Clinical Science

Evaluation of genetic biomarkers for distinguishing benign from malignant thyroid neoplasms

Sapna Nagar, M.D.¹,*, Samreen Ahmed, M.P.H.², Claire Peeples, M.D.¹, Nichole Urban, M.D.¹, Judy Boura, M.S.³, Bryan Thibodeau, Ph.D.², Jan Akervall, M.D., Ph.D.¹,², George Wilson, Ph.D.²,², Graham Long, M.D.¹, Peter Czako, M.D.¹

¹Department of Surgery Beaumont Health System, 3811 West 13 Mile Road, 104-RI, Royal Oak, MI 48073, USA; ²Beaumont BioBank, ³Department of Biostatistics, and ⁴Department of Radiation Oncology, Beaumont Health System, Royal Oak, MI, USA

KEYWORDS:
Follicular thyroid cancer; Fine-needle aspiration; Gene expression; Real-time polymerase chain reaction

Abstract

BACKGROUND: Fine-needle aspiration (FNA) aids in the diagnosis of thyroid nodules. The expression of previously implicated genes was examined to potentially discriminate between benign and malignant thyroid samples.

METHODS: Patients included for study had cytology demonstrating follicular cells of undetermined significance, atypical cells of undetermined significance, follicular neoplasm, or suspicion of malignancy with one of the following postoperative diagnoses: follicular thyroid adenomas, follicular thyroid carcinomas, or follicular variant of papillary thyroid carcinomas (FV-PTCs). FNA and tumor expression of human telomerase reverse transcriptase (hTERT), high-mobility group A2 (HMGA2), and trefoil factor 3/3-galactoside-binding lectin (T/G ratio) were analyzed.

RESULTS: T/G ratios were not significantly different in the malignant and benign groups. HMGA2 was overexpressed in carcinoma states; however, only FV-PTCs were significant (P < 0.006). Tumor hTERT expression was detected in 25% of follicular thyroid carcinomas, whereas 5% of FV-PTCs and 10% of follicular thyroid adenomas had expression. FNA aspirates showed similar results.

CONCLUSIONS: Although HMGA2 and hTERT showed differential expression, they did not consistently differentiate benign from malignant. Further study based on global gene expression is needed to identify markers that could serve as a diagnostic tool.

© 2014 Elsevier Inc. All rights reserved.

Thyroid nodules are common, affecting as many as 4% to 7% of the adult population.¹ Of all thyroid nodules detected clinically, only about 5% to 15% are found to contain malignancy.² Although thyroid carcinoma accounts for only 1% of all new malignancies, most of these (approximately 90% to 94%) will manifest as differentiated thyroid carcinomas.²,³ The ability to differentiate between benign and malignant thyroid conditions remains a challenge. Currently, the most cost-effective and accurate method of determination involves the performance of fine-needle aspiration (FNA).²

FNA is a simple and reliable procedure that is accepted as a standard method for diagnosing a thyroid nodule. However, there are limitations in characterizing cellular...
architecture, and, thus, malignancy cannot be assessed. The Bethesda System was developed to further stratify indeterminate lesions. However, only 20% of patients with inconclusive pathologic findings will have follicular thyroid carcinoma (FTC) on surgical pathology. The ability to identify molecular markers from analysis of FNA samples would be a useful diagnostic tool that could accurately differentiate follicular thyroid adenoma (FTA) from FTC before surgical intervention and could ultimately save the morbidity and cost associated with surgery.

Because an FNA sample has a small cell yield, a sensitive test such as quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR) may be a method that could prove practical in the clinical arena because it can quickly detect gene expression with limited amounts of material. Previous studies conducted on various thyroid tissue specimens have established that qRT-PCR of different genes may differentiate benign from malignant follicular lesions. Foukakis et al. evaluated 26 molecular markers in thyroid tissue using qRT-PCR and proposed 2 models implementing specific genes (including human telomerase reverse transcriptase [hTERT] and trefoil factor 3 [TFF3]) to determine benign versus malignant expression patterns. Notably, most of the studies using this method did not use an FNA sample, which limits the use of this technique during preoperative assessments.

After an extensive literature review, we identified studies that investigated specific gene models that are potential candidate markers to distinguish FTC from FTA. hTERT is a ribonucleoprotein polymerase that directs chromosomal telomerase and is down-regulated in normal cells. However, it has been implicated in tumor growth in both malignant thyroid tissue and FNA samples. TFF3 is a family of peptides with a 3-loop trefoil domain that are mainly synthesized and secreted by epithelial cells. Although the function of these proteins is not yet well understood, previous studies have linked decreased expression to thyroid follicular carcinoma. A third biomarker, high-mobility group A2 (HMGA2) proteins, is a group of DNA-binding proteins that regulate gene transcriptional activity during embryogenesis. The expression of these proteins is low in normal tissue; however, a high expression has been implicated in various malignancies including thyroid neoplasia.

By using potential biomarkers for thyroid carcinoma that have been identified in previous studies (hTERT, T/G ratio, and HMGA2), we sought to determine whether genetic expression could provide diagnostic accuracy in distinguishing FTA from FTC using prospectively collected samples. This modality may be useful in the clinical setting, and in the future, it may serve to direct more targeted future therapies for the thyroid nodule patient.

**Patients and Methods**

The study was approved by our institutional review board, and consent was obtained from all patients before testing. From October 2009 to June 2011, we included in our study all consecutive patients with thyroid nodules showing a follicular lesion of undetermined significance or atypical cells on undetermined significance, follicular neoplasm, and suspicion for malignancy with one of the following postoperative pathologies: follicular or Hürthle cell adenoma, follicular or Hürthle cell carcinoma, or papillary carcinoma of follicular variant (FV-PTC). Patients were recommended surgical intervention in the form of thyroid lobectomy, total thyroidectomy, or lobectomy followed by completion thyroidectomy based on current guidelines and a discussion between the patient and surgeon. Immediately after surgical resection of the specimen, an FNA with a 22-G needle was performed by the surgeon on the ex vivo thyroid tumor in the operating room. To distinguish tumor from normal tissue, gross specimens were obtained by visual inspection and determined by pathology staff. All specimens were placed in RNAlater (Qiagen, Valencia, CA) at the time of retrieval and stored at −80°C until study. Histopathologic diagnosis of all specimens was then confirmed by an independent pathologist.

**Genetic markers**

The following predesigned TaqMan gene-specific primers (Applied Biosystems, Foster City, CA) were used: HMGA2 (assay ID Hs00171569m1), TFF3 (assay ID Hs00173587m1), hTERT (assay ID Hs00162669m1), 3-galactoside-binding lectin (assay ID Hs00173587m1), and gene glyceraldehyde-3’-phosphate dehydrogenase (GAPDH) (assay ID Hs99999905m1). These genes were selected based on previous studies showing promise in the stratification of follicular carcinomas and adenomas.

**RNA extraction and reverse transcription**

Tumor tissues were prepared for RNA extraction by Polytron (Polytron 1200, Kinematica AG, Lucerne, Switzerland) homogenization of 5 mg tissue. FNA aspirates were strained through a 35-μm cell strainer (BD Biosciences, San Jose, CA) and then centrifuged at 150g for 1 minute. Retained cellular material from the FNA or tumor tissue was then subjected to RNA extraction using an RNeasy Mini Kit (Qiagen, Valencia, CA). Complementary DNA (cDNA) was then synthesized by reverse transcription using 40 to 100 ng RNA with 250 ng/μL random hexamers and 200 U SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturer’s protocol.

**Quantitative reverse-transcriptase polymerase chain reaction**

Secondary to the limited amount of cDNA available, samples were preamplified before amplification (Applied Biosystems). In brief, this was performed by incubating cDNA with TaqMan PreAmp MasterMix (Applied
Statistical analysis

All analyses were completed using SAS System for Windows version 9.2 (SAS Institute Inc, Cary, NC). Missing data were omitted with no substitutions or imputations. Demographic and clinical data including age, sex, medical history, and surgical intervention were analyzed. The Fisher exact test and the nonparametric Kruskal-Wallis test were used to examine specific variables.

Concordance between expression in FNA and tissue samples was measured using the GAPDH gene for normalization. Gene expression data were divided into numerical quartiles to account for large jumps in the data. For symmetric comparisons, weighted $\kappa$ statistics with 95% confidence intervals were calculated; the $\kappa$ coefficient can range from 1 to $-1$ with the following indication: less than .4 indicates “poor agreement”, .4 to .75 indicates “acceptable agreement,” and greater than .75 indicates “excellent agreement.” When too few comparisons existed to categorize, the presence or absence of expression was compared using simple $\kappa$ statistics and the McNemar test of agreement.

Significance for FTA versus either FTC or FV-PTC was determined using Wilcoxon rank tests. Because of the small volume of cells in the clinical FNA samples, the RNA expression of the TFF3 gene was measured as a relative expression of TFF3 messenger RNA (mRNA) to galectin-3 mRNA (T/G ratio) as previously described by Takano et al.8 Galectin-3 mRNA is used as an internal control of thyroid-derived cells because of its expression in both benign and malignant thyroid tumors.8,9

Results

Demographic data

A total of 52 patients were included in this study: 21 had FTAs, 12 had FTCs, and 19 had FV-PTCs on postoperative pathology. Women constituted the majority of patients in all groups (83%), and the mean age in years of the entire group was 53.3. A history of previous cancers, family history of thyroid disease, and history of hypo-/hyperthyroidism were similar between groups. Patients with a diagnosis of carcinoma on FNA were more likely to undergo total thyroidectomy or lobectomy with completion thyroidectomy ($P < .0001$). Patient demographics are listed in Table 2.

Genetic expression

We examined mRNA expression of the 3 genes relative to GAPDH (Table 3). In the tumor samples, a nearly 3-fold reduction of the T/G ratio was observed in both FTCs and FV-PTCs relative to FTAs. However, this decrease was not observed in the FNA samples. Next, a T/G ratio cutoff point of 16.0 (as identified previously by Takano et al8 through the construction of a receiver operating characteristic curve) was used to distinguish follicular adenomas from carcinomas (Fig. 1). Applying this cutoff point in FNA aspirates, the T/G ratio was decreased in 8 (67%) of 12 FTCs, 8 (42%) of 19 FV-PTCs, and 9 (43%) of 21 FTAs. In tumor samples, the T/G ratio was lower in 6 (50%) of 12 FTCs, 9 (47%) of 19 FV-PTCs, and 9 (43%) of 21 FTAs. The concordance of T/G ratio expression between tumor tissue and FNA aspirates suggests a slight agreement ($\kappa = .31 \ [ .11 \ to \ .51 ]$).

HMGA2 gene expression in tumor tissue was significantly increased with a 28-fold up-regulation in patients with FV-PTCs compared with patients with FTAs ($P = .006$). An 11-fold increase in expression was also observed in FV-PTC patients relative to FTA patients in the FNA samples ($P = .39$). A 2-fold increase in the expression of HMGA2 was observed in the FTA samples compared with the FTA and FNA samples (Fig. 2). There was some agreement between HMGA2 tumor and FNA expression ($\kappa = .48 \ [ .26 \ to \ .69 ]$).

Because of a limited number of samples that showed expression of hTERT, absence versus presence of gene expression was analyzed. In the tumor samples, hTERT gene expression was observed in 2 (10%) of the FTA samples and 3 (25%) of the FTC samples. There were no FV-PTC

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Amplification efficiency (%)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFF3</td>
<td>96</td>
<td>.999</td>
</tr>
<tr>
<td>LGALS5</td>
<td>95</td>
<td>1.00</td>
</tr>
<tr>
<td>hTERT</td>
<td>92</td>
<td>.999</td>
</tr>
<tr>
<td>HMGA2</td>
<td>97</td>
<td>.998</td>
</tr>
<tr>
<td>GAPDH</td>
<td>91</td>
<td>.999</td>
</tr>
</tbody>
</table>

GAPDH = glyceraldehyde-3’-phosphate dehydrogenase; HMGA2 = high-mobility group A2; LGALS = 3-galactoside-binding lectin; PCR = polymerase chain reaction; hTERT = human telomerase reverse transcriptase; TFF = trefoil factor.
tumor samples with \textit{hTERT} expression. When examining \textit{hTERT} expression in the FNA aspirates, 2 (10\%) in FTAs, 1 (5\%) in FV-PTCs, and 3 (25\%) in FTCs showed the presence of expression (Fig. 3). In detecting the presence of \textit{HMGA2} expression, there was some agreement between tumor and FNA samples ($\kappa = 0.49$ [0.11 to 0.88]).

**Comments**

FNA is essential in the workup of a thyroid nodule.\textsuperscript{1,2} However, there are limitations of this technique’s diagnostic capability for follicular lesions. The primary aim of this study was to validate specific genetic markers already established in the literature as possible biomarkers of thyroid malignancy and to distinguish between benign and malignant follicular thyroid disease in FNA specimens. Genetic markers have been identified as possible markers for FTC, but the methodologies are limited in the clinical setting because of the limited number of samples from FNA. Studies using cDNA microarray assays to identify biomarkers that are expressed in the carcinoma specimens compared with normal tissue have shown that this is a viable strategy for developing molecular diagnostics. However, these studies analyzed actual tumor tissue and not FNA samples.\textsuperscript{7,16–18} If markers could be detected in FNA samples, malignant disease could be detected preoperatively. An advantage of qRT-PCR analysis of FNA samples is the ability to reliably detect these differences with small amounts of RNA.

In our cohort, women constituted the majority of the patients with thyroid pathology. Of the 12 patients diagnosed with FTC, I was found to have a widely invasive FTC, whereas the others were found to have minimally invasive FTCs on postoperative pathology. It was expected and confirmed by our study that patients who were diagnosed with carcinoma in FNA (with either FTCs or FV-PTCs) were more likely to proceed to either total thyroidectomy or partial lobectomy with subsequent completion thyroidectomy compared with patients with a preoperative diagnosis of adenomas, which requires simple lobectomy.

One objective of this study was to determine the concordance in genetic expression between tumor tissue and FNA samples. To show the validity of genetic alterations observed in the FNA of the adenomas and carcinomas, we compared these values with the genetic expression of the same genes within the tumor tissue of adenomas and carcinomas. Because of the large variation in expression levels, it was necessary to group concordance values within quartiles. Comparing tumor tissue and FNA samples, we found only a slight concordance in the \textit{HMGA2} gene and no agreement or concordance using the T/G ratio. We were not able to determine concordance using \textit{hTERT} because of the small number of samples with expression. These findings indicate that the use of qRT-PCR on FNA samples may be sufficient but not equivalent to performing similar studies on actual tissue samples.

Once the concordance was examined, we looked at genetic expression in all samples of \textit{HMGA2}, \textit{TFF3}, and

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Comparison of patient demographics with final histopathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>FTA</td>
</tr>
<tr>
<td>Total, n (%)</td>
<td>21 (40)</td>
</tr>
<tr>
<td>Age (mean, y)</td>
<td>52 ± 13</td>
</tr>
<tr>
<td>Sex</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Female</td>
</tr>
<tr>
<td>Surgical procedure</td>
<td>Lobectomy only</td>
</tr>
<tr>
<td></td>
<td>Total thyroidectomy</td>
</tr>
<tr>
<td></td>
<td>Lobectomy with completion</td>
</tr>
</tbody>
</table>

FTA = follicular thyroid adenoma; FTC = follicular thyroid carcinomas; FV-PTC = follicular variant of papillary thyroid carcinomas.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>The median expression relative to FTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
<td>Tissue type</td>
</tr>
<tr>
<td>T/G</td>
<td>Tumor</td>
</tr>
<tr>
<td></td>
<td>FNA</td>
</tr>
<tr>
<td>\textit{HMGA2}</td>
<td>Tumor</td>
</tr>
<tr>
<td></td>
<td>FNA</td>
</tr>
<tr>
<td>\textit{hTERT}</td>
<td>Tumor</td>
</tr>
<tr>
<td></td>
<td>FNA</td>
</tr>
</tbody>
</table>

FNA = fine-needle aspiration; FTA = follicular thyroid adenoma; FTC = follicular thyroid carcinoma; \textit{HMGA2} = high-morbidity group A2; \textit{hTERT} = human telomerase reverse transcriptase; T/G = trefoil factor 3/3-galactoside-binding lectin.

* signifies $p < 0.006$.

† signifies insufficient samples with \textit{hTERT} expression, median expression relative to FTA was not calculated.
hTERT after qRT-PCR. Previously, Belge et al\textsuperscript{13} showed HMGA2 overexpression in thyroid carcinoma tissue compared with benign disease. Similarly, we found the HMGA2 gene to be overexpressed in FV-PTCs compared with FTAs in the tumor samples only. Expression in the tumor tissue was over 25-fold higher in the carcinoma state, which was statistically significant ($P < .006$).

In examining the gene expression of TFF3, we did not observe a decrease in the T/G ratio in either the tumor tissue or FNA of FTCs as in previous studies.\textsuperscript{8,9} Although there was a trend toward a decreased ratio, this was not significant. In addition, we found a sensitivity of 86\% for TFF3 in the FNA samples and a specificity of 16\%. For tumor tissue, there was both a low sensitivity and specificity for TFF3. It is possible that the TFF3 gene might be useful in identifying carcinomas but may not be useful to rule out benign disease and is therefore not a reliable marker.

A possible explanation for the detection of underexpression of TFF3 in a small subset of patients with FTC is the transformation of previously determined benign lesions into malignant lesions. The significance of this underexpression in the benign state may indicate a precancerous FTA. Because these lesions have been resected, we cannot know the natural history of this subset of FTAs and their potential to develop malignancy in the future. Therefore, there may be greater value in identifying markers that could reliably rule out the benign state because a majority of indeterminate follicular lesions are benign.

Saji et al\textsuperscript{5} looked at hTERT expression in FNA samples and found that there was no expression in benign thyroid tissue, whereas there was increased expression in carcinoma samples. We were not able to complete a similar analysis of hTERT as with the other genes because of an inconsistent expression pattern of hTERT in all samples. We did observe slightly more expression in tumor and FNA samples of carcinoma when compared with the adenomas. This difference was not significant, and, therefore, we cannot draw any conclusions to suggest hTERT as an adequate marker of disease.

There are some limitations to this study. The largest was the modest number of samples studied, limiting the power of the study and creating the possibility of type 2 errors. A larger sample size is needed to confirm the trends observed in this study. We were unable to further stratify and analyze the FTC samples into minimally invasive versus widely invasive carcinomas because of only 1 sample with widely invasive pathology. Separating these pathologic groups could show different outcomes both on the molecular level and the clinical level. Hypotheses have been proposed indicating that minimally invasive FTCs may behave similarly to FTAs. As a result, any genetic alterations may not be observed in our particular subset of patients with FTCs.\textsuperscript{10} Similarly, it must be noted that the FNAs were performed ex vivo as opposed to in vivo as in the normal clinical setting. To extrapolate data from this to the clinical setting would require further study to show similar findings using both methodologies.

In summary, the usage of gene expression analysis in FNA samples showing follicular or Hürthle cells to determine whether the lesions truly represent benign or malignant disease seems plausible, especially with the use of qRT-PCR. This would be useful in the clinical setting in

![Figure 1](image1.png)  
**Figure 1** mRNA expression of T/G expression relative to GAPDH. T, tumor; F, FNA.

![Figure 2](image2.png)  
**Figure 2** mRNA expression of HMGA2 relative to GAPDH shown in a logarithmic scale (log10). T, tumor; F, FNA.

![Figure 3](image3.png)  
**Figure 3** Percent expression of hTERT among tumor and FNA samples.

FTA is the transformation of previously determined benign lesions into malignant lesions. The significance of this underexpression in the benign state may indicate a precancerous FTA. Because these lesions have been resected, we cannot know the natural history of this subset of FTAs and their potential to develop malignancy in the future. Therefore, there may be greater value in identifying markers that could reliably rule out the benign state because a majority of indeterminate follicular lesions are benign.
which information about the malignant potential of such lesions could be determined before surgical intervention. Findings from this study fail to support the relationship of TFF3, HMGA2, or hTERT genes with follicular thyroid carcinoma when analyzed separately; however, future studies are needed to investigate these genes in combination with others as potential markers for malignancy. Our results suggest that there is not an obvious correlation between the expression of the selected genes found in tumor samples and samples obtained by FNA. Discerning why there is a lack of concordance could be an important step in better using FNA for tumor diagnosis.

References