed in HCA and not expressed in HCC and FTC libraries, because they had better potential to improve our test accuracy. Eighteen transcripts (Table 1) were chosen for validation. As the first step, we compared the means of the RE values for each transcript. The mean of RE values in the HCA was significantly higher than in malignant tumors for 17 of 18 transcripts, as predicted by SAGE ($P < 0.01$, Mann-Whitney-Wilcoxon test). The greater difference in expression was observed for *ALDH1L1*, *ATF5*, *KLK4*, *KLK1*, *DVL3*, *PVALB*, *SH3BGRL*, *TBC1D1*, *NUDT22*, and *TSPAN3* ($P < 0.001$) (Supplemental Table 1 published on The Endocrine Society's Journals Online web site at http://jcem.endojournals.org). All transcripts with $P < 0.05$ were used in the second step.

In second step, an ROC curve, graphical representation of the trade off between the false negative and false positive rates for every possible cut off, was generated. Cut-off levels with the optimum diagnostic efficiency derived from the ROC curves are described in Table 2. Among the genes, the observed specificity was 0.92 (1 of 13 false negative) for 15 genes. Therefore, the best predictive markers were those with highest sensitivities. *KLK1* (sensitivity = 0.97; area under ROC curve = 0.971) and *PVALB* (sensitivity = 0.94 and area under ROC curve = 0.894) were considered as the best candidate markers and were tested by IHC analysis (Supplemental Fig. 1). The mean for *ALDH1L1*, *KLK1*, *PVALB*, and *NUDT22*, which had higher sensitivity

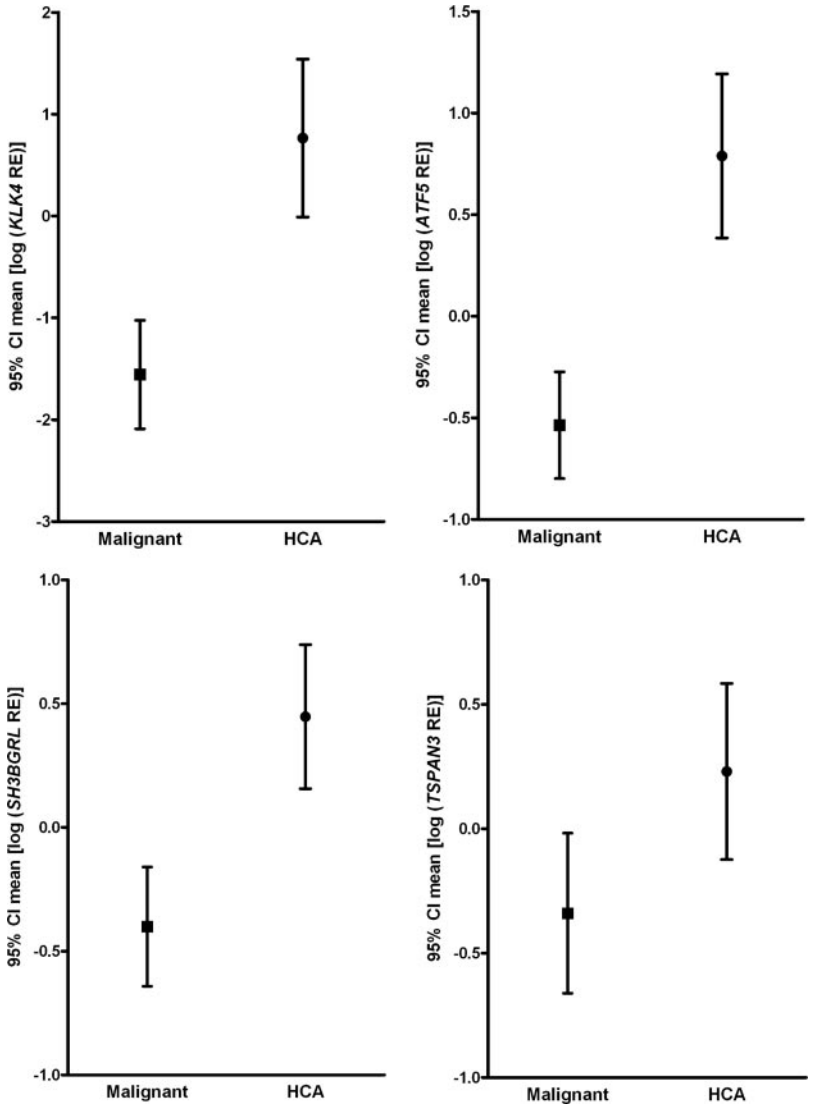

FIG. 3. Graphs showing results of qPCR for *ATF5*, *KLK4*, *SH3BGRL*, and *TSPAN3* transcripts evaluated in the HCA *vs.* malignant groups. The results are expressed as mean of log transformed RE values [95% confidence interval (CI)]. All transcripts are expressed in HCA while their expression is reduced (or absent) in most of malignant tumors. Specificity and sensitivity are detailed in Table 2.

(0.79–0.97), are shown in Fig. 2. *ATF5*, *KLK4*, *SH3BGRL*, and *TSPAN3* had lower sensitivity (0.68–0.74) (Fig. 3).

PVALB had best predicted sensitivity and specificity in paraffin-embedded thyroid tumors

On the basis of the results of ROC curve, we tested the best predictor's diagnostic markers (PVALB and KLKL1) by IHC analysis in a panel of paraffin-embedded thyroid sections selected from patients completely different from set 1.

Our original set of samples included 82 samples (27 HCA, 16 FTA, 14 HCC, 15 FTC, and 10 FVPTC). The pathological diagnosis was performed blinded. When the code was broken we found that 75 samples were in agreement with the pathological diagnosis when PVALB was used. Five samples that were originally classified as HCA by standard histological criteria were classified as malignant by our markers (*i.e.* PVALB was negative). Two samples that were diagnosed as HCC by standard histology were positive for PVALB. The performance of KLK1 was inferior given that 60 samples were diagnosed in accordance with pathological diagnosis. Six samples originally classified as HCA were negative for KLK1 while 16 malignant tumors stained positive.

All samples that were misclassified by our model were rereviewed by the two pathologists. Additional blocks were selected, and a deeper analysis was performed. Invasion through the tumor capsule and/or vascular invasion was found in four HCA, and therefore the four tumors were reclassified as HCC. The diagnosis was changed based on histological criteria. Interestingly, the four samples that were reclassified as HCC were negative for both PVALB and KLK1.

Our set of samples, therefore, included 23 HCA, 16 FTA, 18 HCC, 15 FTC, and 10 FVPTC. Overall, 22 of 23 HCA scored positive while 41 of 43 malignant tumors were negative (2 HCC) for PVALB (Fig. 4). KLK1 was positive in 21 of 23 HCA and negative in 27 of 43 malignant tumors (Fig. 5). PVALB and KLK1 were not expressed in matched normal thyroid tissues and FTAs. This analysis demonstrated not only that PVALB has a high accuracy and therefore can be used as preoperative diagnosis marker but also that it could be used to help postoperative diagnosis.

The combination of markers that best achieves high sensitivity and specificity is PVALB and C1orf24

To develop a more accurate and realistic antibody-based test, the previous identified carcinoma markers were also tested in combination in the aforementioned set of 82 samples. The carcinoma markers (ARG2, ITM1, and C1orf24) were positive in all malignant tumors while DDIT3 was negative in a FTC. Because all FTA subjects have negative C1orf24 and all carcinomas were C1orf24-positive, the use of a C1orf24 negative marker in the prediction of FTA gives 100% sensitivity and 100% specificity. ITM1 was the second best prediction marker of FTA, because it was positive in two FTAs with solid/trabecular pattern (Supplemental Table 2). Overall, this new validation analysis not only confirmed our previous findings but also demonstrated that C1orf24 is the best predictor of carcinoma. As expected from our initial analysis, all carcinoma markers presented false-positive results in all HCAs, substantiating the need to find HCA markers.

Our final panel detects malignancy when PVALB is negative and C1orf24 is positive. It detects HCA when PVALB is positive and C1orf24 is positive. It detects FTA when both markers are negative. Overall, this test gives 95.35% sensitivity, 97.44% specificity, 97.6 PPV, and 95% NPV, based on the 82 samples.

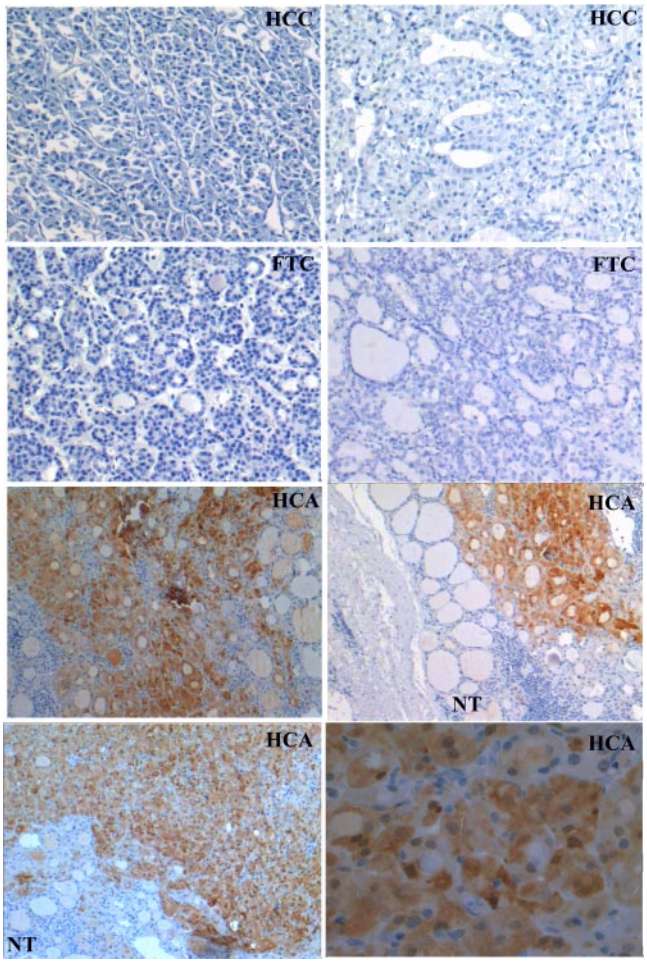

FIG. 4. Immunohistochemical analysis of PVALB in paraffin-embedded sections of thyroid samples. HCAs exhibited strong brown immunostaining for PVALB. In contrast, FTCs, HCCs, and negative control exhibited no immunoreactivity. Normal thyroid (NT) tissue is negative for PVALB, adjacent to tumor area that was positive for PVALB. In this study, PVALB was best in detecting HCA in 22 of 23 samples. Only 2 of 43 minimally invasive malignancies stained positive.

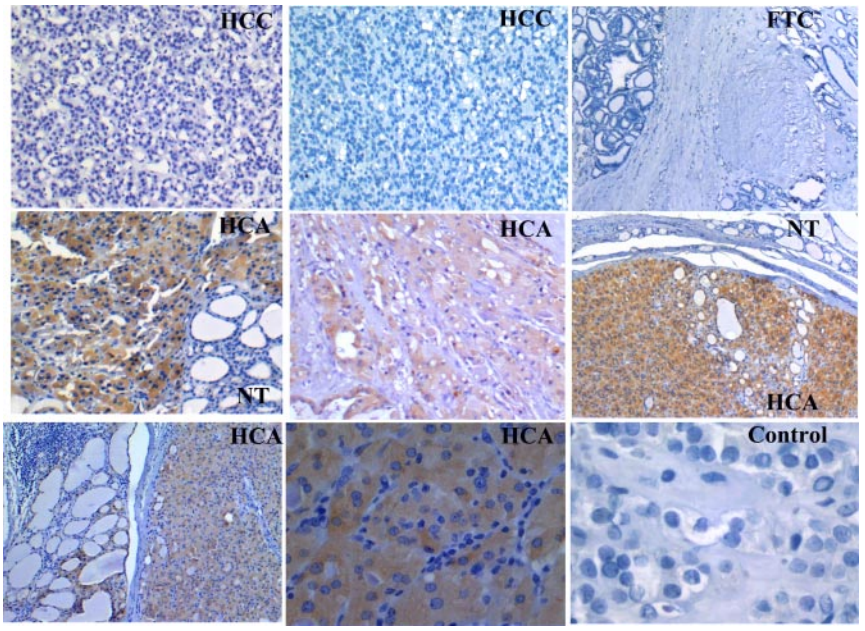

FIG. 5. Immunohistochemical analysis of KLK1 in paraffin-embedded sections of thyroid samples. HCAs exhibited strong brown immunostaining for KLK1. In contrast, FTC, most of HCC, and negative control exhibited no immunoreactivity. Adjacent normal thyroid (NT) tissue is negative for KLK1 at the same time as the tumor area is positive for PVALB.

Discussion

The major aim of this study was to identify new HCA biomarkers. This is important for two reasons. First, it helps improve a developing diagnostic test (4, 5) and allows us to more accurately identify HCA as a benign lesion. A more accurate test would make it easier to determine which patients should get total thyroidectomy, partial thyroidectomy, or followed without surgery for a benign tumor. Second, it helps to understand the biology of Hurtle Cell tumors which are seldom studied.

Although Hürthle tumors are rare, their prevalence is higher than the percentage of non–Hürthle cell follicular tumors displaying signs of malignancy (14). Moreover, the existence of many lesions of thyroid composed of Hürthle cells raises the problem of the differential diagnosis on FNA (15, 16), thereby arguing for the necessity of finding new HCA diagnostic markers.

Because just morphological features alone are not sufficient for improvements in the accuracy of FNA, molecular biomarkers are the logical choice to help distinguish tumors subtypes and improve presurgical diagnosis.

In this study we performed a gene expression screen using SAGE, which revealed more than 100 transcripts that were highly expressed in HCA compared with HCC. The abundance of these transcripts was assessed in the previously described FTA and FTC libraries (4). A total of 18 potential biomarkers were validated in a set of thyroid tumors by qPCR. Using this strategy we report sensitivity, specificity, and area under ROC curve for the classification of HCA and malignant thyroid tumors. Based on this analysis, the two best predicted HCA markers and our four previous identified carcinoma markers (4, 5) were than validated by IHC analysis in an independent set of 82 paraffin-embedded thyroid tumors. Importantly, by testing the carcinoma markers in this additional set of samples would not only allow evaluate the performance of these markers in Hürthle tumors but also define the best combination of markers for detecting carcinoma.

The data presented here allowed us to complete a panel of markers to classify all thyroid tumors commonly diagnosed as indeterminate on FNA with high sensitivity, specificity, and reproducibility. This assay is based on combination of a HCA marker and carcinoma marker, PVALB, and C1orf24, respectively (Fig 6). Using this combination, we report an overall 95% sensitivity, 97% specificity, 97% positive predicted value, and 95% negative predicted value, based on the 82 samples.

Importantly, the sensitivity and specificity of our test is 100% when we considered the most common thyroid tumor subtypes that can be source of diagnostic error on FNA such as FTC, PTC, and FVPTC (4, 5 and present study). According to our analysis, all possible malignancies would undergo surgery. Our test may also eliminate unnecessary surgery for FTA and hyperplasia, because the sensitivity and specificity approaches 100% when FTA and hyperplasia without Hürthle components are evaluated (5). The only two samples regarded as cancers but

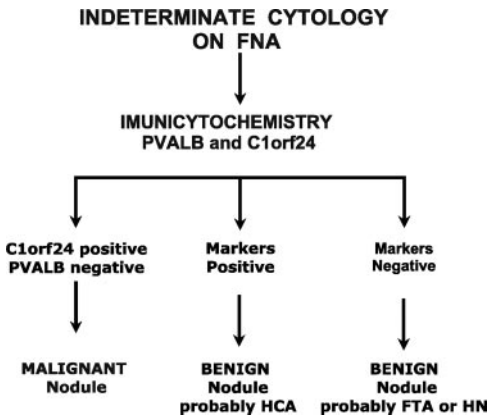

FIG. 6. The predicted best combination based on two antibody-based tests (PVALB and C1orf24) in paraffin-embedded thyroid samples and their application in nodules classified as indeterminate on FNA cytology. HN, Hyperplastic nodule.

called benign by our testing were two minimally invasive HCC. If this minimally invasive HCC are less aggressive tumors, only the follow-up will determine. Of note, 21 HCC were correctly classified by our test.

To our knowledge, no other test so far has achieved such a high sensitivity and specificity when included this wide range of thyroid tumors. Most other studies that achieved similar accuracy for thyroid tumor classification excluded HCA, because HCA was a source of false-positive results (16, 17).

From the clinical standpoint, these markers can also be used further on sections of the removed tumors to detect cancers that are missed in some cases by the present histology-only based diagnosis. In our investigation, four thyroid tumors that were previously classified as HCA on final histology and that were positive for carcinomas markers and negative for PVALB when reevaluated by two pathologist demonstrated invasion thorough the tumor capsule and/or vascular invasion. Based on the standard histological criteria, the pathologist reclassified these cases as HCC. Therefore, these markers have the potential to improve diagnostic accuracy over present standard of care.

Our results also provide an opportunity to better understand the biology of Hürthle tumors. The results of clustering analysis suggests that there is a different genetic signature when comparing Hürthle to follicular tumors, as previously described in the literature (18–23). Although PVALB was chosen as best predictor markers and we focus on markers of clinical utility, all transcripts validated here were differentially expressed between HCA and malignant thyroid tumors. Although our findings open an entire new field and new questions need to be answered, they may help understand clinical and biological behavior of this little investigated tumor class. The list of genes reported here, unlike other follicular derived carcinomas, Hürthle carcinomas are more refractory to treatment with radioactive iodine, have a greater incidence of nodal metastasis at presentation, higher recurrence rate, and the overall survival is reduced (3, 24).

The mechanism by which *PVALB* expression, a new HCA marker reported here, is lost in thyroid carcinomas is unclear, although *PVALB* is located at a frequently lost genomic locus. Chromosome 22 was the most common loss identified in HCC and for widely invasive thyroid carcinoma. Additionally, chromosome 22 losses were identified in HCC from patients who died of disease, and it may be of prognostic value in both FTC and HCC (25). The function of PVALB is not definitively known, but it appears to be a calcium ion binding protein that is expressed in several tissues such is brain and kidney. Interestingly, *PVALB* was described as down-regulated in clear cell and papillary renal carcinomas while it was expressed in chomophobe type, suggesting its potential to classify subtypes of renal carcinomas (26).

In summary, we identified new HCA markers that, in combination with our previous carcinomas markers, can distinguish a wide range of thyroid lesions with high sensitivity and specificity. Additionally, this immunopanel can be easily performed on FNA material, which is an inexpensive nonsurgical diagnostic tool already in clinical use for the initial evaluation of a thyroid nodule. This approach to thyroid nodules could have an important clinical impact on the management of a thyroid nodule and/or for the postoperative evaluation of thyroid biopsy tissue sections.

Acknowledgments

Address all correspondence and requests for reprints to: Janete Cerutti, Ph.D., Genetic Bases of Thyroid Tumor Laboratory, Rua Pedro de Toledo 669, 11° andar, Federal University of São Paulo, 04039-032, São Paulo, SP, Brazil. E-mail: j.cerutti@unifesp.br.

This work was supported by National Institutes of Health Grant CA113461 and São Paulo State Research Foundation (FAPESP) Grant 05/60330-8. J.M.C. and S.J.S. are investigators of the Brazilian Research Council (CNPq), G.O. is scholar from FAPESP, and G.J.R. is the Irving J. Sherman, M.D. Research Professor.

Disclosure Summary: The authors have nothing to declare.

References

1. **Mazzaferri EL** 1993 Management of a solitary thyroid nodule. N Engl J Med 328:553–559
2. **Ezzat S, Sarti DA, Cain DR, Braunstein GD** 1994 Thyroid incidentalomas. Prevalence by palpation and ultrasonography. Arch Intern Med 154:1838–1840
3. **Baloch ZW, LiVolsi VA, Asa SL, Rosai J, Merino MJ, Randolph G, Vielh P, DeMay RM, Sidawy MK, Frable WJ** 2008 Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. Diagn Cytopathol 36:425–437
4. **Cerutti JM, Delcelo R, Amadei MJ, Nakabashi C, Maciel RM, Peterson B, Shoemaker J, Riggins GJ** 2004 A preoperative diagnostic test that distinguishes benign from malignant thyroid carcinoma based on gene expression. J Clin Invest 113:1234–1242
5. **Cerutti JM, Latini FR, Nakabashi C, Delcelo R, Andrade VP, Amadei MJ, Maciel RM, Hojaij FC, Hollis D, Shoemaker J, Riggins GJ** 2006 Diagnosis of suspicious thyroid nodules using four protein biomarkers. Clin Cancer Res 12:3311–3318
6. **Matsumoto F, Fujii H, Abe M, Kajino K, Kobayashi T, Matsumoto T, Ikeda K, Hino O** 2006 A novel tumor marker, Niban, is expressed in subsets of thyroid tumors and Hashimoto's thyroiditis. Hum Pathol 37:1592–1600
7. **Saha S, Sparks AB, Rago C, Akmaev V, Wang CJ, Vogelstein B, Kinzler KW, Velculescu VE** 2002 Using the transcriptome to annotate the genome. Nat Biotechnol 20:508–512

8. **Lal A, Lash AE, Altschul SF, Velculescu V, Zhang L, McLendon RE, Marra MA, Prange C, Morin PJ, Polyak K, Papadopoulos N, Vogelstein B, Kinzler KW, Strausberg RL, Riggins GJ** 1999 A public database for gene expression in human cancers. Cancer Res 59: 5403–5407
9. **Galante PA, Trimarchi J, Cepko CL, de Souza SJ, Ohno-Machado L, Kuo WP** 2007 Automatic correspondence of tags and genes (ACTG): a tool for the analysis of SAGE, MPSS and SBS data. Bioinformatics 23:903–905
10. **Zhang L, Zhou W, Velculescu VE, Kern SE, Hruban RH, Hamilton SR, Vogelstein B, Kinzler KW** 1997 Gene expression profiles in normal and cancer cells. Science 276:1268–1272
11. **Suzuki R, Shimodaira H** 2006 Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 22: 1540–1542
12. **Cerutti JM, Oler G, Michaluart Jr P, Delcelo R, Beaty RM, Shoemaker J, Riggins GJ** 2007 Molecular profiling of matched samples identifies biomarkers of papillary thyroid carcinoma lymph node metastasis. Cancer Res 67:7885–7892
13. **Oler G, Camacho CP, Hojaij FC, Michaluart Jr P, Riggins GJ, Cerutti JM** 2008 Gene expression profiling of papillary thyroid carcinoma identifies transcripts correlated with BRAF mutational status and lymph node metastasis. Clin Cancer Res 14:4735–4742
14. **Máximo V, Sobrinho-Simões M** 2000 Hurthle cell tumours of the thyroid. A review with emphasis on mitochondrial abnormalities with clinical relevance. Virchows Arch 437:107–115
15. **Belchetz G, Cheung CC, Freeman J, Rosen IB, Witterick IJ, Asa SL** 2002 Hurthle cell tumors: using molecular techniques to define a novel classification system. Arch Otolaryngol Head Neck Surg 128: 237–240
16. **Montone KT, Baloch ZW, LiVolsi VA** 2008 The thyroid Hurthle (oncocytic) cell and its associated pathologic conditions: a surgical pathology and cytopathology review. Arch Pathol Lab Med 132: 1241–1250
17. **Barroeta JE, Baloch ZW, Lal P, Pasha TL, Zhang PJ, LiVolsi VA** 2006 Diagnostic value of differential expression of CK19, Galectin-3, HBME-1, ERK, RET, and p16 in benign and malignant follicular-derived lesions of the thyroid: an immunohistochemical tissue microarray analysis. Endocr Pathol 17:225–234
18. **Sahin M, Allard BL, Yates M, Powell JG, Wang XL, Hay ID, Zhao Y, Goellner JR, Sebo TJ, Grebe SK, Eberhardt NL, McIver B** 2005 PPARγ staining as a surrogate for PAX8/PPARγ fusion oncogene expression in follicular neoplasms: clinicopathological correlation and histopathological diagnostic value. J Clin Endocrinol Metab 90:463–468
19. **Máximo V, Botelho T, Capela J, Soares P, Lima J, Taveira A, Amaro T, Barbosa AP, Preto A, Harach HR, Williams D, Sobrinho-Simões M** 2005 Somatic and germline mutation in GRIM-19, a dual function gene involved in mitochondrial metabolism and cell death, is linked to mitochondrion-rich (Hurthle cell) tumours of the thyroid. Br J Cancer 92:1892–1898
20. **Masood S, Auguste LJ, Westerband A, Belluco C, Valderama E, Attie J** 1993 Differential oncogenic expression in thyroid follicular and Hurthle cell carcinomas. Am J Surg 166:366–368
21. **Farrand K, Delahunt B, Wang XL, McIver B, Hay ID, Goellner JR, Eberhardt NL, Grebe SK** 2002 High resolution loss of heterozygosity mapping of 17p13 in thyroid cancer: Hurthle cell carcinomas exhibit a small 411-kilobase common region of allelic imbalance, probably containing a novel tumor suppressor gene. J Clin Endocrinol Metab 87:4715–4721
22. **Cheung CC, Ezzat S, Ramyar L, Freeman JL, Asa SL** 2000 Molecular basis off hurthle cell papillary thyroid carcinoma. J Clin Endocrinol Metab 85:878–882
23. **Chiappetta G, Toti P, Cetta F, Giuliano A, Pentimalli F, Amendola I, Lazzi S, Monaco M, Mazzuchelli L, Tosi P, Santoro M, Fusco A** 2002 The RET/PTC oncogene is frequently activated in oncocytic thyroid tumors (Hurthle cell adenomas and carcinomas), but not in oncocytic hyperplastic lesions. J Clin Endocrinol Metab 87:364–369
24. **Mills SC, Haq M, Smellie WJ, Harmer C** 2009 Hurthle cell carcinoma of the thyroid: Retrospective review of 62 patients treated at the Royal Marsden Hospital between 1946 and 2003. Eur J Surg Oncol 35:230–234
25. **Erickson LA, Jalal SM, Goellner JR, Law ME, Harwood A, Jin L, Roche PC, Lloyd RV** 2001 Analysis of Hurthle cell neoplasms of the thyroid by interphase fluorescence in situ hybridization. Am J Surg Pathol 25:911–917
26. **Chen YT, Tu JJ, Kao J, Zhou XK, Mazumdar M** 2005 Messenger RNA expression ratios among four genes predict subtypes of renal cell carcinoma and distinguish oncocytoma from carcinoma. Clin Cancer Res 11:6558–6566