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 non palpable 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...
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 long SAGE 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

<table>
<thead>
<tr>
<th>Tag sequence</th>
<th>HCA Count</th>
<th>HCC Count</th>
<th>FTC Count</th>
<th>Transcript description</th>
<th>GenBank accession no.</th>
<th>Locus</th>
<th>Gene ontology</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTGGACAAGGACGACCAAG</td>
<td>89</td>
<td>0</td>
<td>0</td>
<td>PARVALB, Parvalbumin</td>
<td>NM_002854</td>
<td>22q12-q13.1</td>
<td>Calcium ion binding</td>
</tr>
<tr>
<td>GGGCTAGCTGCTGGCTTTG</td>
<td>51</td>
<td>0</td>
<td>0</td>
<td>KLK1, Kallikrein 1</td>
<td>BC005313</td>
<td>19q13.3</td>
<td>Tissue kallikrein activity</td>
</tr>
<tr>
<td>CTGACCGCGGAGTACCC</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>OBSCN, Obscurin, cytoskeletal calmodulin</td>
<td>NM_052843</td>
<td>1q24.13</td>
<td>Protein-tyrosine kinase activity</td>
</tr>
<tr>
<td>TCTGACATTTGGACGCC</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>ALDH1L1, Aldehyde dehydrogenase 1 family member L1</td>
<td>BC027241</td>
<td>3q21.2</td>
<td>Methyltransferase activity</td>
</tr>
<tr>
<td>TCTCGGCGCCGGCGAGGG</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>KLK4, Kallikrein 4 (prostate enamel matrix prostate)</td>
<td>NM_004917</td>
<td>19q13.41</td>
<td>Peptidase activity</td>
</tr>
<tr>
<td>TGGGTATCTCTACTACT</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>GPPT1, Glutamin-fructose-6-phosphate transaminase 1</td>
<td>NM_002056</td>
<td>2p13</td>
<td>Transferase activity</td>
</tr>
<tr>
<td>TACCTCTCAAGCCAG</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>ATP1A3, ATPase Na+/K+/H3 polypeptide</td>
<td>NM_152296</td>
<td>19q13.31</td>
<td>ATP binding</td>
</tr>
<tr>
<td>AACTCTTTTCAGTGT</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>NUDT22, Nudix (nucleoside diphosphate linked moiety X-type) motif 22</td>
<td>NM_005528</td>
<td>11q13</td>
<td>Heat shock protein binding</td>
</tr>
<tr>
<td>CCCTGGAATCTCCTCTCTT</td>
<td>15</td>
<td>0</td>
<td>1</td>
<td>TSPAN33, Tetraspanin 33</td>
<td>NM_005724</td>
<td>15q24.3</td>
<td>Development, activation, growth &amp; motility</td>
</tr>
<tr>
<td>CTAAGCTTCTCCAGAGA</td>
<td>15</td>
<td>0</td>
<td>1</td>
<td>SMARC02, SWNSF related matrix associated actin dependent regulator of chromatin subfamily d member 3</td>
<td>NM_003078</td>
<td>7q35-q36</td>
<td>Transcription regulator activity</td>
</tr>
<tr>
<td>GGCGCCAGCGGGCTCAA</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>TUSC2, Tumor suppressor candidate 2</td>
<td>NM_007275</td>
<td>3p21.3</td>
<td>Negative regulation of cell cycle</td>
</tr>
<tr>
<td>TCCAGTGGTTTTTTTG</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>TBC1D1, (tre-2AUSP6 BUB2 cd16) domain family member 1</td>
<td>NM_015173</td>
<td>4p14</td>
<td>GTPase activator activity</td>
</tr>
<tr>
<td>GGATTGAGATAGG</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>DVL3, Dishevelled dsh homolog 3 (Drosophila)</td>
<td>NM_004423</td>
<td>3q27</td>
<td>Signal transducer activity</td>
</tr>
<tr>
<td>CTAACCTCCTTGTCGCGT</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>ENAH, Enabled homolog (Drosophila)</td>
<td>NM_018212</td>
<td>1q24.12</td>
<td>Actin binding</td>
</tr>
<tr>
<td>TGCAAACCTTICTGTTT</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>LOC289833, Similar to hypothetical protein MGC27019</td>
<td>NM_001033515</td>
<td>Unknown</td>
<td>Microtubule cytoskeleton organization and biogenesis</td>
</tr>
<tr>
<td>TCACTCCTGACGACCCT</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>ATF5, Activating transcription factor 5</td>
<td>BC005174</td>
<td>19q13.3</td>
<td>Transcription factor activity</td>
</tr>
<tr>
<td>TTCCTGTACAAATT</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>CHUK, Conserved helix-loop-helix ubiquitins kinase</td>
<td>NM_001278</td>
<td>10q24-q25</td>
<td>Protein Serine/threonine kinase activity</td>
</tr>
<tr>
<td>TATATTCTTATGTGCT</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>SH3BGLR, SH3 domain binding glutamic acid-rich protein like</td>
<td>NM_003022</td>
<td>Xq13.3</td>
<td>SH3 domain binding</td>
</tr>
</tbody>
</table>

*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 NlaIII 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 samples 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 FTC. 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 (R8) (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

<table>
<thead>
<tr>
<th>Gene</th>
<th>Area under the ROC curve</th>
<th>Cut-offa</th>
<th>Number of true carcinoma</th>
<th>Sensitivity</th>
<th>CI sensitivity</th>
<th>Specificity</th>
<th>CI specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ALDH1L1</td>
<td>0.941</td>
<td>0.6199</td>
<td>30</td>
<td>0.88</td>
<td>(0.69-0.96)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>2. TUSC2</td>
<td>0.787</td>
<td>0.7099</td>
<td>23</td>
<td>0.68</td>
<td>(0.43-0.83)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>3. ATFS</td>
<td>0.903</td>
<td>0.6499</td>
<td>23</td>
<td>0.68</td>
<td>(0.43-0.83)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>4. OBSCN</td>
<td>0.781</td>
<td>0.8799</td>
<td>25</td>
<td>0.74</td>
<td>(0.54-0.90)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>5. KLK4</td>
<td>0.889</td>
<td>0.1999</td>
<td>23</td>
<td>0.68</td>
<td>(0.43-0.83)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>6. KLK1</td>
<td>0.971</td>
<td>0.9999</td>
<td>33</td>
<td>0.97</td>
<td>(0.84-1.00)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>7. CHUK</td>
<td>0.753</td>
<td>0.3699</td>
<td>21</td>
<td>0.62</td>
<td>(0.38-0.82)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>8. GFTP1</td>
<td>0.787</td>
<td>0.6099</td>
<td>20</td>
<td>0.59</td>
<td>(0.36-0.81)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>9. DVL3</td>
<td>0.833</td>
<td>0.6899</td>
<td>18</td>
<td>0.53</td>
<td>(0.30-0.78)</td>
<td>11</td>
<td>0.85</td>
</tr>
<tr>
<td>10. PVCLB</td>
<td>0.894</td>
<td>1.3899</td>
<td>32</td>
<td>0.94</td>
<td>(0.80-0.99)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>11. SHBGR1</td>
<td>0.842</td>
<td>0.6999</td>
<td>24</td>
<td>0.71</td>
<td>(0.49-0.87)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>12. SMARC3</td>
<td>0.810</td>
<td>0.6099</td>
<td>19</td>
<td>0.56</td>
<td>(0.33-0.79)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>13. TBC1D1</td>
<td>0.873</td>
<td>0.3199</td>
<td>18</td>
<td>0.53</td>
<td>(0.30-0.78)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>14. ATPIA3</td>
<td>0.790</td>
<td>0.0319</td>
<td>8</td>
<td>0.24</td>
<td>(0.03-0.65)</td>
<td>11</td>
<td>0.85</td>
</tr>
<tr>
<td>15. NUDT22</td>
<td>0.882</td>
<td>1.0199</td>
<td>27</td>
<td>0.79</td>
<td>(0.58-0.92)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>16. TSPAN3</td>
<td>0.869</td>
<td>1.5199</td>
<td>25</td>
<td>0.74</td>
<td>(0.54-0.90)</td>
<td>12</td>
<td>0.92</td>
</tr>
<tr>
<td>17. ENAH</td>
<td>0.735</td>
<td>0.3299</td>
<td>13</td>
<td>0.38</td>
<td>(0.14-0.68)</td>
<td>12</td>
<td>0.92</td>
</tr>
</tbody>
</table>

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.

**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

![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, PV1/2 and KLK1 distinguish HCA and malignant class with higher specificity and sensitivity, 0.94 and 0.97, respectively (as detailed in Table 2).](image-url)
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 \( (P < 0.01, \text{Mann-Whitney-Wilcoxon test}) \). The greater difference in expression was observed for \textit{ALDH1L1}, \textit{ATF5}, \textit{KLK4}, \textit{KLK1}, \textit{DVL3}, \textit{PVALB}, \textit{SH3BGRL}, \textit{TBC1D1}, \textit{NUDT22}, and \textit{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. \textit{KLK1} (sensitivity = 0.97; area under ROC curve = 0.971) and \textit{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 \textit{ALDH1L1}, \textit{KLK1}, \textit{PVALB}, and \textit{NUDT22}, which had higher sensitivity

FIG. 3. Graphs showing results of qPCR for \textit{ATF5}, \textit{KLK4}, \textit{SH3BGRL}, and \textit{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. $\text{ATF5}$, $\text{KLK4}$, $\text{SH3BGR1}$, and $\text{TSPAN3}$ 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.
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–Hu¨rthle cell follicular tumors displaying signs of malignancy (14). Moreover, the existence of many lesions of thyroid composed of Hu¨rthle 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 then 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 carcinomas.

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
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).

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, 11º 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


