Role of Gene Expression Profiling in Cytologically Indeterminate Thyroid Nodules

Fulvio Basolo1*, Agnese Priotetti1, Riccardo Giannini1, Hurng-Sheng Wu2, Paolo Vitti3, Paolo Miccoli1

1Department of Surgical, Medical, Molecular Pathology and Critical Area, University of Pisa; 2Honorable Dean Asia Institute Tele-Surgery Show-Chwan Memorial Hospital; 3Department of Experimental and Clinical Medicine, University of Pisa, Pisa, Italy

Abstract

Palpable thyroid nodules are a common clinical entity found in 4% to 7% of the general population. Ultrasonography and fine-needle aspiration (FNA) cytology are worldwide accepted methods to discriminate between benign and malignant lesions. Nevertheless, 10-30% of the biopsied nodules exhibit "indeterminate" cytology. The objective of the new reporting systems is to encourage further investigation of these nodules and to lower the overall incidence of false-negative FNA results. In particular, molecular testing for a panel of somatic mutations is feasible using thyroid FNA material and has been revealed as a powerful adjunct to visual microscopic evaluation, since 60-70% of the thyroid cancers harbor at least one known genetic alteration. A sensitive preoperative test that would allow safe identification of clearly benign nodules with indeterminate cytology raises the prospect of avoiding a purely diagnostic surgery. Here, we reviewed the more recent knowledge of gene expression profiling in cytologically indeterminate thyroid nodules and we reported our experience on nCounter Analysis System from NanoString Technologies.

Keywords: Thyroid, indeterminate cytology, gene expression profiling

Introduction

Palpable thyroid nodules are a common clinical entity found in 4% to 7% of the general population (1-2). Both American and European Thyroid Association have recently published guidelines for their evaluation. Ultrasonography and fine-needle aspiration (FNA) cytology (FNAC) are worldwide accepted methods to discriminate between benign and malignant lesions. For those nodules that required biopsy, FNAC provides sufficient information to classify most nodules as benign (72%; ranging 62-85%), whereas approximately 5% (ranging 1-8%) of nodules are cytologically malignant, most often papillary thyroid carcinoma (PTC) (3-4). Nevertheless, 10-30% of the biopsied nodules exhibit “indeterminate” cytology, including categories III (atypia of undetermined significance, AUS/follicular lesion of undetermined significance, FLUS), IV (folicular neoplasm or suspicious for follicular neoplasm, FN), and V (suspicious for malignancy) according to The Bethesda System for Reporting Thyroid Cytology and TIR3A, TIR3B, and TIR4 from the new Italian Consensus for the Classification and Reporting of Thyroid Cytology (5-6). Current guidelines recommend surgical resection for most of these nodules to permit adequate pathologic evaluation, although repeat biopsy is supported for categories III and TIR3A when the risk of malignancy is felt to be sufficiently low (<10%) (7). However, based on surgical histopathological results, in one metareview only 34% (14-48%) of the cytologically indeterminate nodules were found to be histopathologically malignant, revealing significant rates of invasive, cumbersome and expensive surgery (8).

The objective of the new reporting systems is to encourage further investigation of these nodules and to lower the overall incidence of false-negative FNA results. In particular, molecular testing for a panel of somatic mutations is feasible using thyroid FNA material and has been revealed as a powerful adjunct to visual microscopic evaluation, since 60-70% of thyroid cancers harbor at least one known genetic alteration. Several studies have detected v-raf murine sarcoma viral oncogene homolog B (BRAF), rearranged in transformation/PTC1 (RET/PTC), or rat sarcoma (RAS) mutations in thyroid FNAC samples and have suggested that the detection of these genes may improve the conclusiveness of FNAC diagnoses (9-12). When indeterminate aspirations were analyzed for the presence of BRAF and RAS mutations and for RET/PTC and PAX8-PPARY 1 (peroxisome proliferator-activated receptor gamma 1) gene rearrangements, mutations were found in 16% of cases (13). These genetic markers have high specificity and a high predictive value, and therefore identify which indeterminate nodules are malignant (12). On this basis marker positivity can assume a cost-reducing proposition, particularly due to the reduction in 2-stage thyroidectomy (lobectomy followed by completion thyroidectomy) (14).

Though useful, these markers have limited sensitivity and a limited predictive value and therefore fail to detect more than 30% of cancers (13; 15). This rate is too high to be helpful in making the choice between watchful waiting and diagnostic thyroid surgery. A more sensitive preoperative test that would allow safe identification of clearly benign nodules with indeterminate cytology raises the prospect of avoiding a purely diagnostic surgery.

A further specification of FNAC diagnosis can be achieved by the use of gene expression signatures in addition to the screening for known thyroid cancer related mutations. Their purpose is to have high sensitivity and a high negative predictive value (16).

Microarray analysis

The microarray platform has emerged as a potential preoperative diagnostic test. Applications of this technology have included tumor subclassification, prediction of response to
expression profiling in indeterminate thyroid nodules. JBCM 2015; 4(2):75-80

MicroRNA expression profiling

MicroRNAs (miRNAs) are small (approximately 22 nucleotide lengths), noncoding single-stranded RNAs that constitute a novel class of gene regulators (24). The miRNAs are transcribed from endogenous DNA and processed from primary transcript (pri-miRNA) to a hairpin precursor (pre-miRNA) comprising 2 strands: the leading strand used to produce the mature miRNA and the passenger strand that is believed to be degraded. Mature miRNAs target and bind to transcripts and interfere with their translation into protein or cause messenger RNA (mRNA) degradation (25). miRNAs are closely involved in human cancer pathogenesis by regulating the biological behaviors as cell growth, adhesion, differentiation, and apoptosis.

Investigations of the miRNA expression patterns of FTC and PTC compared to benign thyroid tissues have been done and identified several differentially expressed miRNAs. Pallante et al. have analyzed the genome-wide miRNA expression profile in 30 human FTC samples vs. 10 normal thyroid tissue samples using a microarray containing oligonucleotide probes corresponding to 245 human precursors and mature miRNAs. A subset of miRNAs was found to be overexpressed in FTC samples. In particular, miRNA (miR)-221, -222 and -181b were overexpressed in most of the FTCs analyzed, and also in fine needle aspiration biopsies originating from patients affected by PTC. Blocking miR-221 function by antisense methodology led to a reduced cell growth of a human PTC cell line, while its overexpression led to increased colony formation, indicating a critical role of miR-221 overexpression in thyroid carcinogenesis (27). Furthermore, Nikiforova et al. analyzed 60 FNA samples from histologically confirmed malignant and benign thyroid neoplasms (23 PTCs, nine follicular carcinomas, eight follicular adenomas, four anaplastic cancers, four poorly differentiated carcinomas, two medullary carcinomas, five normal thyroid tissues, and five hyperplastic nodules). They observed that miRNA expression profiles had substantial variability within specific tumor types. They confirmed the overall highest up-regulation of miR-221, miR-222, and miR-146b in thyroid PTCs but showed that these miRNAs are not equally expressed in all tumors of this type and mostly up-regulated in tumors carrying BRAF mutation. Moreover, their data suggested that miR-181b was overexpressed in virtually all types of follicular cell-derived thyroid tumors and also in thyroid hyperplastic nodules. They also identified several additional miRNAs that were highly up-regulated in PTCs. One of them, miR-187, appeared to be the most up-regulated miRNA in tumors harboring RET/PTC rearrangement and RAS mutations but was expressed at significantly lower levels in tumors with BRAF mutation. miR-224, in contrast, was a pan-PTC marker because it was overexpressed, albeit at lower levels, in all PTCs (28).

More recently, Vriens et al. analyzed 104 tissue samples from patients with benign and malignant thyroid lesions as well as 125 indeterminate FNA samples (29). They demonstrated that 4 miRNAs (miR-100, miR-125b, miR-138, and miR-768-3p) were down- and upregulated in PTC, follicular thyroid carcinoma and anaplastic carcinoma. Moreover miR-138 even did so for thyroid FNA samples. The negative predictive value of miR-138 for distinguishing between benign and malignant was only 81%; however, it was 100% for distinguishing the follicular type. Thus, authors suggested to consider close follow-up based on testing miR-138 and forgo a diagnostic thyroidectomy. Interestingly, target gene of miR-138 is a human telomerase reverse transcriptase (hTERT) that has a role in cellular senescence and whose deregulated expression in somatic cells is known to be involved in thyroid neoplasm oncogenesis. Mutations of TERT had been found in aggressive thyroid neoplasms, as poorly differentiated and anaplastic tumors (30).

Afirma gene expression classifier

Genomic classifier above described resulted to be limited in their sensitivity and their usefulness has not been validated in large group patients. The Afirma gene expression classifier (AGEC) is a proprietary diagnostic assay developed by Veracyte Inc (South San Francisco, CA). It relies on a “benign gene expression fingerprint” to identify those indeterminate FNA with a high negative predictive value, similar to the probability of malignancy for an initial “benign” cytologic diagnosis (31). AGEC is essentially a “rule-out” test for thyroid cancer. The assay analyzes the mRNA expression of a panel of 167 genes. Of these 167 genes, 142 are involved in the main classifier (i.e., benign vs. malignant) and the remaining 25 genes filter out rare neoplasms (31). It uses a multidimensional algorithm to identify the signature of a benign thyroid nodule. Because the AGEC is proprietary, the detail of the relative weights given to each of the genes in the panel remains unpublished. Based on this information, it classifies nodules as either “benign” (95% negative predictive value for aspirates
classified as AUS/FLUS according to the Bethesda, and TIR3A according to the Italian Consensus for the Classification and Reporting of Thyroid Cytology, and 94% for aspirates classified as FN and TIR3B, respectively) or “suspicious for malignancy (>50% risk of malignancy).

In 2012, Alexander et al. presented a double-blind, prospective, multicenter validation study on 4812 thyroid FNAs from nodules that measured ≥ 1.0 cm in diameter. An important strength of the study is that it included samples from both community practices as well as those from large academic centers. Five hundred seventy-seven were classified as “indeterminate”, from which only 256 samples ultimately were selected to comprise the study’s cohort of indeterminates (49% AUS/FLUS, 31% FN, and 21% suspicious for malignancy). The overall negative predictive value for AGEC was 93% with a sensitivity of 92%. They found an 8.2% false negative rate based on histologic follow-up that was attributed to insufficient RNA (secondary to insufficient cells) in the FNA samples used for gene expression analysis. Although the overall sensitivity of the Afirma test was very good, the overall specificity at 52% was much lower, with a malignancy rate within the AUS/FLUS group (24%) higher than that reported from both the American and European classification systems (5-15%). Of the 129 AUS/FLUS cases in the study’s cohort, the Afirma test was able to reclassify 55 as “benign”, whereas 74 of the cohort’s “AUS/FLUS” FNAs were classified as “suspicious”; and in histologic follow-up, 46 of the latter were benign (false-positive rate, 62%). That means a large number of unnecessary surgeries. For those FNAs in the “suspicious for a follicular neoplasm/follicular neoplasm” group, the negative predictive value also was at 94%. In contrast, the test did less well for those thyroid FNAs diagnosed as “suspicious for malignancy” (negative predictive value, 85%).

In 2014 Alexander et al. provided a multicenter clinical validation study of the previously reported data about Afirma GEC (32). Their purpose was to analyze the diagnostic performance and use of this assay in the clinical setting, and its impact on clinical decision-making. Three hundred thirty-nine Afirma GEC analyses were performed. 49% were cytologically AUS/FLUS. At the Afirma test, 55% of those cases were “benign”, 40% “suspicious” and 5% “nondiagnostic”. Separately, 47% of nodules were cytologically classified as FN. The Afirma classifier found similar percentage of “benign”, “suspicious” and “nondiagnostic” cases in this cytological group. However, 13 of 339 were suspicious at FNA and Afirma GEC analyses were “benign” in 4 of 13 (31%) cases, and “suspicious” in 9 of 13 (69%) cases.

In total, “cytologically indeterminate/Afirma GEC suspicious” nodules proved cancerous in 44% of patients after surgical removal. According to authors, a benign AGEC result altered clinical care recommendations, as 95% of Afirma GEC “suspicious” nodules were referred for thyroid surgery in comparison to only 2% of Afirma GEC “benign” nodules. It is notable that this assay was only recommended for use in nodules with AUS/FLUS and FN cytology.

Similar results have been discussed in a single institutional experience by Lastra et al. (33). In their study cohort of 132 patients the positive predictive value for malignancy of Afirma GEC in cases diagnosed as “suspicious for follicular neoplasm” was 37%. Moreover, they observed that performing Afirma testing only to those cases diagnosed as AUS/FLUS on repeat FNA would increase the positive predictive value, up to 61%. They concluded that this approach may avoid unnecessary a costly molecular test in a considerable number of cases that are diagnosed as AUS/FLUS on first FNA.

These data are in agreement with those reported by Faquin in his commentary to Alexander’s study (34). He stressed out the importance of FNA repeating in case of indeterminate cytology results before approaching with costly molecular tests. Moreover, he proposed the role of targeted review of AUS/FLUS cases, either by an intralaboratory consensus or by experienced pathologists in difficult AUS/FLUS issues. This application of Afirma GEC to triaged patients would increase its diagnostic potential.

Recently, McIver et al. published an independent study to assess the performance of the Afirma GEC in an academic medical center. A total of 72 samples were sent to Veracyte for GEC analysis. In 12 (17%) of these samples, there was insufficient quantity of mRNA to obtain a result, leaving a total of 60 Afirma results available for analysis. Of these, 16 (27%) were reported as benign, while 44 (73%) were reported as suspicious. Only 15.6% of the nodules that were GEC suspicious, for which final histology was known, have proven malignant. In summary, they demonstrated a lower than expected rate of benign Afirma GEC reports in AUS/FLUS and FN, increasing the number of test needed to avoid one surgery from two to four and raising the questions about the costs of widespread application of this assay.

In addition, they found the positive predictive value of suspicious classifier results to be lower than reported by Alexander (16 vs 38%), so that more than 80% of GEC-suspicious nodules proved to be benign at surgery. This result is consistent with the performance of the classifier, when applied to a group of patients at low risk for malignancy, and it is a reminder that its performance depends critically on the input cytopathology.

In conclusion, as Alexander and other authors pointed out neither cytological nor Afirma GEC results alone should mandate an exact clinical recommendation for all patients. Individualized clinical risk assessment and personalized care recommendations should always be pursued, because clinical symptoms, nodule size and/or sonographic findings can at times be enough to warrant intervention regardless of molecular analysis.

nCounter Analysis System

The nCounter Analysis System from NanoString Technologies offers a simple, cost-effective way to profile hundreds of mRNAs, microRNAs, or DNA targets simultaneously with a high sensitivity and precision. Digital detection of target molecules and high levels of multiplexing eliminate the compromise between data quality and data quantity, producing a gold-standard sensitivity and reproducibility for studies of hundreds of targets. NanoString’s system uses molecular “barcodes” and single-molecule imaging to detect and count hundreds of unique transcripts in a single reaction. Unlike other methods, the protocol does not include any amplification steps that might introduce bias to the results. NanoString offers several key advantages, including sensitivity, reproducibility, technical robustness, and utility for clinical application (35).

Its application has been validated in a consistent group of solid tumors (36–38). Nevertheless, only few studies have been conducted on thyroid neoplasms.

Our group had recently published data on a cohort of 70 patients who presented with either a solitary nodule or a dominant thyroid nodule with a histological follow-up (39). FNA samples were classified as 26 benign, 27 malignant, and 17 indeterminate. The objective of the study was to identify a molecular signature capable of classifying thyroid tumors as benign or malignant in cytologically indeterminate samples using a k-nearest neighbor (K-nn) machine-learning algorithm (40).
The nCounter custom code set used in this study was designed and synthesized by NanoString Technologies. This code set consisted of 82 reporter and capture probe pairs directed against 39 genes (2 pairs for each target) and included 34 genes that are differentially expressed in malignant and benign thyroid nodules, thyroglobulin (TG) to confirm the follicular origin of the cells obtained from FNAC, and 4 housekeeping genes for reference (β-actin (ACTB), β2-microglobulin (B2M), hypoxanthine phosphoribosyltransferase (HPRT), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH)).

The raw data normalization was performed in two steps. The first normalization was based on the positive-spike controls, and the second was based on the housekeeping gene expression data. Indeterminate FNAC samples were then filtered using a TG expression level cutoff. Consistently, nearly all of the indeterminate samples (16 of 17) expressed TG levels above the cutoff value, confirming the follicular origin of the cells obtained from FNAC. Consequently, the one sample with a TG expression level under the cutoff value was excluded. After data processing and filtering, 66 of our initial 70 thyroid FNAC samples (25 benign samples, 25 malignant samples, and 16 indeterminate samples) were selected for investigation of the correlation between gene expression profiles and the true benign or malignant status of indeterminate FNAC samples by cluster analyses. Among the 24 genes, 19 exhibited overexpression in PTCs, and 5 were down-regulated.

In the second step, we analyzed the expression profiles of 16 indeterminate FNAC using the same code set. Specifically, K-nn classified 12 FNAC samples as benign, and only 1 sample was classified as malignant. Three of the 16 samples were unclassified according to the molecular analysis, because the gene expression levels fell between the characteristic levels of the benign and malignant lesion groups. The K-nn algorithm classified 13 of the 16 tested samples; 12 of the 13 were considered benign, and 1 was considered malignant. Ten of the 12 molecularly benign samples were identified as adenomas based on histology, whereas the remaining 2 samples were identified as PTCs. The misclassification of the 2 molecularly benign PTCs may have been because of an overall low quality of sample data. Nevertheless, an underestimation of gene expression could be caused by the presence of several different cell types. From a biologic point of view, our results suggest that a group of thyroid lesions that were considered indeterminate on cytologic analysis has a gene expression profile that is intermediate between the profiles of malignant and benign tumors.

Interestingly, our data are in agreement with those presented by Lubitzi et al. (19). As above mentioned, their cluster analysis of indeterminate FNA by microarray tests showed three distinct groups: one clearly benign, one clearly malignant, and one indeterminate grouping. All five indeterminate samples were deemed so on preoperative FNA. Histological review showed partial features of PTC, hinting that at least some cases in this group might truly represent a borderline lesion between follicular adenoma and follicular variant of PTC.

The issue of clinical and pathological meaning of encapsulated FVPTCs is a topical one. Taken together, the morphologic and molecular features suggest that these lesions need further study to define the sources of their malignant potential and, consequently, to determine the appropriate therapy for patients with PTC.

Conflicts of Interest: None

References


indeterminate thyroid nodules. Cancer Cytopathol 2015; [Epub ahead of print]

40. Berrar D, Bradbury I, Dubitzky W. Instance-based concept learning from multiclass DNA microarray data. BMC Bioinformatics 2006; 7:73.