Combined immunostaining with galectin-3, fibronectin-1, CITED-1, Hector Battifora mesothelial-1, cytokeratin-19, peroxisome proliferator-activated receptor-γ, and sodium/iodide symporter antibodies for the differential diagnosis of non-medullary thyroid carcinoma

Ying Y Liu, Hans Morreau1, Job Kievit2,3, Johannes A Romijn, Nancy Carrasco4 and Johannes W Smit

Department of Endocrinology, C4-R, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands, Departments of 1Pathology, 2Medical Decision Making and 3Surgery, Leiden University Medical Center, NL-2300RC Leiden, The Netherlands and 4Department of Molecular Pharmacology, Albert Einstein College of Medicine, Bronx, New York, NY-10461 USA

(Correspondence should be addressed to J W A Smit; Email: jwasmit@lumc.nl)

Abstract

Objectives: The microscopic distinction between benign and malignant thyroid lesions in clinical practice is still largely based on conventional histology. This study was performed to evaluate the diagnostic value of galectin-3 (Gal-3), Hector Battifora mesothelial-1 (HBME-1), cytokeratin (CK)-19, CBP P300-interacting transactivator with glutamic acid E- and aspartic acid D-rich C-terminal domain (CITED-1), fibronectin (FN)-1, peroxisome proliferator-activated receptor (PPAR)-γ, and intracellular sodium/iodide symporter (iNIS) immunostaining in a large panel of thyroid neoplasms. Our study differed from earlier ones with regard to the identification of optimal semiquantitative cut-off levels using receiver operator curve (ROC) analysis and hierarchical cluster analysis.

Methods: We used tissue arrays containing 177 thyroid tissues: 100 benign tissues (including normal thyroid, Graves disease, multinodular goiter, and follicular adenoma (FA)) and 77 thyroid carcinomas (including papillary thyroid carcinoma (PTC), follicular thyroid carcinoma, and follicular variant of PTC (FVPTC)). Antibody staining was scored semiquantitatively based on the ROC analyses and with hierarchical cluster analysis.

Results: In general, we found overexpression of FN-1, CITED-1, Gal-3, CK-19, HBME-1, and iNIS in malignant thyroid lesions. Gal-3, FN-1, and iNIS had the highest accuracy in the differential diagnosis of follicular lesions. A panel of Gal-3, FN-1, and iNIS, identified by hierarchical cluster analysis, had a 98% accuracy to differentiate between FA and malignant thyroid lesions. In addition, HBME-1 was found to be useful in the differentiation between FA and FVPTC (accuracy 88%).

Conclusion: We conclude that identifying optimal antibody panels with cluster analysis increases the diagnostic value in the differential diagnosis of thyroid neoplasms, the combination of FN-1, Gal-3, and iNIS having the best accuracy (98%).
in a considerable percentage of malignant thyroid tumors (30–32). This expression pattern may therefore be helpful in the distinction between benign and malignant lesions. In the pathogenesis of thyroid tumors, decreased expression of the peroxisome proliferator-activated receptor (PPAR)-γ has been reported (33–35), which may therefore also be used as a diagnostic marker.

Genetic expression studies confirmed the differential expression of Gal-3 and also identified the extracellular matrix component fibronectin-1 (FN-1) and CBP p300-interacting transactivators with glutamic acid–aspartic acid D-rich C-terminal domain (CITED-1) as specific markers for PTC (4, 12, 18).

As no marker by itself has a superior diagnostic value, a combination of markers may be more accurate than any single marker. A combined approach using a panel of HBME-1, Gal-3, and CK-19 was followed by Casey et al. (26), de Matos et al. (27), Scognamiglio et al. (18), Nakamura et al. (36), and Prasad et al. (37). In the study by de Matos et al. the combination had a limited value; in contrast, in the study by Prasad et al. (37), this panel had a high sensitivity and specificity for carcinomas.

In most of the aforementioned studies, fixed cut-off levels for positive staining were used. Therefore, we decided to evaluate the diagnostic value of Gal-3, HBME-1, CK-19, CITED-1, FN-1, NIS, and PPAR-γ in a large panel of thyroid neoplasms. We calculated the cut-off levels of positive staining for each antibody using receiver operator curve (ROC) analyses. We analyzed not only the diagnostic value of each individual marker but also the diagnostic accuracy of panels identified by hierarchical cluster analysis.

Materials and methods

Patients

One hundred seventy-seven histological samples from surgically removed thyroid lesions representing seven different histological thyroid disorders and adjacent normal thyroid tissue were obtained from the pathological archive of the Leiden University Medical Center, the Netherlands. We selected 100 benign thyroid tissues (normal: n = 64, Graves disease (GD): n = 10, multinodular goiter: n = 14, and FA all microfollicular: n = 12) and 77 non-medullary thyroid carcinomas (PTC: n = 53, FTC: n = 13, and follicular variant of PTC (FVPTC): n = 11). All original histological diagnoses were reviewed by two independent observers.

Tissue microarrays

Formalin-fixed, paraffin-embedded blocks routinely prepared from the surgical specimens of thyroid tumors were selected for this study. Representative areas containing tumor or adjacent normal tissues were identified by a pathologist (H M). Triplicate tissue cores with a diameter of 0.6 mm were taken from each specimen (Beecher Instruments, Silver Springs, MD, USA) and arrayed on a recipient paraffin block, using standard procedures (38).

Immunohistochemistry

Four micrometer consecutive tissue sections were cut from each arrayed paraffin block and prepared on pathological slides. The sections were deparaffinized in xylene followed by 0.3% hydrogen peroxide in methanol at room temperature for 20 min for blocking endogenous peroxidase. After rehydration, antigen retrieval treatment was done for CK-19, HBME-1, FN-1, CITED-1, NIS, and PPAR-γ by microwave treatment in 0.01 M citrate buffer (pH 6.0). After 2 h of cooling, endogenous avidin activity blocking was performed for NIS immunostaining by incubation with egg white for 5 min followed by biotin for 15 min. Then the sections were incubated with the following primary antibodies in PBS with 1% BSA overnight at room temperature: anti-Gal-3 monoclonal antibody (MAB) NCL-Gal3, dilution 1:200 (Novocastra, Newcastle, UK); anti-HBME-1 MAB M3505, dilution 1:50 (DakoCytomation, Glostrup, Denmark); anti-CK-19 polyclonal antibody, dilution 1:100 (DakoCytomation); anti-CITED-1 polyclonal antibody ab15096, dilution 1:100 (DakoCytomation); anti-FN-1 polyclonal antibody A0245, dilution 1:2000 (DakoCytomation); anti-NIS polyclonal antibody, dilution 1:200 (39); and PPAR-γ MAB sc-7273, dilution 1:30 (Santa Cruz, CA, USA). Next, the sections were incubated for 30 min with either the biotinylated rabbit anti-mouse conjugate, dilution 1:200, or goat anti-rabbit conjugate, dilution 1:400 (DakoCytomation); anti-Gal-3 monoclonal antibody (MAB) NCL-Gal3, dilution 1:200, followed by incubation for 30 min with the streptavidin–biotin peroxidase conjugate. This step was performed by means of a 10-min incubation with 3,3′-diaminobenzidine tetrachloride substrate in a buffered 0.05 M Tris/HCl (pH 7.6) solution containing 0.002% hydrogen peroxide. Negative controls were stained with the primary antibody omitted and the sections were counterstained with hematoxylin.

Scoring

A semiquantitative assessment of immunohistochemical staining was performed taking into account both the intensity of staining and the percentage of positive cells. The percentage of cells with positive staining was scored as follows: 0–20%: ‘0’, 20–50%: ‘1’, 50–70%: ‘2’, and 70–100% ‘3’. The intensity was scored as faint: ‘0’, intermediate: ‘1’, and intense: ‘2’. Scores for the proportion of positive cells and intensity were added. The total score per sample therefore reached 0–6. Score results for triplicate samples were averaged. Because a distinct intracellular distribution was observed for some
Figure 1 Immunostaining of thyroid tissues with Hector Battifora mesothelial-1 (HBME-1), fibronectin-1 (FN-1), galectin-3 (Gal-3), peroxisome proliferator-activated receptor (PPAR)-γ, CBP/p300-interacting transactivators with glutamic acid E- and aspartic acid D-rich C-terminal domain-1 (CITED-1), cytokeratin-19 (CK-19), and sodium/iodide symporter (NIS). Magnification was 200×. For immunohistochemical staining procedures, see the Materials and methods section. HBME-1 exhibited plasma membrane staining in papillary thyroid carcinoma (PTC) and follicular variant PTC (FVPTC) and was absent in follicular adenoma (FA). FN-1 displayed cytoplasmic staining in PTC. Gal-3 showed cytoplasmic or nuclear staining in thyroid carcinomas. PPAR-γ nuclear staining was observed in benign thyroid lesions. CITED-1 demonstrated cytoplasmic staining in benign and malignant thyroid lesions. CK-19 was overexpressed in PTC. Intracellular NIS was present in FTC and PTC, whereas typical plasma membrane staining was noted in Graves disease.
antibodies, their staining scores were categorized according to these patterns – NIS staining was differentially categorized as plasma membrane (pmNIS) or intracellular (iNIS). Accordingly, FN-1 was also identified as pmFN-1 and iFN-1. Gal-3 was classified as cGal-3 or nuclear Gal-3 (nGal-3). Examples of the staining patterns are depicted in Fig. 1.

**Statistical analyses**

Statistical analyses were performed using SPSS 12.0 (SPSS Inc, Chicago, USA). Initially, staining scores for every individual antibody were expressed as median and ranges per histological category. The next step was the analysis of differences in staining scores for every antibody between all histological categories in a $2 \times 2$ table, using the Kruskal–Wallis test. For each differentially expressed antibody between two histological subclasses, the optimal cut-off values for the distinction between the two histological classes were determined with the ROC analysis. In theory, this could provide different cut-off values for an antibody for different comparisons. Thereafter, the diagnostic validity for each antibody was analyzed for comparisons between histological subclasses, using the Bayesian statistics as sensitivity, specificity, and accuracy.

In addition to the individual protein markers, the analysis of the diagnostic accuracy of panels of antibodies was performed by means of the hierarchical clustering analysis of tissue microarray data using Cluster and TreeView (Cluster and TreeView 2.11; Eisen Lab, University of California at Berkeley, CA, USA) (40). Input for these analyses was the individual staining score per sample for each antibody. $P<0.05$ was considered significant.

**Results**

**Protein expression in thyroid lesions**

The median values, the ranges of the expression of the proteins, and the proportion of samples with staining scores above the cut-off levels are shown in Table 1. Statistically significant differences in protein expression between all the categories of thyroid tissues were investigated in $2 \times 2$ tables, the results of which are given in Table 2.

In general, malignant tumors showed overexpression of Gal-3 (predominantly PTC), iFN-1 (all carcinomas), CK-19 (mostly PTC), HBME-1 (mostly PTC and FTC), and iNIS (mostly PTC and FTC). By contrast, the expressions of PPAR-$\gamma$ and pmNIS were low or absent in thyroid carcinomas. In GD, the expression of pmNIS was abundant, as expected. PPAR-$\gamma$ was also higher in the adjacent normal tissues and benign thyroid lesions. The most prominent differences were observed in PTC.

![Table 1](https://www.eje-online.org)
in comparison with benign lesions and adjacent normal thyroid tissues – PTC showed high expression levels of iFN-1, iGal-3, iNIS, HBME-1, CITED-1, and CK-19 and the absence of PPAR-γ and pmNIS.

In the comparison between FTC and FVPTC, the only differentially expressed protein was HBME-1 (Table 2). FVPTC differed from FA for iFN-1, PPAR-γ, HBME-1, and CK-19, whereas protein expression was different for FN-1, Gal-3, HBME-1, and CK-19 in the comparison between FVPTC and PTC (Table 2). Intra-cellular NIS was mainly observed in PTC and FTC (Table 1). Remarkably, the differences in CITED-1 were not prominent between benign thyroid tissues and malignant lesions in 2×2 comparisons.

**Protein expression in follicular lesions**

Since the clinical distinction between follicular lesions proves to be most difficult, we analyzed the diagnostic value of proteins found to be differentially expressed in follicular lesions (Tables 2 and 3). The highest accuracies were found in the discrimination between PTC and FVPTC. In the comparison between FA and FTC, moderate accuracies were found for FN-1 and iNIS. The distinction between FA and FVPTC had a high accuracy for HBME-1, PPAR-γ, and FN-1. HBME-1 also gave a good discrimination between FVPTC and FTC (accuracy 84%).

**Clustered expression pattern of iGal-3, iFN-1, and iNIS distinguishes benign thyroid tumors from thyroid carcinomas**

To identify the optimal combinations of antibodies, we performed an unsupervised hierarchical cluster analysis including all the tissues and antibodies. The results of this analysis are given in Fig. 2. We found that three markers – iGal-3, iNIS, and iFN-1 – had the highest discriminating power to cluster benign and malignant thyroid lesions – all malignant lesions had positive expression of iGal-3, iNIS, and iFN-1.
staining for at least two out of the three antibodies, iGal-3, iNIS, and iFN-1 (cluster 1), whereas lesions with absent staining for all the three antibodies were all benign, except one (cluster 5). All malignancies were combined because the subgroups were too small to allow separate cluster analysis. We therefore used the combined staining patterns of these antibodies to discriminate between benign and malignant thyroid lesions and FA and malignant thyroid lesions (Table 4). We found that positive staining for two out of the three antibodies, iFN-1, iGal-3, or iNIS, had a high sensitivity (97–98%) and high specificity for thyroid carcinoma (100%).

Discussion

The present study was performed to evaluate the diagnostic values of Gal-3, HBME-1, CK-19, CITED-1, FN-1, PPAR-γ, and NIS staining in a large panel of thyroid neoplasms. Our study differed from the earlier ones with regard to the identification of optimal semiquantitative cut-off levels using the ROC analysis and hierarchical cluster analysis.

We initially analyzed the differentially expressed protein markers comparing all the thyroid tissues. In general, we found the overexpression of FN-1, CITED-1, Gal-3, CK-19, HBME-1, and iNIS in thyroid carcinomas, whereas pmNIS and PPAR-γ showed decreased expression in carcinomas in comparison with benign thyroid tissues.

The most challenging differential diagnosis is between FA and thyroid carcinoma. We found all studied proteins to be differentially expressed between FA and PTC. The differences between FA on the one hand and PTC and FVPTC on the other hand were less prominent, but we found a differential expression of PPAR-γ, HBME-1, Gal-3, iNIS, and FN-1. We did not confirm the differential expression of CITED-1 and CK-19 among FA, FVPTC, and FTC reported by Prasad et al. (37).

CK-19 is the most commonly used CK in investigating thyroid lesions. We and others found that CK-19 is relatively specific for PTC (18, 24, 26, 27). However, in our analyses, CK-19 had limited use in the differential diagnosis of follicular thyroid lesions. This has also been reported by Sahoo et al. (25). In the study by Prasad et al. (37), CK-19 had a sensitivity of 64% for thyroid carcinoma.

Several recent investigations have reported that HBME-1 expression is a useful diagnostic marker for FTC (4, 18, 37). We found HBME-1 expression predominantly in FTC and FVPTC and in a limited number of FA with relatively high accuracy. Therefore, HBME-1 may indeed be useful in the differential diagnosis of FVPTC and FA (accuracy 88%).

We found CITED-1 expression both in benign and malignant thyroid lesions. Although the highest proportion of positive samples was found in PTC (as previously reported (4, 37)) and FVPTC, the high number of positive samples in benign lesions makes CITED-1, in our opinion, a less attractive marker for differential diagnosis as also reported by Scognamiglio et al. (18).

Gal-3 was predominantly expressed in PTC (92%) and to a lesser extent in FTC and FVPTC. Other investigators have used Gal-3 in differentiating FTC from FA in fine needle aspirates (16); however, Gal-3 expression has also been reported in benign thyroid lesions (41, 42). We found a reasonable accuracy (88%) in the differential diagnosis between FA and FVPTC for Gal-3. We also found Gal-3 to be a useful marker in the cluster analysis.

FN-1 was first reported to be overexpressed in PTC (4, 12). In a subsequent study, FN-1 appeared to be a valuable marker for the differentiation of FA and thyroid carcinomas (37). The proportion of FA (40%) positive for FN-1 in our study was higher than that reported by Prasad et al. (37). We found accuracies of 74% for the differentiation between FA and FVPTC and 71% between FA and FTC. Cluster analysis also identified FN-1 as a useful marker.

Although some studies report decreased NIS protein expression in thyroid carcinoma (43), Dohán et al. observed intracellular overexpression of NIS in a large series of human thyroid cancers (30, 31, 40). We confirmed iNIS overexpression in PTC (83% of tissues) and FTC (67% of tissues). As we used the same antibody as Dohan et al. it may well be that the differences with other studies are related to differences in antibody specificity. Furthermore, the differential expression of iNIS between the subtypes of thyroid neoplasms makes it a candidate for differentiating between these lesions. The accuracies of 68% (FA versus FTC) and 77% (PTC versus FVPTC), however, are moderate. iNIS was also identified by cluster analysis as a potential useful

Table 3 Diagnostic value of proteins differentially expressed in follicular thyroid lesions.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Cut-off value</th>
<th>FA</th>
<th>PTC</th>
<th>FTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>FVPTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPAR-γ</td>
<td>0</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBME-1</td>
<td>1</td>
<td>89</td>
<td>63</td>
<td>84</td>
</tr>
<tr>
<td>CK-19</td>
<td>1.5</td>
<td>63</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>iGal-3</td>
<td>2</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nGal-3</td>
<td>1</td>
<td>63</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>iIFN-1</td>
<td>2</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pmFN-1</td>
<td>2</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iNIS</td>
<td>2</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iFN-1</td>
<td>2</td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iNIS</td>
<td>2</td>
<td>65</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The proteins studied are selected from Table 2. The accuracy is a combination of the sensitivity and specificity for the distinction between two lesions. For instance, the accuracy for the distinction of FVPTC and FA by PPAR-γ staining is 74%. FA, follicular adenoma; PTC, papillary thyroid carcinoma; FTC, follicular thyroid carcinoma; FVPTC, follicular variant PTC; Gal-3: galectin-3 (i, intracellular; n, nuclear); NIS, sodium iodide symporter (i, intracellular); FN-1, fibronectin-1 (i, intracellular); CK-19, cytokeratin-19. *For the definition of cut-off values, see text.
Figure 2 Hierarchical cluster analyses using seven antibodies in all thyroid tissues. Intracellular sodium/iodide symporter (iNIS), fibronectin-1 (FN-1), and galectin-3 (Gal-3) were identified as the best predictors of benign or malignant thyroid lesions. The presence of two of these markers (cluster 1) gave a 100% clustering of malignant thyroid lesions, whereas the absence of these proteins (cluster 5) was highly suggestive of benign thyroid lesions. CTL, normal thyroid; HP, hyperplasia; FA, follicular adenoma; FTC, follicular thyroid carcinoma; PTC, papillary thyroid carcinoma; FmixPTC, mixed FTC and PTC; FVPTC, follicular variant PTC.
marker in the discrimination between FA and malignant carcinomas.

PPAR-γ has been found to be downregulated in the experimental models of thyroid carcinoma (33–35). The importance of the downregulation of PPAR-γ is also illustrated in the PPAR-γ/Pax8 rearrangement (44), which was initially observed in a series of FTC. Although the PPAR-γ/Pax8 rearrangement was therefore considered a specific marker for FTC, later studies also reported the rearrangement in benign thyroid lesions (45, 46). We observed decreased PPAR-γ nuclear staining in malignant tumors, whereas in non-malignant lesions, the percentage of positive cells varied from 50 to 100%. Although our results confirmed the decreased expression of PPAR-γ in thyroid carcinoma, the diagnostic accuracy for the differentiation between follicular lesions was limited.

As no marker by itself has a superior diagnostic value, a combination of markers may be more accurate than any single marker. We performed a cluster analysis including all studied tissues and antibodies. To our knowledge, this has not been done before by immunohistochemistry in thyroid lesions. In this study, hierarchical clustering analysis on all valid samples confirmed that thyroid carcinomas, FA, and benign lesions could be categorized with high sensitivity, specificity, and accuracy. Our findings show that a diagnostic immunohistochemical panel comprising Gal-3 and FN-1 was 97% sensitive for all thyroid carcinomas, whereas its specificity was 100%. The diagnostic values of CK-19, CITED-1, and HBME-1 in our series were not sufficient to be included in the panel, which is in line with the results of de Matos et al. (27). However, HBME-1 was found to be a useful marker for the differentiation between FA and FVPTC. Hierarchical clustering was not able to identify FVPTCs as a separate cluster, likely because in common with PTCs in general, they show a high positivity for HBME-1 and CITED-1, and therefore cluster too closely to the main group of PTCs. Prasad et al. also found a limited accuracy for HBME-1, CK-19, and CITED-1 (37). Our study was performed to test the proof of concept of the ROC and hierarchical cluster-based multiple marker staining. This concept has to be validated in the subsequent series and most importantly in cytological preparations. Also, other difficult-to-classify thyroid neoplasms such as minimally invasive follicular carcinomas, follicular variants of papillary carcinoma, as well as FA subclasses should be included in the subsequent studies.

In conclusion, Gal-3, FN-1, and iNIS constitute a useful diagnostic panel in the differential diagnosis of thyroid lesions. The absence of Gal-3, FN-1, and iNIS is highly suggestive of a benign lesion. Moreover, HBME-1 may be useful in the specific differentiation of FVPTC from FA.

Acknowledgements
The research has been supported by a grant from the Netherlands Cancer Foundation KWF.

References
Combined immunostaining for differential diagnosis


Received 29 November 2007
Accepted 5 December 2007