Diagnosis of Suspicious Thyroid Nodules Using Four Protein Biomarkers


Authors' Affiliations: 1Department of Neurosurgery, Johns Hopkins University Medical School, Baltimore, Maryland; Divisions of 2Genetics, 3Endocrinology, Head and Neck, and 4Department of Pathology, Federal University of São Paulo, São Paulo, Brazil; and 5Department of Pathology, Duke University Medical Center, Duke University, Durham, North Carolina

Received 10/11/05; revised 3/30/06; accepted 4/6/06.

Grant support: NIH/Cancer Genome Anatomy Project contract S98-146, NIH/National Cancer Institute grant R21 CA113461, and the Ludwig Trust and the São Paulo State Research Foundation grants 03/10626-2 and 04/15288-0.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: J.M. Cerutti and R.M.B. Maciel are investigators of the Brazilian Research Council. F.R.M. Latin is a scholar from São Paulo State Research Foundation. G.J. Riggins is the recipient of the Irving J. Sherman M.D. Research Professorship. Conflict of interest: The authors declare that they have no affiliations that would constitute a financial conflict of interest relating to the subject matter of this study.

Requests for reprints: Gregory J. Riggins or Janete Cerutti, Johns Hopkins University, CRB II, Room 257, 1550 Orleans Street, Baltimore, MD 21231. Phone: 410-502-2304; Fax: 410-502-5559; E-mail: griggin1@jhmi.edu or cerutt-end@pesquisa.epm.br.

doi:10.1158/1078-0432.CCR-05-2226

Imaging, Diagnosis, Prognosis

Abstract

Purpose: Fine-needle aspiration (FNA) cytology, a standard method for thyroid nodule diagnosis, cannot distinguish between benign follicular thyroid adenoma (FTA) and malignant follicular thyroid carcinoma (FTC). Previously, using expression profiling, we found that a combination of transcript expression levels from DDIT3, ARG2, C1orf24, and ITM1 distinguished between FTA and FTC. The goal of this study was to determine if antibody markers used alone or in combination could accurately distinguish between a wider variety of benign and malignant thyroid lesions in fixed sections and FNA samples.

Experimental Design: Immunohistochemistry was done on 27 FTA, 25 FTC, and 75 other benign and malignant thyroid tissue sections using custom antibodies for chromosome 1 open reading frame 24 (C1orf24) and integral membrane protein 1 (ITM1) and commercial antibodies for DNA damage–inducible transcript 3 (DDIT3) and arginase II (ARG2). FNA samples were also tested using the same antibodies. RNA expression was measured by quantitative PCR in 33 thyroid lesions.

Results: C1orf24 and ITM1 antibodies had an estimated sensitivity of 1.00 for distinguishing FTA from FTC. For the expanded analysis of all lesions studied, ITM1 had an estimated sensitivity of 1.00 for detecting malignancy. Because all four cancer biomarkers did well, producing overlapping confidence intervals, not one best marker was distinguished. Transcript levels also reliably predicted malignancy, but immunohistochemistry had a higher sensitivity. Malignant cells were easily detected in FNA samples using these markers.

Conclusions: We improved this diagnostic test by adding C1orf24 and ITM1 custom antibodies and showing use on a wider variety of thyroid pathology. We recommend that testing of all four cancer biomarkers now be advanced to larger trials. Use of one or more of these antibodies should improve diagnostic accuracy of suspicious thyroid nodules from both tissue sections and FNA samples.

Thyroid nodules are commonly encountered during routine medical care. An estimated 4% to 7% of the adult population develops clinically palpable thyroid nodules during their lifetime. With the advent of ultrasound in medical practice, an increasing number of impalpable thyroid nodules can be detected in 20-67% of the general population (1–3).

Fine-needle aspiration (FNA) is considered the best initial diagnostic tool to evaluate thyroid nodules. Although papillary thyroid carcinomas (PTC) are easily diagnosed by FNA, benign follicular thyroid adenoma (FTA) and the malignant follicular thyroid carcinoma (FTC) cannot be distinguished, yielding an inconclusive diagnosis of ‘suspicious’ or ‘follicular-patterned lesion’ in 10% to 30% of all FNA biopsies (4, 5). Using FNA alone, the cytologic features of benign and malignant follicular tumors overlap. Currently for follicular-patterned lesions, the tumor requires histologic evaluation for capsular and/or vascular invasion, which indicates FTC.

In addition to FTA and FTC, the differential diagnosis of suspicious thyroid nodules from FNA cytology may represent hyperplasia or a follicular variant PTC (FVPTC). FVPTC diagnosis is difficult due to its overlapping features with both benign and malignant follicular thyroid lesions. In addition to FNA being inconclusive for these lesions, diagnostic discrepancies also occur in the final histology for minimally invasive FTC, PTC, and hyperplastic nodules (6–8).

Guidelines for evaluation and management of patients with thyroid nodules suggest that all patients with a cytologic report of follicular-patterned lesions should have a surgical biopsy (9, 10). Only 5% to 20% of these cytologically suspicious...
nODULES, when removed, are malignant on histology, indicating many surgeries could be avoided if there was a more accurate diagnostic test. Additionally, those patients with malignancy may have delayed or miss treatment if a malignancy is missed by final histology alone or because a diagnosis of “suspicious” alone does not prompt a rapid follow-up (11–13). A more acute molecular-based test for thyroid nodules is needed not only to improve treatment decisions but also to potentially reduce the long-term health costs.

Several groups have evaluated the usefulness of biomarkers in the diagnosis of suspicious thyroid lesions. However, when applied in clinical practice, the biomarkers failed to show use as an adjunctive test (14–16).

To address the problem of incomplete diagnosis of thyroid nodules, we previously did gene expression profiling of a FTA, FTC, and normal thyroid tissues (17). This profiling and subsequent validation using quantitative real-time PCR revealed that four genes (DDIT3, ARG2, C1orf24, and ITM1) differed between the two classes, and a linear combination of expression levels distinguished FTC from FTA with an estimated predictive accuracy of 0.83. Two commercial antibodies [DNA damage inducible transcript 3 (DDIT3) and arginase II (ARG2)] were used for further validation in independent FTA and FTC paraffin-embedded samples.

Here, we add new antibodies for chromosome 1 open reading frame 24 (C1orf24) and integral membrane protein 1 (ITM1) to evaluate all four potential markers on a wider range of thyroid pathology. We custom produced polyclonal antibodies for C1orf24 and ITM1 and validated our data by immunohistochemistry in FTA and FTC sections. In addition, we tested the expression levels of all four genes using both immunohistochemistry and quantitative PCR in the types of suspicious thyroid lesions that can produce diagnostic errors. Moreover, we assessed the effectiveness of immunostaining of the four genes for the preoperative diagnosis of suspicious lesions and compared them with final histology on paraffin-embedded sections. The antibodies for C1orf24 and ITM1 were effective in distinguishing FTA from FTC and also distinguished between benign and malignant for the other thyroid pathologies tested. This study shows that all four antibodies are reliable for the detection of cancer antigen expression in FNA material. Our results support the testing of these markers in a larger trial to determine if FNA thyroid nodule diagnosis can be improved as predicted from this study.

Materials and Methods

Materials and Methods

Tumor samples. The first validation set was 127 paraffin-embedded sections from the Department of Pathology, Federal University of São Paulo, Brazil collected from 1997 to 2002 (27 FTA, 25 FTC, 27 FVPTC, 15 FTC, and 33 hyperplastic nodules). A second validation set was 33 samples (11 FVPTC and 22 benign hyperplastic nodules) obtained from patients undergoing thyroid surgery for thyroid disease at Hospital São Paulo, Federal University of São Paulo, Brazil. These samples were frozen immediately after surgical biopsy and stored at −80°C. Final histologic classification was obtained from paraffin-embedded tissue. A third independent set of samples included four FNA specimens and two intraoperative aspirates from nodules previously diagnosed as suspicious; five of six patients underwent surgical treatment, which was determined through evaluating the clinical profile of the biopsy and through cytologic diagnosis. Written informed consent was obtained from each patient. The study was approved by the Ethics and Research Committee of the Federal University of São Paulo.

Antibodies. Peptides corresponding to the COOH-terminal sequence of C1orf24 (910-928; DVKGEQQGQSPSPFPPSEEE) and ITM1 (692-705; YVKKDLNRGLSRT) were synthesized by solid-phase synthesis and purified by reversed-phase high-performance liquid chromatography (Zymed Laboratories, Inc., South San Francisco, CA). Polyclonal antibodies against C1orf24 (also named Niban) and ITM1 were generated by immunizing female New Zealand rabbits with COOH-terminal peptides (Zymed Laboratories). The antibodies against DDIT3 (R-20) and ARG2 (H-64) were purchased from Santa Cruz Biotechnology (Santa Cruz, CA).

Cell culture and transfection. The parental NIH3T3 and the papillary thyroid carcinoma (NPA) cell lines were maintained in DMEM supplemented with 10% FCS serum, 2 mmol/L l-glutamine, and 1% penicillin-streptomycin at 37°C and 5% CO2. NIH3T3 were transfected with 20 μg of plasmid DNA from either pCMV6-XL4 vector encoding the human C1orf24 gene or pCMV6-XL5 vector encoding the ITM1 gene (Origene Technologies, Inc., Rockville, MD) by electroporation. Following a 48-hour incubation, total protein was obtained from the parental NIH3T3 and the transfected cells as previously described (18).

Immunoblot analysis. To show the specificity of the human C1orf24 and ITM1 polyclonal antibodies, whole extracts were obtained from parental NIH3T3, transfected NIH3T3, NPA cell line, three normal thyroid tissues, and a FTC that was used to generate the initial SAGE library (18). Protein concentration was measured with a Bradford assay (Bio-Rad, Hercules, CA). The extracted proteins (40 μg/well) were separated on 10% SDS-polyacrylamide gels and transferred to Immobilon-P polyvinylidene difluoride membrane (Bio-Rad). After transfer, sheets were subjected to Ponceau S (Sigma-Aldrich, Inc., St. Louis, MO) staining, which was consistently identical in all lanes. Membranes were blocked for 1 hour and then incubated with primary antibodies. Anti-ITM1 was used at a dilution of 1:100, and anti-C1orf24 was used at a dilution of 1:200. The membranes were incubated for 1 hour with goat anti-rabbit horseradish peroxidase–conjugated IgG diluted 1:10,000 (Bio-Rad). As a protein loading control, these membranes were stripped and reprobed for anti-α-tubulin (Sigma-Aldrich, Inc.). Proteins were visualized using horseradish peroxidase–conjugated secondary antibodies. The immune complexes were detected using a Super Signal West Pico Chemiluminescent Substrate (Pierce Biotechnology, Inc., Rockford, IL).

Immunohistochemistry analysis. Sections were deparaffinized in xylene and rehydrated, and antigen retrieval was archived using buffer AR-10 (Biogenex, San Ramon, CA) in a steamer for 10 minutes. Sections were then exposed to the primary antibodies for at least 16 hours at 4°C. DDIT3, ARG2, and ITM1 were used at dilution of 1:100, and C1orf24 was used at a dilution of 1:200. The labeled DAKO Envision+ System/horseradish peroxidase and substrate 3,3′-diaminobenzidine were used for immunodetection (DAKO Laboratories, Carpinteria, CA). The slides were counterstained with 1% hematoxylin and permanently mounted. The control for antibody specificity included incubation with rat IgG used at the same concentration as the first antibody (DAKO Laboratories). All slides were blindly scored in a semiquantitative manner as described (17).

cDNA synthesis and quantitative PCR. To determine if the four predictor markers for FTC can correctly classify other benign or malignant thyroid lesions at the RNA level, the expression levels of four genes were analyzed by quantitative PCR. An aliquot of cDNA was used in 20-μl PCR reactions containing Taqman universal PCR master mix, 10 μmol/L of each specific primer and FAM-labeled probes for the target (DDIT3, ARG2, C1orf24, and ITM1) or reference genes (RS3 and QP-C; Taqman Gene Assays on Demand, Applied Biosystems, Foster City, CA). Quantitative PCR reactions were done in triplicate, and the
threshold cycle (Ct) obtained using Applied Biosystems software (Applied Biosystems) was averaged (SD ≤ 1). Gene expression was calculated as described (17, 19).

**FNA cytology.** To test DDT3, ARG2, ITM1, and C1orf24 antibodies for immunocytochemistry, preoperative FNA samples from suspicious nodules were immunostained (n = 4). Two aspirations were done on each patient with ultrasound guidance by an experienced cytopathologist. One aspirate was used to prepare smears, which were air-dried and stained for standard diagnosis. The second part was directly transferred to a vial containing PreservCyt solution (Cytene Corp., Kansas City, MO). Aliquots of 100 μL were placed in 12 Shandon CytoSpin chambers (Shandon, Inc., Pittsburgh, PA) and centrifuged at 1,200 rpm for 10 minutes on silane-coated slides. At least four slides from each case were prepared and fixed in ethanol until immunostaining; immunocytochemistry was done as previously mentioned with no antigen retrieval. Additionally, in two cases, intraoperative aspirate was obtained from nodules diagnosed as suspicious for FVPTC and medullary thyroid carcinoma and prepared as mentioned. The nodules were classified as malignant when >10% of cells were immunoreactive for one to four markers, or negative when immunoreactivity was absent for all markers.

**Statistical analysis.** For the immunohistochemistry data, the markers were analyzed separately and in combination. For each marker (DDT3, ARG2, C1orf24, and ITM1), point estimates and 95% confidence intervals (95% CI) were calculated for sensitivity and specificity. Confidence intervals were calculated using the score CI, which is more desirable than an asymptotic CI when the sample size is small, or the sensitivity is close to 1.0 (20). Sensitivity was defined as the proportion of carcinoma patients who expressed a marker (had a positive test result). Specificity was defined as the proportion of patients with benign lesions who did not express the marker (i.e., had a negative test result). Patients were defined as carcinoma patients or as patients with benign lesions based on histology. To determine if a combination of markers (DDT3, ARG2, ITM1, and C1orf24) could better classify the nodules into benign or malignant class using immunohistochemistry, we separately examined the data sets consisting of two classes (FTC and NPA) and three classes (FVPTC, PTC, and hyperplastic nodules). We used Classification and Regression Trees, used to test whether there is an interaction between markers (21).

For the quantitative PCR data, we first determine if the expression values for four genes were different between benign (n = 22) and malignant (n = 11) thyroid samples. Second, we determined whether a class predictor (benign versus malignant) could be developed using the quantitative PCR data. Because the four genes were selected as being differentially expressed between FTC and FTA, it was possible that they would not distinguish between the benign and malignant thyroid classes in the present study. A Student’s t test was used to compare expression levels of log-transformed data. A comparison was designated as statistically significant if the t statistic was found to be significant, using an α level that had been adjusted (using a Bonferroni adjustment) to keep the family-wise error rate at 0.05. A one-sided test was used because of the expectation that the gene expression in the malignant samples would be higher than the gene expression in the benign samples. For the development of the class predictor, we followed the framework outlined by Radmacher et al. (22) using the compound covariate predictor for gene expression data (22, 23) as described (17).

**Results**

**Specificity of anti-human ITM1 and C1orf24 polyclonal antibodies.** We showed previously that transcript levels for ITM1 and C1orf24 distinguished between FTA and FTC (17). No antibodies were previously available for these genes. Therefore, we generated rabbit polyclonal antibodies to ITM1 and C1orf24 polypeptides and assessed antibody specificity by Western blot. The C1orf24 antibody detected a predicted 130-kDa protein in NIH3T3 cells expressing C1orf24 and the NPA cell line, which we determined previously expresses the C1orf24 transcript (ref. 17; Fig. 1A). C1orf24 was not detected in normal thyroid tissues and the parental NIH3T3 cell line. A smaller band of 70 kDa was detected in NPA and FTC (Fig. 1A). This result is in agreement with a previous study using C1orf24 antibody that also found both bands in renal tumors from Eker rats and human renal carcinoma cell lines but only a 130-kDa band in NIH3T3 transiently transfected (24).

**ITM1** is a new gene that encodes a protein with a high degree of evolutionary conservation (98.5%) between mouse and human. The predicted size of ITM1 (80 kDa) was observed in NIH3T3 expressing ITM1, lower levels in parental NIH3T3, FTC, and NPA cell line but not detected in the normal thyroid tissues (Fig. 1B). A larger band was detected in NIH3T3 transfected with an ITM1 vector. A likely explanation would be that ITM1 might be insoluble and migrated as higher-order complexes (200-250 kDa) rather than monomers (~80 kDa). Similar results were obtained with STT3, which belongs to the ITM1 family. STT3 was found insoluble and resistant to extraction with relatively mild detergents showing an anomalous electrophoretic migration that persisted after deglycosylation (25).

**Immunostaining.** A total of 127 paraffin-embedded samples were analyzed for the expression of DDT3, ARG2, ITM1, and C1orf24. Five of the 32 FTA and 2 of the 27 FTC specimens that were previously analyzed for DDT3 and ARG2 (17) were not available for assessing ITM1 and C1orf24 expression; these seven samples were included in the estimates of sensitivity and specificity of DDT3 and ARG2 (thus, the total sample size for these estimates was 59 instead of 52) but not included in the Classification and Regression Trees analysis.

![Fig. 1. Expression of rabbit polyclonal antibodies for human ITM1 and C1orf24. Western blot analysis of extracted obtained from a NIH3T3 transfected with vector expressing ITM1 (A) or C1orf24 (B). Transfected NIH3T3 (lane 1), parental NIH3T3 (lane 2), FTC cell line (lane 3), FTC (lane 4), normal thyroid tissues (lanes 5-7).](Image)
We found that C1orf24 correctly classified 24 of 27 FTA and 25 of 25 FTC. C1orf24 was immunoreactive in >50% of the cells in most FTCs, moderate staining in a few FTCs, and with weak staining in three FTAs. Therefore, for screening for FTC, C1orf24 is 100% sensitive with 10.8% false-positive rate. ITM1 correctly classified all FTCs and 23 of 27 FTAs (representative results are shown in Fig. 2). ITM1 had a strong staining in all FTCs and moderate staining in three FTAs. Interestingly, the three FTAs that were positive for ITM1 and C1orf24 were also previously misclassified samples by DDIT3 and ARG2 (17). Statistical analysis of the immunohistochemistry results from the four markers showed high sensitivity and specificity for all four markers (Table 1). Although no markers could be declared significantly better than any other (note overlapping CIs in Table 1), we note that the point estimates for sensitivity of ITM1 and C1orf24 were 1.0, whereas the point estimates of sensitivity for DDIT3 and ARG2 were 0.85. The analysis of all four markers in combination (Classification and Regression Trees) showed that one marker (C1orf24) was sufficient to distinguish among the classes, and additional markers did not add much discriminatory power for this sample set.

We next examined the expression of all markers in other benign (hyperplastic nodules) and malignant thyroid lesions (FVPTC and PTC), which can be a source of diagnostic error.

### Table 1. Estimates and 95% CIs of sensitivity and specificity for each marker

<table>
<thead>
<tr>
<th>Marker name</th>
<th>FTC vs FTA</th>
<th>Malignant vs benign</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity (95% CI)</td>
<td>Specificity (95% CI)</td>
</tr>
<tr>
<td>DDIT3</td>
<td>0.85 (0.68-0.94)*</td>
<td>0.90 (0.76-0.97)¹</td>
</tr>
<tr>
<td>ARG2</td>
<td>0.85 (0.68-0.94)*</td>
<td>0.90 (0.76-0.97)¹</td>
</tr>
<tr>
<td>ITM1</td>
<td>1.00 (0.87-1.00)</td>
<td>0.85 (0.68-0.94)</td>
</tr>
<tr>
<td>C1orf24</td>
<td>1.00 (0.87-1.00)</td>
<td>0.89 (0.72-0.96)</td>
</tr>
</tbody>
</table>

*NOTE: Benign: hyperplastic nodule (n = 33) and FTA (n = 27). Malignant: FTC (n = 25), FVPTC (n = 27), and PTC (n = 15).  
*¹n = 23.  
*²n = 32.
The four markers have strong staining in >50% of the cells in most of FVPTC and PTC and no staining in most of hyperplastic nodules, although three hyperplastic nodules had a moderate staining (Fig. 2; Table 1).

When hyperplastic nodules, FVPTC, and PTC were considered along with FTC and FTA, again all four markers showed high sensitivity and could not be statistically distinguished from one another (although the CIs for some pairs of markers barely overlap). However, it is important to emphasize that ITM1 identified all malignant tumors in this sample; the point estimate for sensitivity of ITM1 was 1.0 with a 95% CI of 0.95 to 1.0 (Table 1). Although ITM1 correctly classifies all FVPTCs and PTCs, four hyperplastic nodules were misclassified as malignant.

The point estimate of sensitivity for C1orf24 was lower than ITM1, but the specificity was higher. C1orf24 misclassified two PTCs (2 of 15) and three hyperplastic nodules (3 of 33; Table 1). ARG2 correctly classified 24 of 27 of FVPTC, 30 of 33 of hyperplastic nodules, and 13 of 15 of PTC; DDIT3 correctly classified 23 of 27 of FVPTC, 30 of 33 of hyperplastic nodules, and 12 of 15 of PTC cases (Table 1).

Validation of markers in other benign and malignant thyroid lesions by quantitative PCR. We tested whether or not these new markers could correctly classify other benign and malignant thyroid lesions (n = 33) at transcript levels. The values obtained from quantitative PCR were used for statistical analysis and are summarized in Fig. 3. In the comparison between hyperplastic nodules (n = 22) versus FVPTC (n = 11),

Fig. 3. Relative levels of expression determined by quantitative PCR in 22 samples of hyperplastic nodules (HN) and 11 FVPTC (black columns). Quantitative PCR reactions were done in triplicate, averaged (SD ≤ 1), and normalized to the average of two reference genes (RS8 and OP-C).

Although the comparison between hyperplastic nodules versus FVPTC identified three genes as significantly different (DDIT3, ARG2, and C1orf24), ITM1 was just below the threshold for significance.
three genes were found to be significantly different, using an α level that had been adjusted (α level of 0.0125): DDIT3 \( (P = 0.000016) \), ARG2 \( (P = 0.0062) \), C1orf24 \( (P = 0.00034) \). ITM1 was very close to significance \( (P = 0.014) \).

We then investigated the development of a class predictor. Using leave-one-out cross-validation, 28 of the 33 samples were correctly predicted, for a prediction accuracy of 85% with a two-sided 95% CI of 0.68 to 0.95. In each of the 33 cross-validations, the same three genes (DDIT3, ARG2, and C1orf24) were always selected. To assess the significance of these prediction results, we implemented a permutation test. In none of the 1,000 random permutations were five or fewer samples misclassified, resulting in \( P < 0.001 \). Thus, we declared the results of the prediction analysis significant.

**Investigation of the molecular markers on FNA samples.** We tested our markers on six FNA specimens; five of six patients underwent surgery based on clinical and cytologic evaluation. The immunocytochemistry results were compared with the final histology. Two of six nondiagnostic FNA cases showed positive staining for all markers, which were later diagnosed as a FVPTC and a PTC on final histology. The cytoplasm of most of the cells was stained, consistent with a cell membrane (ITM1) or cytoplasmic protein (ARG2 and C1orf24). For DDIT3, both cytoplasm and nucleus were stained. The third case, a medullary thyroid carcinoma on final histology, was positive for ARG2 and DDIT3. Although immunocytochemistry was done for C1orf24 and ITM1 with positive immunostaining, few cells were present. Two cases showed negative staining for all markers, which were later diagnosed as a colloid goiter and a hyperplastic nodule on final histology. The last case was positive for all markers, but the final histology is not available because the patient refused surgery. Representative results are shown in Fig. 4.

**Discussion**

Currently, the standard diagnostic tool for thyroid nodules is FNA cytology. Now that more thyroid nodules are being detected by high-resolution ultrasound, it is important to increase the accuracy of subsequent diagnostic tests. Although the overall sensitivity and specificity of FNA cytology shows a wide range, FNA has a sensitivity of 25% to 42% for detection of FTC and Hürthle cell carcinoma (12). Therefore, benign and malignant follicular thyroid lesions cannot be distinguished by FNA cytology alone. This lack of accurate diagnosis of many thyroid nodules substantially increases health care costs, leads to avoidable surgery and, perhaps most significantly, may indirectly lead to delayed treatment of malignancies.

We have previously reported that DDIT3, ARG2, C1orf24, and ITM1 transcripts differed between FTA and FTC, and that a linear combination of expression levels had an estimated predictive accuracy of 83% (17). Additionally, we detected DDIT3 and ARG2 with immunohistochemistry and found a concordance with final histology diagnosis of 76% (17).
Because our goal is to improve test accuracy using a small number of antibody markers adaptable to FNA samples, we produced and tested antibodies for C1orf24 and ITM1 in FTA and FTC samples previously evaluated by DDIT3 and ARG2. We also tested other benign and malignant thyroid lesions that frequently provide a challenge to the pathologist.

All four markers were effective in distinguishing FTC from FTA. The point estimates of sensitivity for C1orf24 and ITM1 were both 1.0, with point estimates of specificity equal to 0.87 and 0.85 respectively. Although ITM1 does not provide additional discriminatory information beyond C1orf24, it emerges as a surrogate for C1orf24 (i.e., if C1orf24 was not available, it would be a reliable substitute). Of note, by using these antibodies, we were able to correctly classify two minimally invasive follicular thyroid carcinomas previously misclassified using DDIT3 and ARG2 antibodies alone.

When other benign and malignant lesions were considered along with FTAs and FTCs, all four markers proved effective at discriminating between the two classes. ITM1 could discriminate between benign and malignant groups with an estimated sensitivity of 1.0 (95% CI, 0.95-1.0) and estimated specificity of 0.87 (95% CI, 0.76-0.93). The estimates for C1orf24 were 0.97 (95% CI, 0.90-0.99) for sensitivity and 0.90 (95% CI, 0.80-0.95) for specificity. Although the evidence suggests ITM1 may be more reliable, overlapping CIs suggest that eliminating any of the other three markers is premature, in particular, until a wider range of patient populations are evaluated. Fortunately, any disagreements with histology were conservative errors; there were no cases where a malignant tumor was missed.

We also sought to determine whether transcript levels (rather than protein) from these four markers could reliably distinguish benign from malignant thyroid lesions. When the class predictor markers were tested in hyperplastic nodules and FVPTCs using quantitative PCR, C1orf24, DDIT3, and ARG2 were considered significantly different, and ITM1 was very close to significance. Although transcript levels were also accurate predictors, it was not superior to immunohistochemistry. One difference we observed was that ITM1 was detected in all FVPTC and FTC by immunohistochemistry but not in all FVPTC by quantitative PCR. Our finding suggested that the antibody-based test provided high sensitivity for this test. We advocate the use of an antibody-based test also because of its ability to detect malignant cells in situ and a test more easily implemented by a larger number of pathologists.

We next sought to test whether antibodies could effectively be used in FNA specimens. For the five FNA cases with a follow-up diagnosis, the molecular analysis was consistent with histologic findings. Interestingly, one case with a final diagnosis of medullary thyroid carcinoma stained positive. The diagnosis of medullary thyroid carcinoma is a difficult diagnosis to make preoperatively and usually requires confirmation by immunohistochemical staining for calcitonin, even by experienced endocrine pathologists (26). Future analysis testing these markers in a greater number of medullary thyroid carcinoma will clarify whether or not our markers could identify malignant lesion originating from C cells.

If successfully developed, a simple, easily employed, and low-cost antibody-based test could be incorporated in most laboratories for the routine evaluation for thyroid nodules, and the interpretation of immunohistochemistry stained sections is within the expertise of most pathologists. This would allow clinicians to better indicate to the patient the risk of malignancy and prompt quicker action for high-risk cases. Additionally, immunohistochemistry done in paraffin-embedded sections could be used as an adjunct to identify FTC that might be overlooked by the pathologist.

An important question that emerges from this study is if and how DDIT3, ARG2, ITM1, and C1orf24 are related to thyroid carcinogenesis because their expression is significantly linked to malignancy in this tumor. Reported functions of these genes provide some insight that they may be markers for the malignant progression of these tumors.

The mitochondrial ARG2 hydrolyzes arginine to ornithine plus urea and effectively down-regulates nitric oxide synthesis. The function of ARG2 is not fully understood, but high levels of ARG2 have been described in tumors and cell lines (27, 28). ARG2 was recently found overexpressed in cell lines derived from spontaneous murine mammary tumor and pulmonary metastases, and the authors suggested that ARG2 play a major role in tumor angiogenesis (29). These findings and the fact the ARG2 was found among the up-regulated tumor-related genes identified in a mice model that spontaneously develop FTC through progression of hyperplasia, capsular and vascular invasion, anaplasia, and eventually metastasis to distant organs (30) suggest that ARG2 may play a role in carcinogenesis.

DDIT3 has been found induced during organogenesis, inflammation, and also in response to a variety of cellular stress including glucose deprivation, oxidative stress, and endoplasmic reticulum stress (31, 32). DDIT3 was found induced by anoxia in breast cell lines and by hypoxia in pulmonary artery smooth muscle cells (32–34). DDIT3 may be induced by hypoxia or anoxia in malignant thyroid tumors. Other markers of hypoxia, such as carbonic anhydrase IX, are markers of more aggressive malignancies in a variety of cancer types. Interestingly, DDIT3 was one of the genes up-regulated in FTCs and Hurthle tumors (35, 36) and in a thyroid carcinoma cell line with P53 mutated (37).

ITM1 encodes a highly conserved integral membrane protein that contains 10 to 14 membrane-spanning domains. It is likely involved in protein glycosylation, but its relation to malignant progression is unknown.

The function of C1orf24 and the pathways through which it acts also remain to be elucidated. C1orf24 does, however, contains a DnaJ motif, a feature of heat shock protein that was first described as an oncogene associated with renal carcinogenesis (24).

Although gene expression profiling by cDNA microarray has been used to locate candidate markers that can be used as a diagnostic tool to evaluate thyroid nodules (35, 38–41), there is no single reported marker with the same or higher estimates of sensitivity and specificity as those reported here for C1orf24 or ITM1. However, a reported classification model, based on the expression levels of three genes (PCSK2, CCDN2, and PLAB) allowed the authors to discriminate between FTAs and FTCs with a sensitivity of 100% and specificity of 96.7% (42). PCSK2 and CCDN2 were highly expressed in FTAs, whereas PLAB was highly expressed in FTC. Our SAGE results indicated that PCSK2 was highly expressed in FTC but not expressed in FTA and normal thyroid. Further investigation by quantitative PCR revealed that this marker did not distinguish FTAs from FTCs because, in our hands, it was expressed in about 13% of FTCs and 10% to 40% of FTAs (17). Therefore, our results
predict that immunohistochemistry using one or more of our markers would provide the most accurate diagnostic test to augment FNA cytology or histology of thyroid tumor sections.

In summary, we have shown that all four antibodies are reliable for testing thyroid nodules and indicate that the test will work on FNA samples. We are further improving this test by evaluating additional markers for adenomas, which we hope will improve specificity. To date, the decision about a treatment regimen has been predominantly based on clinical and cytologic evaluation of a thyroid nodule. Our results support the testing of these markers in a larger trial to determine if FNA thyroid nodule diagnosis can be improved.

References

Clinical Cancer Research

Diagnosis of Suspicious Thyroid Nodules Using Four Protein Biomarkers


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/12/11/3311

Cited articles
This article cites 40 articles, 12 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/12/11/3311.full.html#ref-list-1

Citing articles
This article has been cited by 7 HighWire-hosted articles. Access the articles at:
/content/12/11/3311.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.