




# Retrospective analysis of cancer-specific gene expression panel for thyroid fine needle aspiration specimens

Yongchun Wang<sup>1</sup> · Brittany A. McKelvey<sup>1</sup> · Zhiyang Liu<sup>1</sup> · Lisa Rooper<sup>3</sup> · Leslie M. Cope<sup>2</sup> · Martha A. Zeiger<sup>5</sup> · Christopher B. Umbricht<sup>1,3,4</sup> 

Received: 31 May 2021 / Accepted: 18 June 2021 / Published online: 24 June 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

**Background** While molecular testing is a promising strategy for preoperative assessment of cytologically indeterminate thyroid nodules, thyroid fine needle aspiration biopsy (FNA) presents unique challenges for molecular assays, including contaminating peripheral blood mononuclear cells (PBMC) and variable numbers of evaluable epithelial thyroid cells. Moreover, the newly recognized entity, noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), has added an additional challenge to the currently available molecular diagnostic platforms. New diagnostic tools are still needed to correctly distinguish benign and malignant thyroid nodules preoperatively.

**Methods** Twenty-two transcript splice variants from 12 genes we previously identified as discriminating benign from malignant thyroid nodules were characterized in 80 frozen thyroid tumors from 8 histological subtypes. Isoforms detectable in PBMC were excluded, and the 5 most discriminating isoforms were further validated by real-time quantitative PCR (qPCR) on intraoperative FNA samples from 59 malignant tumors, 55 benign nodules, and 23 NIFTP samples. The qPCR threshold cycle values for each transcript were normalized to the thyrocyte-specific thyroid peroxidase isoform 1 (TPO1) and z-transformed. Receiver operating characteristic (ROC) analyses of the composite transcript scores were used to evaluate classification of thyroid FNAs by the 5-gene isoform expression panel.

**Results** A molecular signature was developed by combining expression levels of specific isoforms of CDH3, FNDC4, HMGA2, KLK7, and PLAG1. FNAs containing at least 12–36 thyrocytes were sufficient for this assay. The 5-gene composite score achieved an area under the ROC curve (AUC) of 0.86 for distinguishing malignant from benign nodules, with a specificity of 91%, sensitivity of 75%, negative predictive value of 91%, and positive predictive value of 74%.

**Conclusion** Our newly developed 5-gene isoform expression panel distinguishes benign from malignant thyroid tumors and may help distinguish benign from malignant thyroid nodules in the context of the new NIFTP subtype.

**Keywords** Thyroid cancer · Fine needle aspiration · Indeterminate cytology · Molecular test · Thyrocyte-specific gene expression

---

✉ Christopher B. Umbricht  
cumbrich@jhmi.edu

Yongchun Wang  
ywang25@jhmi.edu

Brittany A. McKelvey  
bavin2@alumni.jh.edu

Zhiyang Liu  
liuzhiyangmf020601@sina.com

Lisa Rooper  
rooper@jhmi.edu

Leslie M. Cope  
lcope1@jhmi.edu

Martha A. Zeiger  
martha.zeiger@nih.gov

<sup>1</sup> Department of Surgery, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>2</sup> Department of Oncology Biostatistics, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>3</sup> Department of Pathology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>4</sup> Department of Oncology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>5</sup> Surgical Oncology Program, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

## Abbreviations

FNA	Fine needle aspiration biopsy
PBMC	Peripheral blood mononuclear cells
AN	Adenomatoid nodules
FA	Follicular adenomas
HA	HÜRTHLE cell adenomas
NIFTP	Noninvasive follicular thyroid neoplasm with papillary-like nuclear features
FVPTC	Follicular variant of papillary thyroid carcinoma
PTC	Papillary thyroid carcinoma
FC	Follicular carcinomas
HC	Hürthle cell carcinomas
qPCR	Real-time quantitative PCR
Ct	Threshold cycle
ROC	Receiver operating characteristic
AUC	Area under the ROC curve
PPV	Positive predictive value
NPV	Negative predictive value

## Background

Thyroid cancer incidence is rapidly increasing worldwide. In the United States, its prevalence has nearly quadrupled over the past 15 years, predominantly due to the increased incidence of papillary thyroid carcinoma (PTC) (SEER 2018). Thyroid nodules are detectable by ultrasound in over 50% of the adult population, and fine needle aspiration (FNA) cytopathology is the most accurate means of pre-operative diagnosis. However, up to 30% of the FNA samples result in indeterminate cytology: atypia of undetermined significance/follicular lesion of undetermined significance (Bethesda III), follicular neoplasm/suspicious for a follicular neoplasm (Bethesda IV), and suspicious for malignancy (Bethesda V) (Ali and Cibas 2018), demonstrating inherent limitations of visual microscopic diagnosis. The situation has been further complicated by the recent reclassification in the American Thyroid Association management guidelines for thyroid tumors of a previously considered malignant subgroup of follicular variant of papillary thyroid carcinoma (FVPTC) to the noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), an indolent neoplasm with questionable malignant potential (Haugen et al. 2016). NIFTP cannot be differentiated from invasive FVPTC by cytopathology, however, and requires histopathologic evaluation for diagnosis (Ali and Cibas 2018; Bongiovanni et al. 2019).

In the last decade, molecular testing has emerged as a promising strategy for the preoperative assessment of thyroid nodules to enable clinicians to better tailor surgical interventions and avoid overtreatment of benign disease. ThyroSeq v3 next generation sequencing (CBL Path, Rye Brook, NY) (Nikiforov and Baloch 2019) Afirma gene expression with recently

upgraded gene sequencing classifiers (Veracyte, South San Francisco, CA) (Endo et al. 2019), and ThygenX/ThyraMIR gene mutation and miRNA analysis (Interpace Diagnostics, Parsippany, NJ) (Lupo et al. 2020) are major molecular tests currently in clinical use in the United States. The DNA-based tests, however, are limited by the prevalence of several of the assessed mutations in benign thyroid lesions (Labourier and Fahey 2021; Livhits et al. 2021), and all tests suffer from a lack of specificity after the introduction of the newly defined NIFTP subtype (Al-Qurayshi et al. 2017; Shrestha et al. 2016; Parajuli et al. 2019; Bose et al. 2019). Indeed, both ThyroSeq and Afirma tests currently report NIFTP as “positive” or “suspicious” for malignant disease. In a recent study, the positive predictive value (PPV) of ThyroSeq mutational analysis panels (including 7-gene ThyroSeq and ThyroSeq V2) decreased from 43 to 14% and Afirma gene expression classifier from 30 to 25% (Jug et al. 2018), if NIFTP was classified as non-malignant.

Before the introduction of the NIFTP subtype, our group performed a transcriptome microarray analysis of 125 tumor samples representing the most common epithelial thyroid tumor diagnoses: adenomatoid nodules (AN), follicular adenomas (FA), Hürthle cell adenomas (HA), follicular carcinomas (FC), Hürthle cell carcinomas (HC), FVPTC, and PTC, and identified over 75 transcripts that were differentially expressed between benign and malignant tumors (Prasad et al. 2008). In a follow-up study, we further characterized 14 of these transcripts by a combination of immunohistochemical and quantitative reverse transcription-PCR assays, and identified a candidate 3-gene panel as a potential preoperative diagnostic tool for FNA samples (Prasad et al. 2012).

In the current study, we selected 12 of the 14 genes based on their diagnostic performance in a receiver operating characteristic (ROC) analysis, to determine if newly characterized specific isoforms further improved their diagnostic performance, including cases of the pathologically confirmed NIFTP tumor subtype. We also investigated isoforms of thyroglobulin and thyroid peroxidase (TPO) to identify a thyrocyte-specific load control, since, in contrast to qualitative markers such as mutations, quantitative molecular assessments are complicated by highly variable samples and admixture of non-thyroid derived cells like peripheral blood mononuclear cells (PBMC). Finally, we tested our 5 best performing candidate isoforms directly in intra-operatively obtained FNA samples.

## Materials and methods

### Clinical samples

Under Institutional Review Board approval, thyroid tumor tissue, intraoperative FNA, and blood specimens were

collected from patients undergoing thyroid surgery at Johns Hopkins Hospital. Intraoperative FNAs were collected from the tumor with a 25-gauge needle syringe immediately prior to resection and preserved in 95% Ethanol at  $-20^{\circ}\text{C}$ , or in RNALater at  $-80^{\circ}\text{C}$ . The FNA site was marked intraoperatively to ensure correlation with the final pathological diagnosis. During pathological prosection, an aliquot of tumor tissue was snap frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . Tumor tissue was identified by hematoxylin and eosin staining of frozen tissue sections, and final surgical pathological diagnoses of samples were confirmed by a pathologist. All FVPTCs were re-reviewed to allow reclassification of NIFTPs from FVPTCs where indicated. PBMCs were isolated from patient blood drawn intraoperatively with Ficoll-Paque Plus (GE Healthcare) and stored at  $-80^{\circ}\text{C}$ .

### RNA isolation and reverse transcription

Total RNA from 80 frozen thyroid tumors (15 ANs, 14 FAs, 10 HAs, 5 NIFTPs, 7 FVPTCs, 7 HCs, 7 FCs, and 15 PTCs) and PBMCs from 31 thyroid patients were isolated with Trizol (Invitrogen) and RNeasy Mini Kit (Qiagen) following the manufacturer's instructions. cDNA was synthesized by reverse transcription with 500 ng of total RNA using SuperScript III reverse transcriptase (Invitrogen). Total RNA from single pass FNA samples was isolated using GenElute Single Cell RNA Purification Kit (Sigma-Aldrich) following the manufacturer's instructions. Because of the variable number of thyroid cells in individual FNA samples, the entire total RNA elution volume (11  $\mu\text{l}$ ) was used for cDNA synthesis.

### PCR of thyroid tumor tissue and PBMC samples

Each PCR assay was performed with 5% of total cDNA using Platinum *Taq* DNA polymerase (Invitrogen) following the manufacturer's instructions. Amplified DNA was analyzed by agarose gel electrophoresis (see below) for 22 isoforms from 12 genes, CEACAM6, CDH3, DIRAS3, DPP4, FNDC4, HMGA2, KLK7, MRC2, SFN, c-KIT, PRSS3, and PLAG1. Additionally, 4 isoforms of thyroglobulin and thyroid peroxidase (TPO) were assessed as potential thyrocyte-specific load control on solid thyroid tumor and PBMC samples. Glyceraldehyde-3-phosphate dehydrogenase (GapDH, NM\_002046) served as a total RNA load control.

### Real-time quantitative PCR (qPCR) on FNA samples

Selected target and reference gene isoforms were tested by real-time qPCR on 159 FNA samples from 6 confirmed surgical histological subtypes of thyroid tumors; AN, FA, HA, NIFTP, FVPTC, and PTC. (See supplemental Table S1 for PCR primer sequences.) Gene expression was quantitated in duplicate using 5% of total cDNA in each assay and Power

SYBR Green Master Mix (Applied Biosystems) on a BioRad iQ5 thermal cycler for 40 cycles. The Ct values of the duplicates were averaged. Serial dilutions of frozen tumor RNA were used to establish the threshold of detectability of the selected thyrocyte-specific reference gene in the assay, and an amplification threshold cycle (Ct) value  $> 30$  was chosen to exclude sample from further analysis for lack of sufficient thyrocytes.

### Data analysis

For the initial selection of candidate gene isoforms on frozen tumor tissue samples, semi-quantitative densitometry using BioRad Quantity One image analysis software was used to select the isoforms showing the highest levels of differential expression between cancer and non-cancer samples on agarose gel electrophoresis images. Expression levels of each isoform were assessed by normalizing to the RNA load control gene GapDH and then z-transformed, so all genes could be assessed on the same scale.

For the subsequent qPCR analysis of FNA samples, the relative expression of each target gene was calculated with respect to the reference load control gene TPO1 [ $\Delta\text{Ct} = \text{Ct}(\text{target}) - \text{Ct}(\text{reference})$ ] and then z-transformed for assessment. When a target gene was undetectable after 40 cycles, the sample was assigned the maximum observed  $\Delta\text{Ct}$  value + 5% of the standard deviation of the  $\Delta\text{Ct}$  values observed for that gene across all samples.

The ability of the candidate genes to distinguish between malignant and benign thyroid tumor subtypes was evaluated using ROC analysis. Overall performance was measured as the AUC.

We then applied Bayes Rule to estimate the positive predictive value (PPV) and negative predictive value (NPV) from our observed sensitivity and specificity using the tumor prevalence rates reported by Steward et al. (2019).

The overall workflow of the study is shown in supplemental Figure S1.

## Results

### Candidate gene isoform selection by reverse transcription-PCR using tumor tissues and PBMCs

The 22 selected transcript isoforms were characterized on frozen tumor tissue samples from 15 ANs, 14 FAs, 10 HAs, 5 NIFTPs, 7 FVPTCs, 7 HCs, 7 FCs, and 15 PTCs, using semi-quantitative PCR. Five transcript isoforms from the following genes: CDH3, FNDC4, HMGA2, KLK7, PLAG1 showed the most differential expression among thyroid cancers, NIFTPs, and benign tumors (Table 1, Fig. 1), as determined by image analysis of agarose gel electrophoresis of

the PCR products. Importantly, none of the isoforms were detectable in PBMCs from 31 patients (Figure S2). Among the candidate thyrocyte load-control isoforms of TPO and thyroglobulin tested, TPO1 was the only isoform not detectable in PBMCs (Figure S3) and stably expressed across the well-differentiated thyroid tumor subtypes (Fig. 1a). Therefore, TPO1 was selected as the load control for thyroid cell content of FNA samples.

One recurrent metastatic HC showed markedly reduced TPO1 levels (Fig. 1a). No FNA samples of this case were available for testing, however.

### Validation of the five candidate gene isoforms on an independent cohort of FNA samples

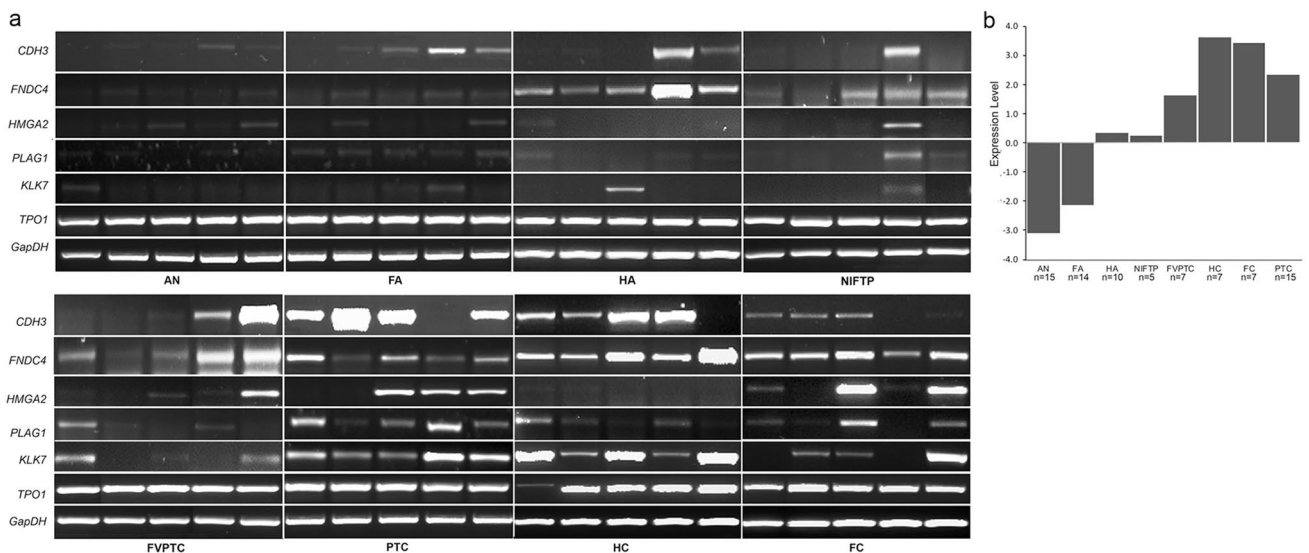
Seven (4.4%) of the 159 FNA samples tested had no detectable TPO1 expression with 40 qPCR cycles and, were therefore excluded from the study. Serial dilutions of thyroid tissue-derived RNA revealed that TPO1 was detectable at

30 qPCR cycles using a minimum of 370 pg of total RNA, which corresponds to the total RNA of approximately 12–36 cells (Nygaard and Hovig 2006), lower than the cytopathological threshold of minimal thyrocyte content in FNA samples (6 clusters of at least 10 follicular epithelial cells on 2 or more slides) (Cibas and Ali 2017). We therefore selected a TPO1 threshold Ct value of 30 to include FNA samples for our study, a criterion met by a total of 137 of 159 (86.2%) FNA samples tested. There was no significant difference in the fail rates across the diagnostic subgroups tested. Table 2 summarizes the patient information for the FNAs used.

The TPO1-normalized  $\Delta$ Ct data were z-transformed to create a z- $\Delta$ Ct score for the expression of each target gene. Figure 2 shows the expression profiles of the selected CDH3, FNDC4, HMGA2, KLK7, PLAG1 isoforms and the composite z- $\Delta$ Ct score in the 6 thyroid tumor subtypes tested (AN, FA, HA, NIFTP, FVPTC, and PTC). Each individual isoform showed higher expression in malignant thyroid tumors, than in benign. When summing the 5 transcripts,

**Table 1** Gene transcript variants selected

Symbol	Reference	Gene name	Isoform
CDH3	NM_001793.5	Cadherin 3	Transcript variant 1
FNDC4	NM_022823.2	Fibronectin type III domain containing 4	Transcript variant 1
HMGA2	NM_003483.4	High mobility group AT-hook 2	Transcript variant 1
KLK7	NM_005046.3	Kallikrein related peptidase 7	Transcript variant 1
PLAG1	NM_002655.2	PLAG1 zinc finger	Transcript variant 1
TPO1	NM_000547.5	Thyroid peroxidase	Transcript variant 1



**Fig. 1** Expression of the 5 gene transcripts and internal controls (TPO1 & GapDH) in solid thyroid tumor frozen samples. **a** Representative gel image showing each of the 5 isoforms and thyrocyte-specific TPO1 in 8 thyroid tumor subtypes. **b** Bar plot of the sum of expression levels of the 5 isoforms (AN,  $n = 15$ ; FA,  $n = 14$ ; HA,

$n = 10$ ; NIFTP,  $n = 5$ ; FVPTC,  $n = 7$ ; HC,  $n = 7$ ; FC,  $n = 7$ ; PTC,  $n = 15$ ). Semi-quantitative densitometry of each isoform was normalized to GapDH and z-transformed so that all genes would be on the same scale

the composite  $z$ - $\Delta$ Ct score exhibited a differential profile among thyroid tumors (Fig. 2f).

An ROC analysis (Fig. 3a) was used to evaluate the ability of our 5-gene isoform expression panel to differentiate benign tumors (AN, FA, HA, and NIFTP,  $n = 78$ ) from malignant tumors (FVPTC and PTC,  $n = 59$ ). Overall performance was measured as the AUC, which was 0.86. Several combinations of sensitivity and specificity, representing points along the ROC curve, are shown in Table 3 as well. One of these combinations is highlighted on the ROC curve and on the strip-plots in Fig. 2. Corresponding to a threshold for the composite expression score ( $z$ - $\Delta$ Ct) of  $-1$ , this value was chosen to maximize sensitivity (75%) while controlling specificity above 90% (actual value = 91%). Forty-four of 59 (74.6%) malignant thyroid tumors had a composite score  $> -1$ , (23/25, 92.0% of PTCs; 21/34, 61.8% of FVPTCs), while only 7 out of 78 (9.0%) benign and NIFTP nodules had a score  $> -1$  (6/23, 26.1% of NIFTPs; 1/14, 7.1% of HAs; 0/16 of FAs; and 0/25, 0% of ANs;  $p < 0.0001$ ).

We used the prevalence rates from a previous large multicenter study (Steward et al. 2019) to apply Bayes Rule to estimate the NPV and PPV from our observed sensitivity and specificity, resulting in an NPV of 91% and a PPV of 74% (Table 3). Variations of this calculation, assuming malignant sample prevalence rates ranging from 20 to 30%, are shown in Supplemental Table S2.

The panel also significantly differentiated the NIFTPs from the malignant PTCs and FVPTCs (26.1% of NIFTPs versus 74.6% of cancers with scores  $> -1$ ,  $p = 0.0002$ ). Further, the comparison between NIFTP versus invasive FVPTC showed a statistically significant separation ( $p < 0.05$ ) (Table 4).

## Discussion

In the past decade, molecular testing has emerged as a promising method to increase the accuracy of the preoperative diagnosis of malignant thyroid tumors. Several molecular diagnostic tests, including RNA based gene expression and multi-panel mutation genotyping analysis are commercially available for clinical use (Steward et al. 2019; Nikiforov 2017). Available tests remain, however, limited by relatively low specificity and PPV (Samulski et al. 2016; Brauner et al. 2015). Furthermore, the recent reclassification of a subgroup of malignant FVPTC to clinically “benign” NIFTP has further complicated the situation, since prior publications assessing the performance of commercially available molecular tests were based upon NIFTP being categorized as malignant. Indeed, most NIFTPs were reported as suspicious/malignant by expression based Afirma or mutation and gene fusion based ThyroSeq in several studies (Jug et al. 2018; Samulski et al. 2016; Sahli et al. 2017; Jiang et al. 2016).

Our group previously carried out a series of studies to identify genetic markers for distinguishing cancer from benign thyroid tumors using genome-wide gene expression arrays (Prasad et al. 2008,2012). In this study, we have developed a molecular test for evaluating preoperative thyroid FNAs taking NIFTP lesions into consideration, by further characterizing isoforms of our previously profiled gene candidates.

Taking advantage of the improved annotation of the human genome over the last decade, we first explored expression of a broad range of splice variants of our 12 gene set in frozen thyroid tumor samples to select candidates which were differentially expressed in benign and malignant

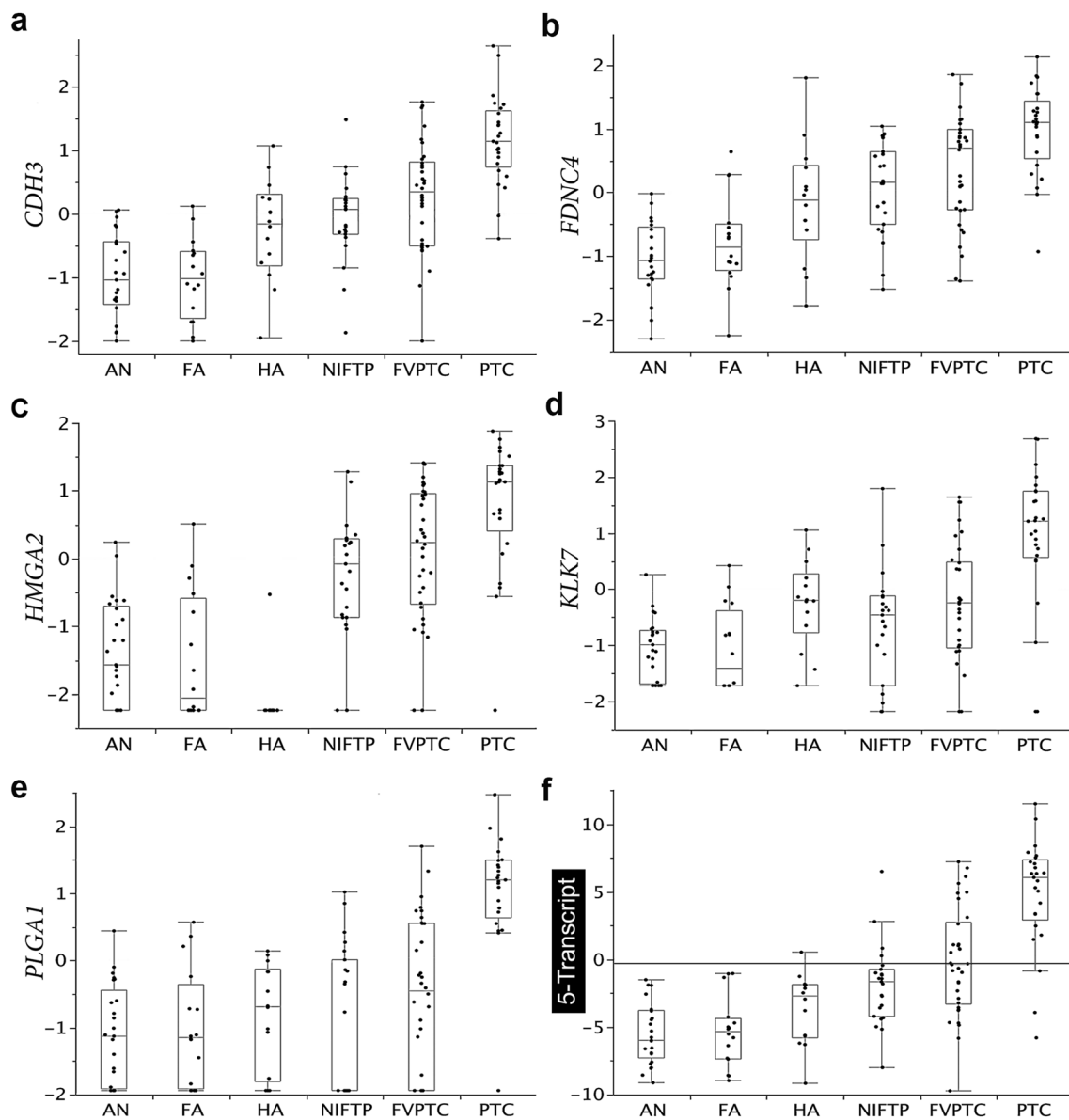
**Table 2** Patient information of the FNA study cohort

	AN	FA	HA	NIFTP	FVPTC	PTC
Sample size, $n$	25	16	14	23	34	25
Sex, M/F	8/17	4/12	3/11	7/16	5/29	9/16
Age, years	48.1 (27–68)	46.1 (18–74)	53.6 (29–80)	50.1 (27–72)	44.4 (19–76)	41.5 (18–59)
Nodule size, cm	3.2 (0.8–7.8)	2.8 (1.0–7.5)	3.1 (1.1–7.0)	2.9 (0.9–8.0)	2.5 (0.6–6.0)	2.4 (0.6–5.0)
Bethesda I, $n$	1					
Bethesda II, $n$	11	3		3	4	
Bethesda III, $n$	5	1		12	7	1
Bethesda IV, $n$	7	10	13	2	8	1
Bethesda V, $n$	1			4	9	5
Bethesda VI, $n$		1		2	6	18
Indeterminate cytology*, $n$		1	1			

AN adenomatoid nodule, FA follicular adenomas, HA Hürthle cell adenoma, NIFTP noninvasive follicular thyroid neoplasm with papillary-like nuclear features, FVPTC follicular variant of papillary thyroid carcinoma, PTC papillary thyroid carcinoma

\*Clinical FNA obtained before current Bethesda classification available





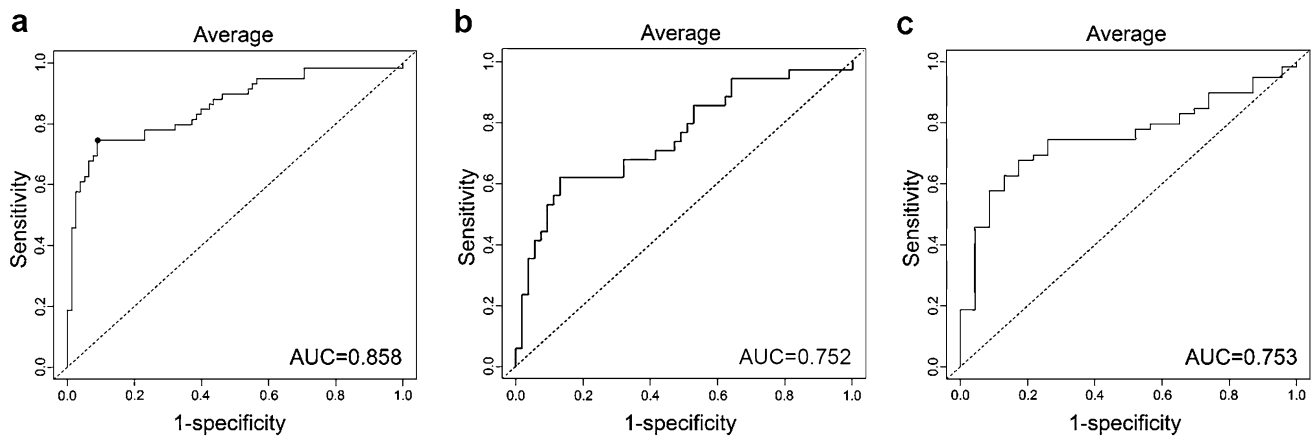
**Fig. 2** The expression profiles of the 5 gene expression isoforms in FNAs. The box plots show the 5th, 25th, 50th, 75th and 95th percentiles of the sample expression  $z$ - $\Delta$ Ct scores in each thyroid tumor subgroup. Panel **f** shows the sum of 5-transcript composite  $z$ - $\Delta$ Ct

score. The expression reference line is the  $z$ - $\Delta$ Ct score of -1, the threshold corresponding to the sensitivity of 75% at the specificity of 91%. AN,  $n=25$ ; FA,  $n=16$ ; HA,  $n=14$ ; NIFTP,  $n=23$ ; FVPTC,  $n=34$ ; PTC,  $n=25$

neoplasms. In this study, the 5 isoforms of CDH3, FNDC4, HMGA2, KLK7, and PLAG1 we identified showed potential in differentiating different thyroid tumor subtypes, and importantly, were chosen because they were not expressed in PBMCs. In FNA samples, quantitative molecular tests must address the contribution from PBMCs. Positive or negative selection using antibody-coated magnetic beads can minimize their contribution, but their use decreases overall assay sensitivity (data not shown). In the absence of selection, the load control reference gene needs to reflect the number of thyrocytes in FNA samples rather than a standard total

RNA content measure. In our study, we tested two isoforms of TPO and two of thyroglobulin. Only TPO1 was found to have constant levels of expression across differentiated thyroid tumor subtypes (Fig. 1a) and importantly, was undetectable by PCR after 40 cycles in any of the 31 PBMC samples obtained perioperatively (Figure S3). We did observe, however, a decrease in TPO1 expression in one large HC metastasis, possibly a consequence of loss of differentiation in this recurrent advanced tumor.

The 5-isoform panel was tested using our intraoperative FNA samples. In this study, only a single pass of needle



**Fig. 3** ROC Curves of the diagnostic power of the 5-transcript panel in FNAs. **a** Benign (AN, FA, HA, and NIFTP,  $n=78$ ) vs. malignant (FVPTC and PTC,  $n=59$ ) thyroid tumors. The dot on the ROC curve is the threshold for the composite expression  $z\text{-}\Delta\text{Ct}$  score of

$-1$ , corresponding to the sensitivity (75%) at specificity 91%. **b** Thyroid follicular neoplasms, malignant FVPTC ( $n=34$ ) vs. FA ( $n=16$ ) and NIFTP ( $n=23$ ). **c** NIFTP ( $n=23$ ) vs. malignant thyroid tumors, FVPTC and PTC ( $n=59$ )

**Table 3** Performance of the 5-transcript panel in benign vs. malignant FNAs

Sensitivity, %	Specificity, %	PPV, %	NPV, %
78	76	53	91
75	81	58	90
75	86	65	91
75	91	74	91
61	96	85	88
19	100	100	78

**Table 4** Performance of the 5-transcript panel in differentiating benign vs. malignant follicular lesions or NIFTP vs. malignant FNAs

NIFTP, HA, and FA vs. FVPTC		NIFTP vs. FVPTC and PTC	
Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %
59	75	69	78
56	81	68	83
53	87	63	87
44	91	58	91
29	96	46	96
9	100	19	100

aspirate from each tumor was used for our analysis. To ensure the reliability of the assay, only samples reaching a TPO1 threshold of detectability of 30 cycles or less were included in this study. Only 7 (4.4%) of FNA samples showed no detectable TPO1, and 137 (86.2%) of 159 FNA samples met our threshold and produced reliable gene expression profiles. This compares favorably with reported 2–20% of cases yielding cytopathologically non-diagnostic results with two to five FNA passes (Cibas and Ali 2009).

In our study, quantitative PCR data generated from 137 qualified FNA samples demonstrated a differential expression among benign and malignant thyroid tumors. The ROC analysis shows our 5-isoform panel has an 86% ability to distinguish benign thyroid tumors (AN, FA, and HA) and NIFTP from malignant tumors (FVPTC and PTC), with a specificity of 91% at a sensitivity of 75%, resulting in an NPV of 91% and a PPV of 74%.

The thyroid follicular-patterned lesions, FA, NIFTP, and FVPTC, contribute the most to the indeterminate cytologies. Molecularly, they all are frequently associated with RAS mutations (Esapa et al. 1999). Currently, histologic evaluation of capsule and vascular invasion is necessary for diagnosis of NIFTP. Thus, an accurate diagnosis of NIFTP is impossible by preoperative cytology or mutation based molecular tests. Nevertheless, NIFTP is an indolent lesion with < 1% risk of recurrence (Hodak et al. 2016), and should not be treated as thyroid cancer, although it may warrant resection as potential premalignant lesion.

The NIFTP classification is new. Currently, the reclassified indolent NIFTP is still considered as a surgical disease by most endocrinologists. Its separation from malignant FVPTC may significantly impact clinical treatment decisions, leading to lesser surgical and other ablative procedures, and potentially simple observation as therapeutic options, which are currently under active investigation. More studies, especially prospective long-term follow-up studies, are needed to evaluate its behavior, progression, and optimal management. Finding tools for accurate preoperative identification of NIFTP will promote the study and management for this lesion. The data presented here show our newly developed 5-isoform panel may reduce cytologically and molecularly indeterminate diagnoses.

Our study also has a number of limitations, foremost the number of available intraoperative FNAs, which also do not exactly replicate the standard preoperative diagnostic FNAs typically obtained percutaneously in a clinic setting. We limited ourselves to epithelial thyroid tumors, and were unable to obtain FNA samples from the infrequent FC and HC cases encountered in the timespan samples were collected for this study. Therefore, although the available FC and HC tissue samples had high scores, suggesting the selected isoform panel may do well diagnostically, the performance of the panel in these cases remains unknown for FNAs.

## Conclusions

In conclusion, we have developed a 5-transcript model combining specific splice variants of HMGA2, PLAG1, KLK7, FNDC4, and CDH3 to better characterize thyroid nodules using the technically challenging but clinically relevant diagnostic FNA samples. Further validation trials will be needed to develop this panel as a diagnostic assay to guide preoperative surgical decision making for thyroid tumors.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00432-021-03706-3>.

**Author contributions** MZ and CU planned and designed the study. YW, BM, and ZL performed and analyzed the experiments. LR was the study pathologist and LC the study bioinformatician. All participated in writing and editing the manuscript.

**Funding** This work was supported by a grant from the TEDCO-Maryland Innovation Initiative #2016-MII-3379, awarded to CBU.

**Availability of Data and Material** The datasets used during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** None of the authors declare any conflicts of interest or competing financial interests.

**Ethical approval** Under Johns Hopkins Institutional Review Board approval, thyroid tumor tissue, intraoperative FNA, and blood specimens were collected from patients undergoing thyroid surgery at Johns Hopkins Hospital.

**Consent to participate** This manuscript does not contain individual patient data.

## References

Ali SZ, Cibas E (2018) The Bethesda system for reporting thyroid cytopathology. In: Ali SZ, Cibas ES (eds) 2nd edn. Springer International Publishing. XV, 236

- Al-Qurayshi Z, Deniwar A, Thethi T, Mallik T, Srivastav S, Murad F et al (2017) Association of malignancy prevalence with test properties and performance of the gene expression classifier in indeterminate thyroid nodules. *JAMA Otolaryngol Head Neck Surg* 143(4):403–408. Available from <https://www.mendeley.com/catalogue/association-malignancy-prevalence-test-properties-performance-gene-expression-classifier-indetermina/>
- Bongiovanni M, Giovannella L, Romanelli F, Trimboli P (2019) Cytological diagnoses associated with noninvasive follicular thyroid neoplasms with papillary-like nuclear features according to the Bethesda system for reporting thyroid cytopathology: a systematic review and meta-analysis. *Thyroid* 29(2):222–228
- Bose S, Sacks W, Walts AE (2019) Update on molecular testing for cytologically indeterminate thyroid nodules. *Adv Anat Pathol* 26(1):114–123
- Brauner E, Holmes BJ, Krane JF, Nishino M, Zurakowski D, Hennessey J V et al (2015) Performance of the afirma gene expression classifier in hürthle cell thyroid nodules differs from other indeterminate thyroid nodules. *Thyroid* 25(7): 789–796. Available from <https://www.ncbi.nlm.nih.gov/pubmed/25962906>
- Cibas ES, Ali SZ (2009) The Bethesda system for reporting thyroid cytopathology. *Am J Clin Pathol* 19(11):1159–1165
- Cibas ES, Ali SZ (2017) The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid* 27(11):1341–1346
- Endo M, Nabhan F, Porter K, Roll K, Shirley LA, Azaryan I et al (2019) Afirma gene sequencing classifier compared with gene expression classifier in indeterminate thyroid nodules. *Thyroid* 29(8):1115–1124
- Esapa CT, Johnson SJ, Kendall-Taylor P, Lennard TWJ, Harris PE (1999) Prevalence of Ras mutations in thyroid neoplasia. *Clin Endocrinol (oxf)* 50(4):529–535
- Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE et al (2016) 2015 American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* 26(1):1–133
- Hodak S, Tuttle RM, Maytal G, Nikiforov YE, Randolph G (2016) Changing the cancer diagnosis: the case of follicular variant of papillary thyroid cancer—Primum non Nocere and NIFTP. *Thyroid* 26(7):869–871
- Jiang XS, Harrison GP, Datto MB (2016) Young investigator challenge: molecular testing in noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Cancer Cytopathol* 124(12):893–900
- Jug RC, Datto MB, Jiang XS (2018) Molecular testing for indeterminate thyroid nodules: performance of the afirma gene expression classifier and thyroseq panel. *Cancer Cytopathol* 126(7):471–480
- Labourier E, Fahey TJ (2021) Preoperative molecular testing in thyroid nodules with Bethesda VI cytology: clinical experience and review of the literature. *Diagn Cytopathol* 49(4):E175–E180
- Livhits MJ, Zhu CY, Kuo EJ, Nguyen DT, Kim J, Tseng C-H et al (2021) Effectiveness of molecular testing techniques for diagnosis of indeterminate thyroid nodules a randomized clinical trial visual abstract supplemental content. *JAMA Oncol* 7(1):70–77. Available from <https://jamanetwork.com/>
- Lupo MA, Walts AE, Sistrunk JW, Giordano TJ, Sadow PM et al (2020) Multiplatform molecular test performance in indeterminate thyroid nodules. *J Am Soc Cytopathol*. <https://doi.org/10.1016/j.jasc.2020.07.127>
- Nikiforov YE (2017) Role of molecular markers in thyroid nodule management: then and now. *Endocr Pract* 23(8):979–988
- Nikiforov YE, Baloch ZW (2019) Clinical validation of the ThyroSeq v3 genomic classifier in thyroid nodules with indeterminate FNA cytology. *Cancer Cytopathol* 127(4):225–230. <https://doi.org/10.1002/cncy.22112>



- Nygaard V, Hovig E (2006) Options available for profiling small samples: a review of sample amplification technology when combined with microarray profiling. *Nucleic Acids Res* 34(3): 996–1014. Available from <https://www.ncbi.nlm.nih.gov/pubmed/16473852>
- Parajuli S, Jug R, Ahmadi S, Jiang XS (2019) Hürthle cell predominance impacts results of Afirma gene expression classifier and ThyroSeq molecular panel performance in indeterminate thyroid nodules. *Diagn Cytopathol* 47(11):1177–1183
- Prasad NB, Somervell H, Tufano RP, Dackiw APB, Marohn MR, Califano JA et al (2008) Identification of genes differentially expressed in benign versus malignant thyroid tumors. *Clin Cancer Res* 14(11):3327–3337
- Prasad NB, Kowalski J, Tsai HL, Talbot K, Somervell H, Kouniavsky G et al (2012) Three-gene molecular diagnostic model for thyroid cancer. *Thyroid* 22(3):275–284
- Sahli ZT, Umbricht CB, Schneider EB, Zeiger MA (2017) Thyroid nodule diagnostic markers in the face of the new NIFTP category: time for a reset? *Thyroid* 27(11):1393–1399
- Samulski TD, LiVolsi VA, Wong LQ, Baloch Z (2016) Usage trends and performance characteristics of a “gene expression classifier” in the management of thyroid nodules: an institutional experience. *Diagn Cytopathol* 44(11):867–873
- SEER (2018) Surveillance epidemiology and end results program. Cancer Stat Facts: thyroid cancer. National cancer institute (<https://seer.cancer.gov/statfacts/html/thyro.html>)
- Shrestha RT, Evasovich MR, Amin K, Radulescu A, Sanghvi TS, Nelson AC et al (2016) Correlation between histological diagnosis and mutational panel testing of thyroid nodules: a two-year institutional experience. *Thyroid* 26(8):1068–1076
- Steward DL, Carty SE, Sippel RS, Yang SP, Sosa JA, Sipos JA et al (2019) Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: a prospective blinded multicenter study. *JAMA Oncol* 5(2):204–212

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.