1

Identification of two distinct molecular subtypes of noninvasive follicular neoplasm with papillary-like nuclear features (NIFTP) by digital RNA counting

Riccardo Giannini[1*] PhD (riccardo.giannini@dc.unipi.it), Clara Ugolini[2*] MD (clara.ugolini@gmail.com), Anello Marcello Poma[1] PhD (marcellopoma@gmail.com), Maria Urpì[1] BSc (maria_23uc@hotmail.com), Cristina Niccoli[2] BSc (cristinaniccoli88@gmail.com), Rossella Elisei[3] MD (rossella.elisei@med.unipi.it), Massimo Chiarugi[1] MD (massimo.chiarugi@med.unipi.it), Paolo Vitti[3] MD (paolo.vitti@med.unipi.it), Paolo Miccoli[1] MD (paolo.miccoli@med.unipi.it) and Fulvio Basolo[1,2] MD (fulvio.basolo@med.unipi.it).

[1]Department of Surgical, Medical, Molecular Pathology and Critical Area, University of Pisa, Pisa, Italy;

[2] Department of Laboratory Medicine, Anatomic Pathology, AOUP, Pisa, Italy;

[3]Department of Experimental and Clinical Medicine, University of Pisa, Pisa, Italy.

*These authors contributed equally to this work.

Running title: Gene expression profiling of NIFTPs.

2

Abstract

Background. The follicular variant (FV) of papillary thyroid cancer (PTC) is one of the most common variants of PTC. Clinically, non-infiltrative FVPTC is considered a low-risk variant of PTC, and the noninvasive encapsulated forms of FVPTC represent a group of thyroid tumors with a particularly good prognosis. Consequently, these neoplasms have been very recently reclassified as "non-invasive follicular neoplasms with papillary-like nuclear features (NIFTP)".

From a molecular standpoint, NIFTP appears to be similar to follicular neoplasms; however, only limited data are currently available regarding their gene expression profile.

Methods. The aim of this study was to identify specific molecular signatures of 26 NIFTPs compared to those of 19 follicular adenomas and 18 infiltrative FVPTCs (IFVPTCs). A nanoString custom assay was used to perform mRNA expression analysis. All cases were also genotyped for *BRAF, N-, H- and K-RAS* mutations. Samples were grouped on the basis of gene expression profiles by Pearson's correlation and non-negative matrix factorization (NMF) clustering analysis. Finally, the uncorrelated shrunken centroid (USC) machine-learning algorithm was used to classify the samples.

Results. The results revealed distinct expression profiles of follicular adenomas (FAs) and infiltrative FVPTCs. NIFTP samples can exhibit different expression profiles, more similar to FAs (FA-like) or to IFVPTCs (IFVPTC-like), and these different expression profiles largely depend on the presence of different mutations (RAS or BRAF).

Conclusion. In conclusion, although further validation of the model is required by using a larger group of prospective cases, these data reinforce the hypothesis that IFVPTC-like NIFTPs might represent precursors of IFVPTC.

3

Introduction

The diagnosis of thyroid cancer has shifted over time, changing from a relatively simple classification during the 1960s and 1970s (1) to highly complicated subclassifications more recently (2). The follicular variant (FV) is one of the most common variants of papillary thyroid carcinoma (PTC), occurring in approximately 20-40% (3,4) of PTC patients, with an incidence of FV increasing exponentially over time (5, 6). The most reliable morphological criteria for the diagnosis of this malignancy are papillary carcinoma nuclear features and tumor capsule invasion or infiltration. Clinically, FVPTC has been suggested to be more similar to minimally invasive follicular thyroid cancer, a low-risk lesion, than to conventional PTC (7-11). Moreover, the recent publication of an integrated genomic characterization (The Cancer Genome Atlas-TGCA) of papillary thyroid carcinoma assists pathologists in identifying molecular similarities and differences between the various subtypes of thyroid malignancy (1).

The encapsulated forms of FVPTC (EFVPTC) represent a group of thyroid tumors with an overall good prognosis (12) in terms of an indolent disease course with an extremely low recurrence rate, likely to be less than 1% within the first 15 years. Therefore, EFVPTC with encapsulation or clear demarcation, follicular growth pattern with less than 1% of papillae, no psammoma bodies, less than 30% of solid/trabecular/insular growth pattern and nuclear score 2-3, absence of vascular or capsular invasion, absence of tumor necrosis and low mitotic activity have been recently reclassified as "non-invasive follicular neoplasm with papillary-like nuclear features (NIFTP)" (13).

The objective of the present study was to identify a specific NIFTP molecular signature distinct from that of follicular adenomas and infiltrative FVPTCs by using the nanoString mRNA expression analysis and the uncorrelated shrunken centroid (USC) machine-learning algorithm.

Materials and Methods

Study group

The study included 62 patients with diagnoses of follicular adenomas (FA) and follicular variant of papillary carcinomas who underwent total/near-total thyroidectomy at the Department of Surgical, Medical, Molecular Pathology and Critical Area of the University of

4

Pisa, Italy, between 2013 and 2015. Hematoxylin & eosin-stained sections of neoplasms obtained from the archives of the University of Pisa section of Pathology were re-evaluated independently by two pathologists (C.U., F.B.). For all cases, the eventual capsular and/or vascular invasiveness and the infiltration of thyroid and extra-thyroidal tissues were evaluated. A diagnostic concordance rate of 98% was achieved between the two investigators. Discordant cases were rejected. The tumors were diagnosed and classified according to WHO 2004 histopathological criteria (2).

To avoid contamination with other cellular types, cases with any grade of thyroiditis were excluded, and only neoplasms with approximately 50% of neoplastic cellularity were considered for examination. This retrospective study was conducted anonymously and in compliance with the principles of the Helsinki Declaration of 1975. Informed consent for the molecular analysis was obtained one day before surgery together with the surgical consent.

NIFTP selection criteria

Thirty cases already diagnosed as EFVPTCs were selected. At time of diagnosis, all these cases were "in toto" sampled. For the aim of this study all paraffin-embedded blocks were re-cut and independently re-evaluated by the same two pathologists (C.U. and F.B.) and re-classified as NIFTPs. Diagnoses were made according to the morphological criteria reported by Nikiforov et al. (13), which include encapsulation or clear demarcation, follicular growth pattern with less than 1% of papillae, no psammoma bodies, less than 30% of solid/trabecular/insular growth pattern and nuclear score 2-3, absence of vascular or capsular invasion, absence of tumor necrosis and mitotic activity less than 3 per 10 high power fields. Among the 30 selected cases, 4 did not match the above criteria and were therefore excluded.

5

RNA purification

Total RNA was purified from 18 FAs, 26 NIFTPs and 18 infiltrative FVPTCs (IFVPTCs). After a standard deparaffinization procedure was performed, 2-3 formalin-fixed, paraffin-embedded (FFPE) tissue sections (5 µm in thickness) were submitted to manual microdissection. Then, RNA was extracted and purified using a Qiagen RNeasy FFPE kit (Qiagen, Hilden, Germany), according to the manufacturer's recommendations. RNA concentrations were measured using an Xpose instrument (Trinean, Gentbrugge, Belgium).

NanoString nCounter assay

The nCounter custom code set used in this study was designed and synthesized by nanoString Technologies (nanoString Technologies, Seattle, Washington, USA). A total of 75 genes were included in the code set based on their observed differential expression between both FAs vs FVPTCs/PTCs and FAs vs FTCs (Table 1) (14-25). Four housekeeping genes were also included in the code set for normalization purposes (*beta actin [ACTB], b2-microglobulin [B2M], hypoxanthine phosphoribosyl transferase* [*HPRT*], and *glyceraldehyde-3-phosphate dehydrogenase* [*GAPDH*]). Total RNA (125 ng) from FFPE tissue was analyzed using the nanoString nCounter Analysis System (nanoString Technologies, Seattle, Washington, USA). All procedures related to mRNA quantification, including sample preparation, hybridization, detection, and scanning, were performed following the manufacturer's instructions. Gene expression profiling using nanoString was performed on RNA extracted from the 62 thyroid neoplasms. Raw data normalization was performed in 2 steps, as previously reported (26).

DNA extraction and BRAF and RAS gene family mutational analyses

Two unstained 10-µm paraffin sections from each sample were first deparaffinized with xylene prior to DNA purification using a DNeasy FFPE kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions.

DNA was eluted in 40 µL TE buffer and then immediately processed for polymerase chain reaction (PCR) amplification. Detection of mutations in *NRAS*, *HRAS* and *KRAS* exons 2-3 and *BRAF* exon 15 was performed according to standard procedures (28-30) using direct

6

DNA sequencing on a 3130 Genetic Analyzer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

Unsupervised sample clustering by Pearson's correlation and non-negative matrix factorization (NMF)

All the gene expression profiles were used as input for both Pearson and NMF methods.

To group the samples by related gene expression profiles, cluster analyses were performed using Pearson's correlation by the average linkage method between samples and genes on the normalized data (26).

NMF was employed to identify stable sample clusters. The NMF parameters were as follows: rank, k = 2 to k = 6 clusters; number of runs to build consensus matrix = 15; error function = divergence; and maximum interactions = 1000. The optimal clustering result was determined using the higher cophenetic correlation value between the clusters.

Sample classification by uncorrelated shrunken centroids

The samples were classified by expression profiling from the normalized counts for each analyzed gene using the uncorrelated shrunken centroid (USC) algorithm (27). This algorithm includes two steps: the training step and the test step. The aim of this algorithm is to build a "classifier" using the given training set and to use the classifier to predict the classes of the *test set*.

The first step, or training set, cross-validated the capability of the model to correctly classify the samples regarding the fixed reference classes of samples over a range of parameters (shrinkage threshold Δ and correlation threshold ρ).

The second step involved the use of the classifier to predict the class to which the unknown samples belonged or to identify the most similar class based on phenotypic characteristics.

In our study, the datasets of the FA and IFVPTC samples were considered as the two reference classes in the training set. The *test set* included the NIFTP samples, for which the classes were assumed to be unknown to the algorithm. For computing purposes, we fixed a maximum Δ threshold of 10 and a minimum ρ threshold of 0.5 (correlation step 0.1), and

7

we performed a 10-fold cross validation. The cross validation was run 15 times to obtain a

reliable quantification of errors. Hence, we tested the chosen training sets on the *test set*.

8

Results

Study group

A total of 62 cases were eligible for study inclusion: 18 adenomas, 18 IFVPTCs (Figure 1A) and 26 NIFTP (3) (Figure 1B). All cases exhibited a complete follicular growth pattern and a total absence of thyroiditis. Each tumor had more than 60% neoplastic cells.

Sample genotyping

Overall, mutations were identified in 13 (50.0%) of the 26 NIFTPs and 10 (62.5%) of the 16 IFVPTCs (two IFVPTCs could not be analyzed).

Specifically, among the NIFTPs, the 13 mutated samples included 1 *BRAF*(V600E), 3 *BRAF* (K601E), 6 *NRAS* (Q61R) and 3 *HRAS* (Q61R) mutations. Finally, among the IFVPTCs, the 10 mutated samples had 5 *BRAF* (V600E), 4 *NRAS* (Q61R) and 1 *HRAS* (Q61R) mutations.

Data normalization and sample clustering

None of the samples were excluded after data normalization. Figure 2 shows the sample clusters obtained using the entire code set without data reduction. NMF analysis showed a higher cophenetic correlation (0.9844) for rank 2 than for the two-cluster model, indicating the former to be more stable (data not shown). Specifically, by Pearson clustering, the two sample clusters showed correlations of r=0.23 for the left-sided cluster (C1) and r=0.30 for the right-sided cluster (C2). The C1 cluster consisted of 28 samples: 15 FAs, 11 NIFTPs and 2 IFVPTCs. The C2 cluster consisted of a total of 34 samples: 16 IFVPTCs, 15 NIFTPs and 3 FAs. Within the C1 cluster, one NIFTP sample harbored an HRAS(Q61R) mutation. In contrast, in the C2 cluster, 22 of 34 (64.7%) samples harbored mutations. Specifically, 12 out of 15 (80%) NIFTPs had 2 *HRAS* (Q61R), 6 *NRAS* (Q61R), 1 *BRAF* (V600E), and 3 *BRAF* (K601E) mutation, whereas 10 of 14 (71.4%) IFVPTCs had 1 *HRAS* (Q61R), 4 *NRAS* (Q61R) and 5 *BRAF* (V600E) mutations.

9

USC training set

The training set cross-validation generates a list of results that report the accuracy (i.e., the number of errors) of the classification and the average number of genes used for the classification together with the shrinkage threshold (Δ) and the correlation threshold (ρ). Briefly, we selected the 27-gene classifier because it generated the lowest number of average classification errors. Specifically, the 27-gene classifier produced 0.7 average errors (Δ=1, ρ=1). The following is an alphabetical list of the 27 gene symbols pertaining to the selected classifier: *CITED1, CKBB, COL9A3, CSGALNACT1, DGKI, DIO1, DIO2, FN1, IPCEF1, LRP4, LRRK2, MET, NPC2, ODZ1, PROS1, QPCT, RXRG, SCEL, SDC4, SERPINA1, SLC26A4, SLC5A8, TFCP2L1, TFF3, TG, TIMP1, TPO*.

USC test set

The NIFTP samples were classified using the previously described classifier. Table 2 includes a summary of the classification results, indicating the class assigned, the discriminant score for the assignment (e.g., sample x is assigned to class k with the minimum discriminant score), and the sample genotypes. In brief, the 27-gene model classified the 26 NIFTPs as 13 IFVPTCs and 13 FAs. With regard to genotype, 11 out of 13 (84.6%) harbored mutations in the IFVPTC group and 2 out of 13 (15%) had mutations at the four analyzed loci of the FA group.

10

Discussion

Recently, morphological and molecular studies have redefined the diagnostic, clinical and surgical approaches to thyroid tumors. In particular, EFVPTCs have been extensively evaluated and, considering the criteria established by Nikiforov et al., the non-invasive form of EFVPTC has been re-classified as NIFTP (13). From a molecular standpoint, follicular pattern thyroid neoplasms are associated with RAS gene mutations (10, 31, 32). Furthermore, Yoo et al. reported the association between encapsulated forms of FVPTC and RAS gene family alterations (33). Nikiforov et al. demonstrated that the NIFPTs analyzed in their study harbored clonal molecular alterations (78% of the cases) with RAS mutations being the most common, thus proposing RAS-mutated NIFTPs as the putative precursors of invasive malignancies (13).

In the past decade, gene expression profiles of different thyroid lesions have been analyzed with particular attention to the differential diagnosis between benign and malignant lesions.

In 2003, Barden *et al.* analyzed the gene expression signature of 10 follicular adenomas and of 7 follicular thyroid cancers, showing that the expression profiles of 105 genes were significantly different between benign versus malignant lesions (22). More recently, Borup *et al.,* in 2010, analyzed global transcriptome signatures and found a strong correlation between down-regulation of genes involved in growth arrest, apoptosis and follicular carcinoma. Based on these results, a gene expression panel that represents a strong genetic signature in the differential diagnosis of follicular adenoma or carcinoma has been proposed (14). In 2004, Mazzanti *et al.* built two panels of 10 and 6 genes able to discriminate benign from malignant thyroid tumors, although they were unable to detect significant differences among different malignant tumor subtypes (PTC and FVPTC) (17).

Interestingly, a meta-analysis by Griffith *et al.* identified 12 genes *(MET, TFF3, SERPINA1, TIMP1, FN1, TPO, TGFA, QPCT, CRABP1, FCGBP, EPS8, PROS1)* that may form a gene panel for the differential diagnosis of thyroid tumors (20). In 2011, Vierlinger *et al.* conducted a meta-analysis of 4 microarray datasets on PTCs and nodular goiters and identified a single gene (*SERPINA1*) as a potent mRNA marker for PTC diagnosis with 99% accuracy (19).

More recently, a comparison of microarray expression profiles between FVPTCs and follicular adenomas was reported in 2015 by Schluten *et al*. (16). These authors analyzed

11

whole-transcript arrays of 6 FVPTCs, 7 FAs and 7 normal thyroid tissues, identifying 55 genes differentially expressed (40 up-regulated and 15 down-regulated) in FVPTCs versus FAs. Among the most significantly "deregulated" genes, 8 (*GABBR2, NRCAM, ECM1, HS6ST2, RXRG, IPCEF1, GPR155, PCP4, GRIP1*) appear to be the most promising (16).

In addition, Yoo *et al.* reported a study including FAs, minimally invasive FCs and PTCs (classical and follicular variant) in which gene expression analysis revealed three molecular subtypes, regardless of their histological types (Non-*BRAF*-Non-RAS, *BRAF*-like and RAS-like). In the same study, the authors showed that the transcriptome of miFTCs or encapsulated FVPTCs was indistinguishable from that of FAs (33).

To date, few data are available regarding the gene expression profile of NIFTPs.

Wong *et al.* reported an evaluation of 63 cases of Afirma GEC-suspicious thyroid nodules, previously diagnosed by fine-needle aspiration (FNA) (34/63) as AUS/FLUS or SFN (29/63) (34). Interestingly, at histopathological follow-up, 16 cases were classified as FVPTCs and 14 (88%) of them were eventually classified as NIFTPs. According to these results, the authors suggested that Afirma-suspicious cases (with a preoperative FNA diagnosis of AUS/FLUS or SFN) should be treated only by diagnostic lobectomy (34).

In the present study, we investigated the expression profile of NIFTPs in comparison to those of adenomas and infiltrative follicular variants of PTC. The aim was to describe the expression signature of this new class of neoplasm by defining a list of genes that may be able to characterize it from a molecular standpoint. In particular, we applied customized NanoString expression analysis, which does not require retro-transcription or amplification.

Our results revealed distinct expression profiles for FAs and IFVPTCs that resulted in the separation of cases into two clusters (C1 and C2), with only a few exceptions. The NIFTP samples were equally subdivided within the two above-mentioned clusters because half of the samples exhibited expression profiles similar to those of FAs and the other half had profiles comparable to those of IFVPTCs. Notably, the samples within the C1 cluster did not harbor mutations at the four studied loci except for one *HRAS* (Q61R)-mutated NIFTP. In contrast, the majority of the samples within the C2 cluster, including the NIFTP samples, harbored a *BRAF* (K601E), *HRAS* (Q61R) or *NRAS* (Q61R) mutation. In general, this finding is in accordance with the lower oncogenic potential and risk of malignancy of these

12

mutations (35) than that of *BRAF*(V600E), which has been associated with infiltrative and more aggressive tumors (36, 37).

Surprisingly, in our series, we also found an NIFTP with a *BRAF* (V600E) mutation. For this reason, the case was re-studied by two additional trained pathologists in an independent manner. The case was sampled *in toto*. The entire lesional capsule was re-evaluated but appeared intact. New sections from paraffin blocks were prepared, but the tumor capsule was not infiltrated (Figure 3). No clear papillae nor wide solid areas were found. This unusual observation could be explained by taking into account two different considerations: 1) in rare cases NIFTP could harbor a *BRAF* (V600E) mutation as a pre-infiltrative lesion; 2) there is an important technical limitation for pathologists with regard to the identification of capsular and/or vascular invasion, as it is essentially impossible to completely evaluate the tumor capsule (1,38).

Once the genetic and mRNA heterogeneity of NIFTPs was apparent, we tried to further validate the results with a machine-learning classification approach using the USC algorithm (27). This algorithm allows the identification of small subgroups of genes and is capable of discriminating between classes without making any preliminary assumptions from the data. In this manner, we chose a 27-gene classifier because of its good performance on the training set of samples. Next, we tested the NIFTP (test) set and the 27-gene classifier, which assigned 50% of NIFTPs as IFVPTCs and 50% of NIFTPs as FAs. Therefore, 13 NIFTPs were classified as IFVPTCs by a molecular/statistical tool. Furthermore, the majority of the IFVPTC-assigned samples (84.64%) were mutation-positive.

As reported in the study of Nikiforov *et al.* (13) in which NIFTPs with mutations in *BRAF*, *RAS* or other genes were considered precursors of infiltrative lesions, our data support that mutated NIFTPs also exhibit a likely malignant expression profile (NIFTP-LM). These data reinforce the hypothesis that they might represent a precursor of IFVPTC. In contrast, wild-type NIFTP with an adenoma-like gene expression profile (NIFTP-LA) might represent the so-called "benign counterpart" of these neoplasms (Figure 4).

From a pragmatic standpoint, these results could be prospectively applied in clinical practice, with particular regard to pre-surgical fine-needle aspiration biopsy (FNAB). In fact, it is widely accepted that NIFTPs cannot be identified with certainty by fine-needle

13

aspiration cytology because their nuclear features are quite similar to those of <u>infiltrative</u> follicular variant papillary carcinomas (39).

In conclusion, we suggest that our classifier, in addition to a genotyping analysis (34), could be used for nodules with indeterminate diagnoses by cytology to obtain additional data that might be useful for the diagnostic and prognostic characterization of these tumors. In fact, because of the benign clinical behavior of NIFTPs, preoperative recognition of these neoplasms would have important implications for the therapeutic and surgical strategies employed to avoid patient overtreatment (39, 40).

14

Disclosure statement: No competing financial interests exist.

Corresponding author: Fulvio Basolo MD, Ph.D.

Department of Surgical, Medical, Molecular Pathology and Critical Area, University of Pisa

Via Roma 57, 56126 Pisa, Italy.

Phone: +39 050 993720; Fax: +39 050 992481; email: fulvio.basolo@med.unipi.it

15

References

1. Asa SL, Giordano TJ, LiVolsi VA 2015 Implications of the TCGA genomic characterization of papillary thyroid carcinoma for thyroid pathology: does follicular variant papillary thyroid carcinoma exist? Thyroid 25:1–2.

2. Asa SL, Giordano TJ, LiVolsi VA 2016 Pathology and Genetics of Tumors of Endocrine Organs, 3rd ed. IARC Press, Lyon.

3. Shi X, Liu R, Basolo F, Giannini R, Shen X, Teng D, Guan H, Shan Z, Teng W, Musholt TJ, Al-Kuraya K, Fugazzola L, Colombo C, Kebebew E, Jarzab B, Czarniecka A, Bendlova B, Sykorova V, Sobrinho-Simões M, Soares P, Shong YK, Kim TY, Cheng S, Asa SL, Viola D, Elisei R, Yip L, Mian C, Vianello F, Wang Y, Zhao S, Oler G, Cerutti JM, Puxeddu E, Qu S, Wei Q, Xu H, O'Neill CJ, Sywak MS, Clifton-Bligh R, Lam AK, Riesco-Eizaguirre G, Santisteban P, Yu H, Tallini G, Holt EH, Vasko V, Xing M 2016 Differential clinicopathological risk and prognosis of major papillary thyroid cancer variants. J Clin Endocrinol Metab 101:264–274.

4. Jung CK, Little MP, Lubin JH, Brenner AV, Wells SA, Sigurdson AJ, Nikiforov YE 2014 The increase in thyroid cancer incidence during the last four decades is accompanied by a high frequency of BRAF mutations and a sharp increase in RAS mutations. J Clin Endocrinol Metab 99:E276–E285.

5. Vaccarella S, Franceschi S, Bray F, Wild CP, Plummer M, Dal Maso L 2016 Worldwide thyroid-cancer epidemic? The increasing impact of overdiagnosis. N Engl J Med 375:614–617.

6. Tallini G, Tuttle RM, Ghossein RA 2017 The History of the Follicular Variant of Papillary Thyroid Carcinoma. J Clin Endocrinol Metab 102:15-22.

7. Widder S, Guggisberg K, Khalil M, Pasieka JL 2008 A pathologic re-review of follicular thyroid neoplasms: the impact of changing the threshold for the diagnosis of the follicular variant of papillary thyroid carcinoma. Surgery 144:80–85.

8. Piana S, Frasoldati A, Di Felice E, Gardini G, Tallini G, Rosai J 2010 Encapsulated well-differentiated follicular-patterned thyroid carcinomas do not play a

16

significant role in the fatality rates from thyroid carcinoma. Am J Surg Pathol
34:868–872.

9. Vivero M, Kraft S, Barletta JA 2013 Risk stratification of follicular variant of
papillary thyroid carcinoma. Thyroid 23:273–279.

10. Rivera M, Ricarte-Filho J, Knauf J, Shaha A, Tuttle M, Fagin JA, Ghossein RA 2010
Molecular genotyping of papillary thyroid carcinoma follicular variant according
to its histological subtypes (encapsulated vs infiltrative) reveals distinct BRAF
and RAS mutation patterns. Mod Pathol 23:1191–1200.

11. Kakudo K, Bai Y, Liu Z, Ozaki T 2012 Encapsulated papillary thyroid carcinoma,
follicular variant: a misnomer. Pathol Int 62:155–160.

12. Baloch ZW, Shafique K, Flannagan M, Livolsi VA 2010 Encapsulated classic and
follicular variants of papillary thyroid carcinoma: comparative clinicopathologic
study. Endocr Pract 16:952–959.

13. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LDR,
Barletta JA, Wenig BM, Al Ghuzlan A, Kakudo K, Giordano TJ, Alves VA,
Khanafshar E, Asa SL, El-Naggar AK, Gooding WE, Hodak SP, Lloyd RV, Maytal G,
Mete O, Nikiforova MN, Nosé V, Papotti M, Poller DN, Sadow PM, Tischler AS,
Tuttle RM, Wall KB, LiVolsi VA, Randolph GW, Ghossein RA 2016 Nomenclature
revision for encapsulated follicular variant of papillary thyroid carcinoma: A
paradigm shift to reduce overtreatment of indolent tumors. JAMA Oncol
2:1023-1029.

14. Borup R, Rossing M, Henao R, Yamamoto Y, Krogdahl A, Godballe C, Winther O,
Kiss K, Christensen L, Høgdall E, Bennedbaek F, Nielsen FC 2010 Molecular
signatures of thyroid follicular neoplasia. Endocr Relat Cancer 17:691–708.

15. Finley DJ, Zhu B, Barden CB, Fahey TJ 2004 Discrimination of benign and
malignant thyroid nodules by molecular profiling. Ann Surg 240:425–36;
discussion 436.

16. Schulten HJ, Al-Mansouri Z, Baghallab I, Bagatian N, Subhi O, Karim S, Al-Aradati
H, Al-Mutawa A, Johary A, Meccawy AA, Al-Ghamdi K, Al-Hamour O, Al-Qahtani
MH, Al-Maghrabi J 2015 Comparison of microarray expression profiles between

17

follicular variant of papillary thyroid carcinomas and follicular adenomas of the thyroid. BMC Genomics 16:S7.

17. Mazzanti C, Zeiger MA, Costouros NG, Umbricht C, Westra WH, Smith D, Somervell H, Bevilacqua G, Alexander HR, Libutti SK, Costourous N 2004 Using gene expression profiling to differentiate benign versus malignant thyroid tumors. Cancer Res 64:2898–2903.

18. Chung KW, Kim SW, Kim SW 2012 Gene expression profiling of papillary thyroid carcinomas in Korean patients by oligonucleotide microarrays. J Korean Surg Soc 82:271–280.

19. Vierlinger K, Mansfeld MH, Koperek O, Nöhammer C, Kaserer K, Leisch F 2011 Identification of SERPINA1 as single marker for papillary thyroid carcinoma through microarray meta analysis and quantification of its discriminatory power in independent validation. BMC Med Genomics 4:30.

20. Griffith OL, Melck A, Jones SJM, Wiseman SM 2006 Meta-analysis and meta-review of thyroid cancer gene expression profiling studies identifies important diagnostic biomarkers. J Clin Oncol 24:5043–5051.

21. Huang Y, Prasad M, Lemon WJ, Hampel H, Wright FA, Kornacker K, LiVolsi V, Frankel W, Kloos RT, Eng C, Pellegata NS, de la Chapelle A 2001 Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. Proc Natl Acad Sci U S A 98:15044–15049.

22. Barden CB, Shister KW, Zhu B, Guiter G, Greenblatt DY, Zeiger MA, Fahey TJ 2003 Classification of follicular thyroid tumors by molecular signature: results of gene profiling. Clin Cancer Res 9:1792–1800.

23. Zhao J, Leonard C, Gemsenjäger E, Heitz PU, Moch H, Odermatt B 2008 Differentiation of human follicular thyroid adenomas from carcinomas by gene expression profiling. Oncol Rep 19:329–337.

24. Lacroix L, Pourcher T, Magnon C, Bellon N, Talbot M, Intaraphairot T, Caillou B, Schlumberger M, Bidart JM 2004 Expression of the apical iodide transporter in human thyroid tissues: a comparison study with other iodide transporters. J Clin Endocrinol Metab 89:1423–1428.

18

25. Chiappetta G, Ferraro A, Vuttariello E, Monaco M, Galdiero F, De Simone V, Califano D, Pallante P, Botti G, Pezzullo L, Pierantoni GM, Santoro M, Fusco A 2008 HMGA2 mRNA expression correlates with the malignant phenotype in human thyroid neoplasias. Eur J Cancer 44:1015–1021.

26. Giannini R, Torregrossa L, Gottardi S, Fregoli L, Borrelli N, Savino M, Macerola E, Vitti P, Miccoli P, Basolo F 2015 Digital gene expression profiling of a series of cytologically indeterminate thyroid nodules. Cancer Cytopathol 123:461–470.

27. Shen Q, Shi WM, Kong W 2009 New gene selection method for multiclass tumor classification by class centroid. J Biomed Inform 42:59–65.

28. Salvatore G, Giannini R, Faviana P, Caleo A, Migliaccio I, Fagin JA, Nikiforov YE, Troncone G, Palombini L, Basolo F, Santoro M 2004 Analysis of BRAF point mutation and RET/PTC rearrangement refines the fine-needle aspiration diagnosis of papillary thyroid carcinoma. J Clin Endocrinol Metab 89:5175–5180.

29. Proietti A, Sartori C, Borrelli N, Giannini R, Materazzi G, Leocata P, Elisei R, Vitti P, Miccoli P, Basolo F 2013 Follicular-derived neoplasms: morphometric and genetic differences. J Endocrinol Invest 36:1055–1061.

30. Nikiforov YE, Steward DL, Robinson-Smith TM, Haugen BR, Klopper JP, Zhu Z, Fagin JA, Falciglia M, Weber K, Nikiforova MN 2009 Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. J Clin Endocrinol Metab 94:2092–2098.

31. Nikiforova MN, Nikiforov YE 2008 Molecular genetics of thyroid cancer: implications for diagnosis, treatment and prognosis. Expert Rev Mol Diagn 8:83–95.

32. Cancer Genome Atlas Research Network 2014 Integrated genomic characterization of papillary thyroid carcinoma. Cell 159:676–690.

33. Yoo SK, Lee S, Kim SJ, Jee HG, Kim BA, Cho H, Song YS, Cho SW, Won JK, Shin JY, Park DJ, Kim JI, Lee KE, Park YJ, Seo JS 2016 Comprehensive analysis of the transcriptional and mutational landscape of follicular and papillary thyroid cancers. PLOS Genet 12:e1006239.

19

34.   Wong KS, Angell TE, Strickland KC, Alexander EK, Cibas ES, Krane JF, Barletta JA
        2016 Noninvasive follicular variant of papillary thyroid carcinoma and the
        Afirma gene-expression classifier. Thyroid 26:911–915.

35.   Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE,
        LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN 2015 Impact
        of the multi-gene ThyroSeq next-generation sequencing assay on cancer
        diagnosis in thyroid nodules with atypia of undetermined significance/follicular
        lesion of undetermined significance cytology. Thyroid 25:1217–1223.

36.   Lupi C, Giannini R, Ugolini C, Proietti A, Berti P, Minuto M, Materazzi G, Elisei R,
        Santoro M, Miccoli P, Basolo F 2007 Association of BRAF V600E mutation with
        poor clinicopathological outcomes in 500 consecutive cases of papillary thyroid
        carcinoma. J Clin Endocrinol Metab 92:4085–4090.

37.   Basolo F, Torregrossa L, Giannini R, Miccoli M, Lupi C, Sensi E, Berti P, Elisei R,
        Vitti P, Baggiani A, Miccoli P 2010 Correlation between the BRAF V600E
        mutation and tumor invasiveness in papillary thyroid carcinomas smaller than
        20 millimeters: analysis of 1060 cases. J Clin Endocrinol Metab 95:4197–4205.

38.   Mete O, Asa SL 2011 Pathological definition and clinical significance of vascular
        invasion in thyroid carcinomas of follicular epithelial derivation. Mod Pathol
        24:1545–1552.

39.   Maletta F, Massa F, Torregrossa L, Duregon E, Casadei GP, Basolo F, Tallini G,
        Volante M, Nikiforov YE, Papotti M 2016 Cytological features of "noninvasive
        follicular thyroid neoplasm with papillary-like nuclear features" and their
        correlation with tumor histology. Hum Pathol 54:134-142.

40.   Faquin WC, Wong LQ, Afrogheh AH, Ali SZ, Bishop JA, Bongiovanni M,
        Pusztaszeri MP, VandenBussche CJ, Gourmaud J, Vaickus LJ, Baloch ZW 2016
        Impact of reclassifying noninvasive follicular variant of papillary thyroid
        carcinoma on the risk of malignancy in the Bethesda system for reporting
        thyroid cytopathology. Cancer Cytopathol 124:181–187.

20

Table 1. Custom panel of deregulated genes in thyroid carcinomas.

GENES	APPROVED NAME	TRANSCRIPT	DEREGULATED IN*	EXPRESSION	REFERENCE(S)
Growth factors & receptors					
AGTR1	Angiotensin II receptor type 1	NM_000685.4	FTC	DOWN	12
CTGF	Connective tissue growth factor	NM_001901.2	FTC	DOWN	12
GABBR2	Gamma-aminobutyric acid type B receptor subunit 2	NM_005458	FTC, FVPTC, PTC	UP	13, 14
KIT	KIT proto-oncogene receptor tyrosine kinase	NM_000222.2	FVPTC, PTC	DOWN	14, 15
MET	MET proto-oncogene, receptor tyrosine kinase	NM_000245	FTC, FVPTC, PTC	UP	16-19
NR4A1	Nuclear receptor subfamily 4 group A member 1	NM_001202233.1	FTC	DOWN	12
NR4A3	Nuclear receptor subfamily 4 group A member 3	NM_006981.3	FTC	DOWN	12
PPARG	Peroxisome proliferator activated receptor gamma	NM_005037.5			
PTCHD4	Patched domain containing 4	NM_001013732	FVPTC	UP	14
RXRG	retinoid X receptor gamma	NM_001256570	FVPTC, PTC	UP	14, 16
Extracellular matrix & cellular adhesion					
CDH3	Cadherin 3	NM_001793	PTC	UP	16, 17
COL9A3	Collagen type IX alpha 3	NM_001853.3	FTC, FVPTC, PTC	DOWN	18
CSGALNACT1	Chondroitin sulfate N-acetylgalactosaminyltransferase 1	NM_001130518	FVPTC	DOWN	14
DPP4	Dipeptidyl-peptidase 4	NM_001935	FTC, FVPTC, PTC	UP	17-19
ECM1	Extracellular matrix protein 1	NM_001202858	FVPTC	UP	14
FN1	Fibronectin 1	NM_212482	FTC, FVPTC, PTC	UP	13, 18, 19
IGSF1	Immunoglobulin superfamily member 1	NM_001170961	FVPTC	UP	14
LAMB3	Laminin subunit beta 3	NM_000228	PTC	UP	17
LGALS3	Lectin, galactoside-binding, soluble, 3	NM_001177388.1	FTC, FVPTC, PTC	UP	18, 19
PLP1	Proteolipid protein 1	NM_000533	FVPTC	UP	14
PROS1	Protein S (alpha)	NM_000313 2.12	FTC, FVPTC, PTC	UP	16-20
SDC4	Syndecan 4	NM_002999	PTC	UP	17, 19
SERPINA1	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	NM_000295	FTC, FVPTC, PTC	UP	17-19
SLIT1	Slit guidance ligand 1	NM_003061	FVPTC	UP	14
TIMP1	TIMP metallopeptidase inhibitor 1	NM_003254	FTC, FVPTC, PTC	UP	13, 17-19
Thyroid metabolism					
DIO1	Deiodinase, iodothyronine, type I	NM_000792.5	FTC, FVPTC, PTC	DOWN	18, 19
DIO2	Deiodinase, iodothyronine, type II	NM_000793.5	FVPTC, PTC	DOWN	15, 19
DUOX1	Dual oxidase 1	NM_017434.4			
DUOX2	Dual oxidase 2	NM_014080.4			
SLC26A4	Solute carrier family 26 (anion exchanger), member 4	NM_000441.1	FTC, PTC	DOWN	12, 22, 23
SLC5A5	Solute carrier family 5 (sodium/iodide cotransporter), member 5	NM_000453.2	FTC, PTC	DOWN	19, 22
SLC5A8	Solute carrier family 5 (sodium/monocarboxylate	NM_145913.3	FTC, PTC	DOWN	22

21

Symbol	Name	Accession	Subtype	Direction	Ref
	cotransporter), member 8				
TG	Thyroglobulin	NM_003235.4			
TPO	Thyroid peroxidase	NM_000547.5	FTC, FVPTC, PTC	DOWN	13, 18, 19
Transcription factor & cell cycle regulation					
ASPM	Abnormal spindle microtubule assembly	NM_001206846	FTC	UP	12
BUB1B	BUB1 mitotic checkpoint serine/threonine kinase B	NM_001211	FTC	UP	12
CCNB2	Cyclin B2	NM_004701	FTC	UP	12
CDCA5	Cell division cycle associated 5	NM_080668	FTC	UP	12
CENPF	Centromere protein F	NM_016343	FTC	UP	12
CEP55	Centrosomal protein 55 kDa	NM_001127182	FTC	UP	12
EGR2	early growth response 2	NM_000399.3	FTC	DOWN	12, 20, 21
FOSB	FBJ murine osteosarcoma viral oncogene homologue B	NM_001114171.1	FTC,PTC	DOWN	12, 19
HMGA2	High mobility group AT-hook 2	NM_003483.4	FVPTC, PTC	UP	14, 23
KIF4A	Kinesin family member 4A	NM_012310	FTC	UP	12
MELK	Maternal embryonic leucine zipper kinase	NM_001256685.1	FTC	UP	12
NEK2	NIMA related kinase 2	NM_001204182.1	FTC	UP	12
NUSAP1	Nucleolar and spindle associated protein 1	NM_001129897.1	FTC	UP	12
PRC1	Protein regulator of cytokinesis 1	NM_001267580.1	FTC	UP	12
TFCP2L1	Transcription factor CP2-like 1	NM_014553.2	FVPTC	DOWN	14
TPX2	TPX2, microtubule-associated	NM_012112.4	FTC	UP	12
Signal transduction					
CHI3L1	Chitinase 3-like 1	NM_001276	FTC, FVPTC, PTC	UP	17-19
CITED1	Cbp/p300-interacting transactivator, with Glu/Asp rich carboxy-terminal domain, 1	NM_004143	FTC, FVPTC, PTC	UP	17-19
DGKI	Diacylglycerol kinase iota	NM_004717.2	FVPTC	DOWN	14
LRP4	LDL receptor related protein 4	NM_002334	FTC, FVPTC, PTC	UP	18, 19
LRRK2	Leucine-rich repeat kinase 2	NM_198578	FVPTC	UP	14
TENM1	Teneurin transmembrane protein 1	NM_001163278	FVPTC, PTC	UP	14, 19
PBK	PDZ binding kinase	NM_001278945.1	FTC	UP	14
SDPR	Serum deprivation response	NM_004657.5	FTC	DOWN	12
Others					
CKB	Creatine kinase, brain	NM_001823.4	FTC, FVPTC, PTC	DOWN	13
CWH43	Cell wall biogenesis 43 C-terminal homolog	NM_001286791.1	FVPTC	DOWN	14
DNASE1L3	Deoxyribonuclease I-like 3	NM_001256560.1	FTC	DOWN	12
GGCT	Gamma-glutamylcyclotransferase	NM_024051	PTC	UP	17
IPCEF1	Interaction protein for cytohesin exchange factors 1	NM_001130699.1	FVPTC, PTC	DOWN	14, 17
LIPH	Lipase H	NM_139248	FVPTC	UP	14
MPPED2	Metallophosphoesterase domain containing 2	NM_001584	FVPTC, PTC	DOWN	15,17
NPC2	Niemann-Pick disease, type C2	NM_006432	PTC	UP	17
PDLIM4	PDZ and LIM domain 4	NM_003687	PTC	UP	17
PRSS23	Protease, serine 23	NM_007173	PTC	UP	17
QPCT	Glutaminyl-peptide cyclotransferase	NM_012413	FTC, FVPTC, PTC	UP	17, 18
RRM2	Ribonucleotide reductase regulatory	NM_001034.3	FTC	UP	12

	subunit M2				
SCEL	Sciellin	NM_001160706	FVPTC, PTC	UP	14, 16, 19
SERTM1	Serine rich and transmembrane domain containing 1	NM_203451.2	FVPTC	DOWN	14
TFF3	Trefoil factor 3	NM_003226	FTC, FVPTC, PTC	DOWN	13, 17-20
TOP2A	Topoisomerase (DNA) II alpha	NM_001067.3	FTC	UP	12
UBE2C	Ubiquitin conjugating enzyme E2C	NM_001281741.1	FTC	UP	12
Housekeeping					
ACTB	Actin, beta	NM_001101			
B2M	Beta-2-microglobulin	NM_004048			
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	NM_002046.5			
GAPHD_v2	Glyceraldehyde-3-phosphate dehydrogenase, transcript variant 2	NM_001256799.2			
HPRT1	Hypoxanthine phosphoribosyltransferase 1	NM_000194			

*Compared to Follicular Adenoma (FA).

FTC: Follicular Thyroid Carcinoma; FVPTC: Follicular Variant of Papillary Thyroid Carcinoma; PTC: Papillary Thyroid Carcinoma.

23

Table 2. Results of the classification, indicating the class assigned, the discriminant score for the assignment and the sample genotypes.

Sample	27-GENE CLASSIFIER		Genotype
	Class assigned*	DS	
NIFTP 15	IFVPTC-LIKE	13.08	BRAF K601E
NIFTP 16	FA-LIKE	16.83	WT
NIFTP 27	FA-LIKE	10.10	HRAS Q61R
NIFTP 28	IFVPTC-LIKE	6.48	NRAS Q61R
NIFTP 29	IFVPTC-LIKE	11.33	NRAS Q61R
NIFTP 43	FA-LIKE	6.23	WT
NIFTP 44	IFVPTC-LIKE	11.89	WT
NIFTP 52	FA-LIKE	8.36	WT
NIFTP 53	FA-LIKE	13.18	WT
NIFTP 54	IFVPTC-LIKE	13.08	NRAS Q61R
NIFTP 55	IFVPTC-LIKE	25.50	BRAF V600E
NIFTP 56	IFVPTC-LIKE	16.79	WT
NIFTP 57	FA-LIKE	14.25	WT
NIFTP 6	IFVPTC-LIKE	11.79	NRAS Q61R
NIFTP 7	FA-LIKE	6.23	NRAS Q61R
NIFTP 71	IFVPTC-LIKE	10.31	BRAF K601E
NIFTP 72	FA-LIKE	17.02	WT
NIFTP 73	FA-LIKE	17.66	WT
NIFTP 74	IFVPTC-LIKE	16.77	HRAS Q61R
NIFTP 76	IFVPTC-LIKE	29.94	HRAS Q61R
NIFTP 77	FA-LIKE	16.25	WT
NIFTP 78	IFVPTC-LIKE	15.02	NRAS Q61R
NIFTP 79	FA-LIKE	6.63	WT
NIFPT84	IFVPTC-LIKE	12.22	BRAF K601E
NIFTP 95	FA-LIKE	14.93	WT
NIFTP 96	FA-LIKE	8.75	WT

FA: Follicular Adenoma; IFVPTC: Infiltrative Follicular Variant of Papillary Thyroid Carcinoma; NIFTP: "non-invasive follicular tumour with papillary-like nuclear features".
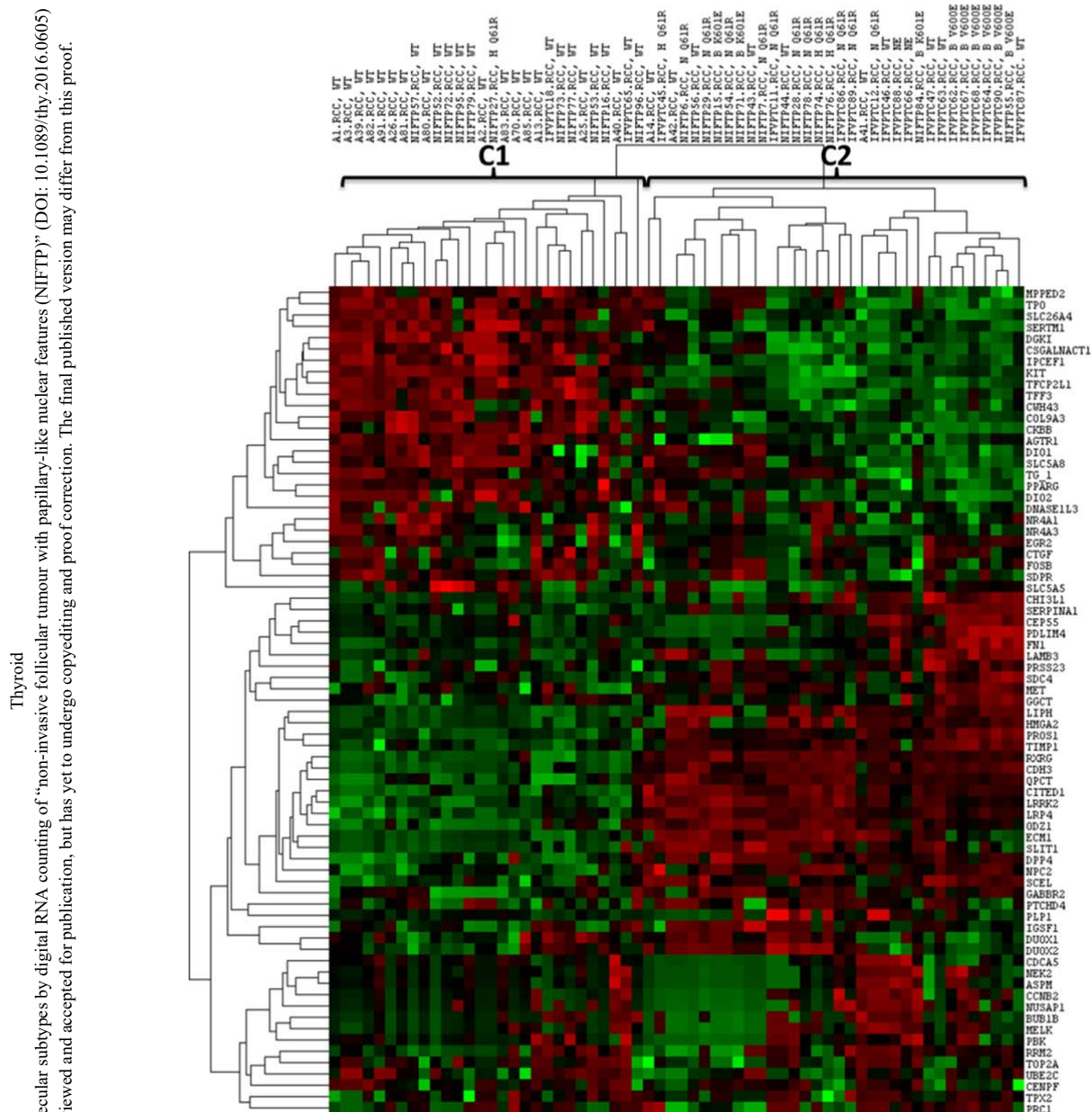
24

Figure legends.

Figure 1
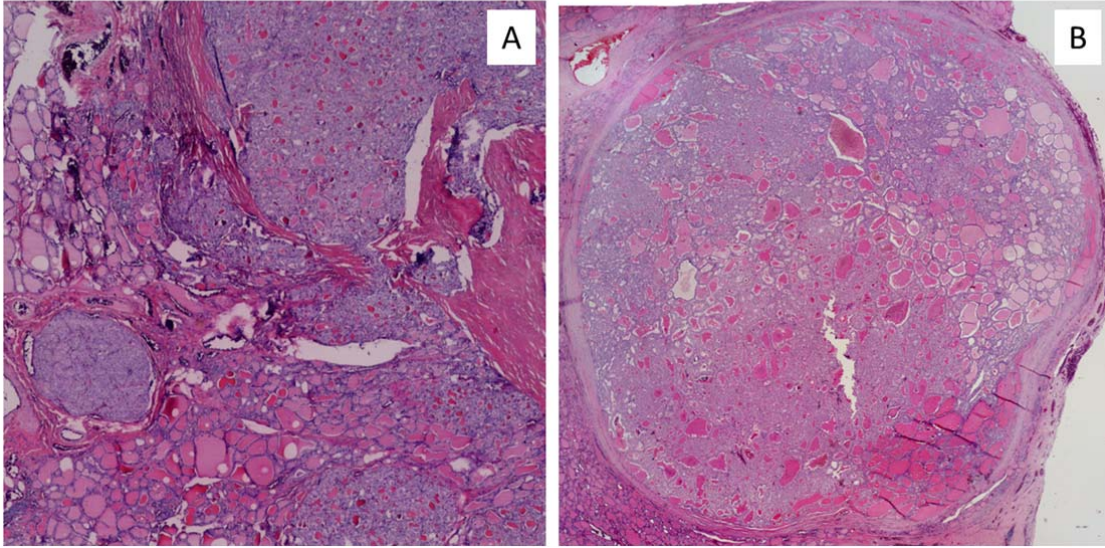
Figure 1: (A) Haematoxylin and eosin (H&E)-stained slides from an IFVPTC; (B) Haematoxylin and eosin (H&E)-stained slides from a whole section of an NIFTP. Both images were reconstituted from a series of digital images (20X magnification) using the Photomerge tool in Adobe Photoshop CS3 (Adobe Systems).

25

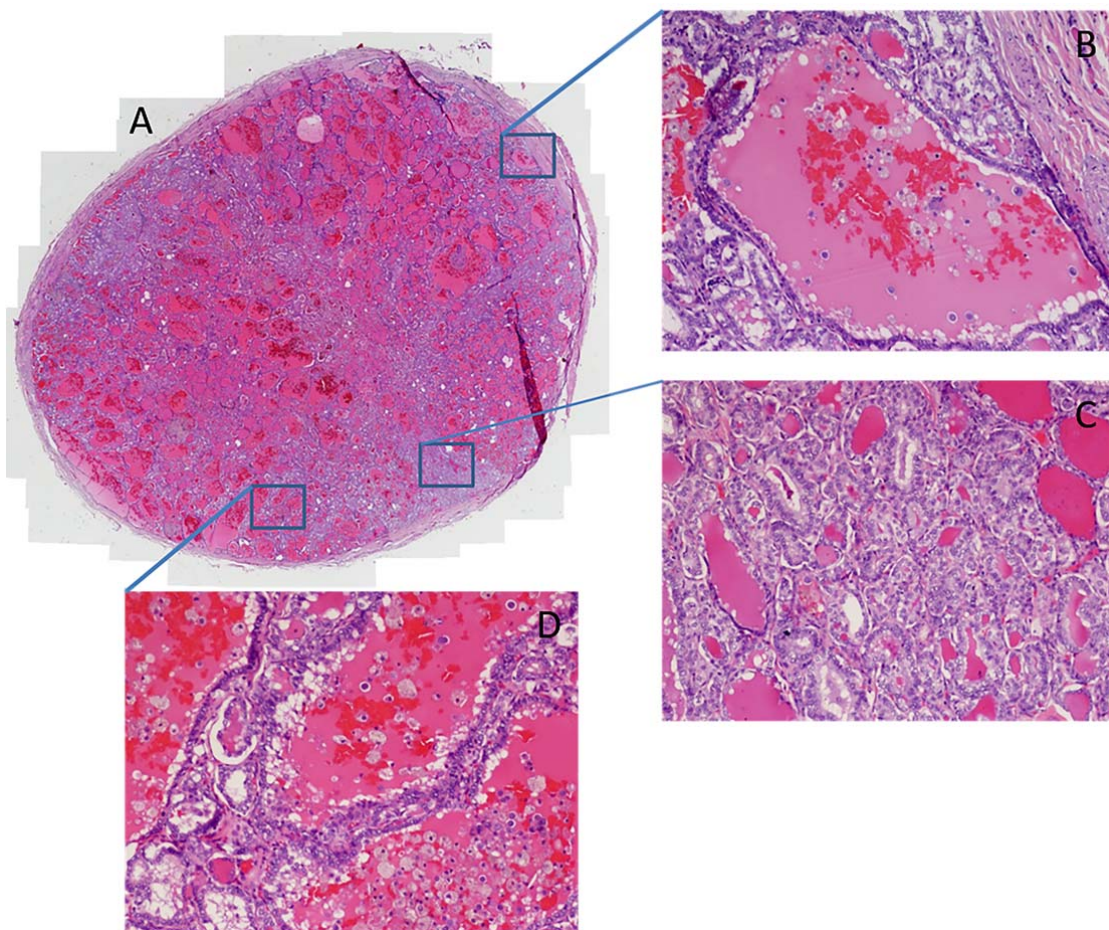## Figure 2

Figure 2: Pearson unsupervised hierarchical clustering of A (FA), NIFTP and IFVPTC samples are illustrated according to gene expression levels as measured by the nCounter System (NanoString). Each column represents a single sample, and each row represents a single gene. Red indicates a high level of expression relative to the mean expression, and green indicates a low level of expression relative to the mean expression. C1 and C2 indicate the two main samples clusters. RCC is the file extension. WT indicates the absence of mutations; H Q61R indicates *HRAS* Q61R mutation; N Q61R indicates *NRAS* Q61R

26

mutation; B K601E indicates BRAF K601E mutation; B V600E indicates BRAF V600E mutation. NE indicates genotype not available.

Figure 3: NIFTP harbouring BRAFV600E mutation: (A) Haematoxylin and eosin (H&E)-stained slides from a whole section of the BRAFV600E mutated NIFTP. Image was reconstituted from a series of digital images (20X magnification) using the Photomerge tool in Adobe Photoshop CS3 (Adobe Systems); (B, C and D) Higher magnification (100X magnification) of different areas.

28
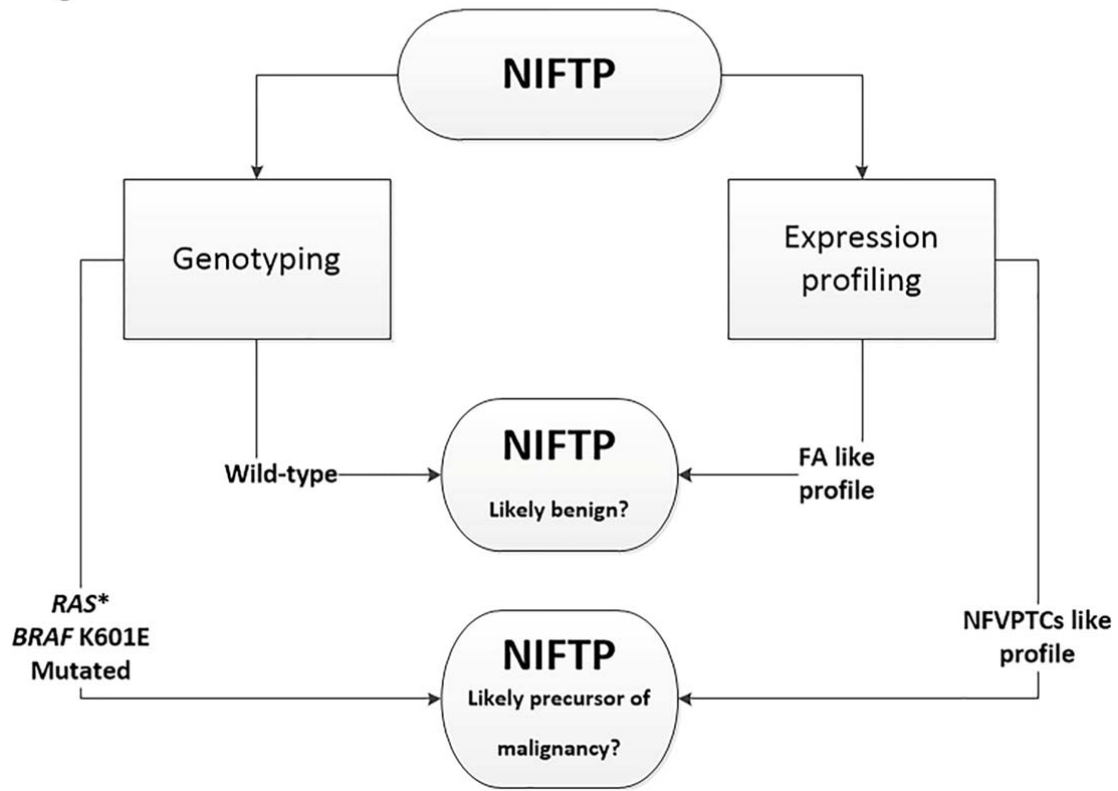
Figure 3

Figure 4: Schematic representation of putative "molecular classification" of NIFTP.