UNIVERSITY OF PISA

Research Doctorate School in
BIOLOGICAL AND MOLECULAR SCIENCES

Course in
EXPERIMENTAL AND MOLECULAR ONCOLOGY
SSD: MED/06
XXIV CYCLE (2009-2011)

THESIS TITLE:

Molecular marker combinations preoperatively differentiate benign from malignant thyroid tumors

Supervisor:
Generoso Bevilacqua

Candidate:
Sara Tomei

Sara Tomei

Tutor:
Chiara Mazzanti

"The important thing is not to stop questioning. Curiosity has its own reason for existing.
One cannot help but be in awe when he/she contemplates the mysteries of
eternity, of life, of the marvelous structure of reality.
It is enough if one tries merely to comprehend a little of this mystery every day.
Never lose a holy curiosity."

Albert Einstein

To Riccardo and Silvia

INDEX

ABBREVIATIONS

AUC: Area Under the Curve

BN: benign nodule

BNN: Bayesian Neural Network

B2M: beta-2 microglobulin

CD: cytological diagnosis

CI: confidence interval

DDI2: DNA-damage inducible 1 homolog 2

ev: expression value

FNA: Fine-needle Aspiration

FNAC: FNA Cytology

FNAB: FNA Biopsy

HD: histological diagnosis

IFP: indeterminate follicular proliferation

IRB: Internal Review Board

MAPK: Mitogen Activated Protein Kinase

NATH: N-acetyl transferase human

PCA: Principal Component Analysis

PI3K: Phosphoinositide 3-kinase

PNN: Probabilistic Neural Network

PTC: Papillary thyroid carcinoma

qPCR: quantitative-PCR

ROC: Receiver Operating Characteristic

SCF: Stem cell factor

SPTC: Suspicious PTC

SYNGR2: Synaptogyrin 2

TC1: thyroid cancer-1

TRK: tyrosine receptor kinase

WT: wild type

ABSTRACT

Background. The initial presentation of thyroid carcinoma is through a nodule and the best way nowadays to evaluate it is by fine-needle aspiration (FNA). However many thyroid FNAs are not definitively benign or malignant, yielding an indeterminate or suspicious diagnosis which ranges from 10 to 25% of FNAs. The development of molecular initial diagnostic tests for evaluating a thyroid nodule is needed in order to define optimal surgical approach for patients with uncertain diagnosis pre- and intra-operatively.

A large amount of information has been collected on the molecular tumorigenesis of thyroid cancer. A low expression of KIT gene has been reported during the transformation of normal thyroid epithelium to papillary carcinoma suggesting a possible role of the gene in the differentiation of thyroid tissue rather than in the proliferation. Moreover, several gene expression studies have shown differential gene expression signatures between malignant and benign thyroid tumors.

The aim of the current study was to determine the diagnostic utility of a molecular assay based on the gene expression of a panel of molecular markers (KIT, SYNGR2, C21orf4, Hs.296031, DDI2, CDH1, LSM7, TC1, NATH) plus BRAF mutational status to distinguish benign from malignant thyroid neoplasm.

Methods. The mRNA expression level of 9 genes (KIT, SYNGR2, C21orf4, Hs.296031, DDI2, CDH1, LSM7, TC1, NATH) was analyzed by quantitative Real-Time PCR (qPCR) in 93 FNA cytological samples. To evaluate the diagnostic utility of all the genes analyzed, we assessed the area under the curve (AUC) for each gene individually and in combination. BRAF exon 15 status was determined by capillary sequencing. A gene expression computational model (Neural Network Bayesian Classifier) was built and a multiple-variable analysis was then performed to analyze the correlation between the markers.

Results. While looking at KIT expression, we have found a highly preferential decrease rather than increase in transcript of KIT in malignant thyroid lesions compared to the benign ones. To explore the diagnostic utility of KIT expression in thyroid nodules, its expression values were divided in four arbitrarily defined classes, with class I characterized by the complete

silencing of the gene. Class I and IV represented the two most informative groups, with 100% of the samples found malignant or benign respectively. The molecular analysis was proven by ROC (receiver operating characteristic) analysis to be highly specific and sensitive improving the cytological diagnostic accuracy of 15%.

The AUC for each significant marker was further assessed and ranged between 0.625 and 0.900, thus all the significant markers, alone and in combination, can be used to distinguish between malignant and benign FNA samples. The classifier made up of KIT, CDH1, LSM7, C21orf4, DDI2, TC1, Hs.296031 and BRAF had a predictive power of 88.8%. It proved to be useful for risk stratification of the most critical cytological group of the indeterminate lesions for which there is the greatest need for accurate diagnostic markers.

Conclusion. The genetic classification obtained with such a model is highly accurate and may provide a tool to overcome the difficulties in today's pre-operative diagnosis of thyroid malignancies

Keywords: thyroid cancer, Fine-needle aspiration (FNA), computational model, Receiver Operating Characteristic (ROC) analysis, Area Under the Curve (AUC), pre-operative diagnosis.

1. INTRODUCTION

Thyroid cancer is the most common malignancy of endocrine organs. The vast majority of thyroid tumors arise from thyroid follicular epithelial cells, whereas 3-5% of cancers originate from parafollicular or C cells. The follicular cell-derived cancers are further subdivided into well-differentiated papillary carcinoma and follicular carcinoma, poorly differentiated carcinoma (also known as insular carcinoma) and anaplastic (undifferentiated) carcinoma (1, 2). Follicular adenoma is a benign tumor that may serve as a precursor for some follicular carcinoma. Less-differentiated thyroid cancers, namely poorly differentiated carcinoma and anaplastic carcinoma, can develop *de novo*, although many of them arise through the process of step-wise dedifferentiation of papillary and follicular carcinoma (Fig. 1) (3).

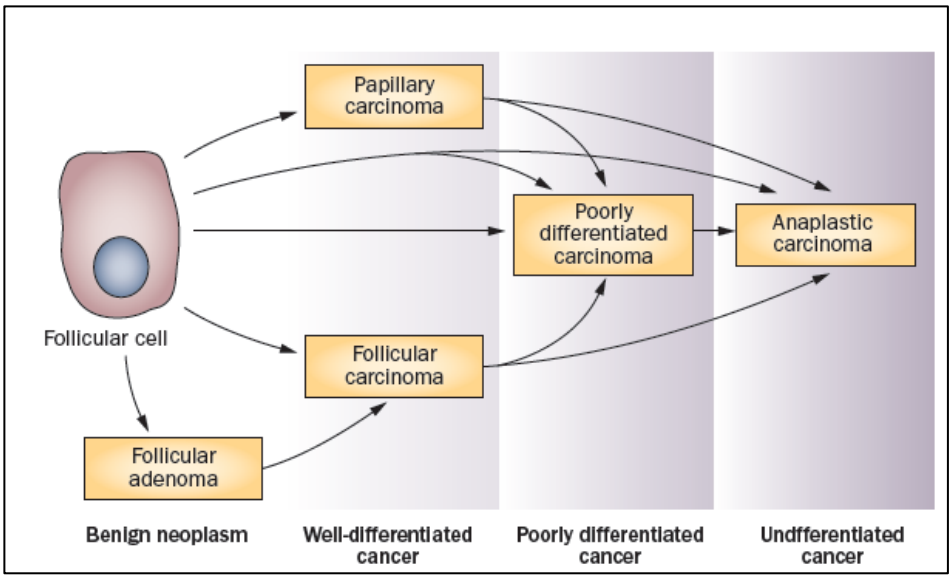

Figure 1. Scheme of step-wise dedifferentiation of follicular cell-derived thyroid cancer. From *Expert Reviews of Molecular Diagnostics* **8** (1), 83-89 (2008).

The worldwide incidence of thyroid cancer has been steadily increasing and has almost tripled over the past 30 years in the US and in the industrialized countries (4-7).

According to European Cancer Observatory data, age-standardized incidence rate was 3.1 cases and 8.8 cases per 100 000 in men and women, respectively in Europe in 2008 and such

incidence rates have steadily increased over the recent decades (8, 9). More than 30 000 cases are newly-diagnosed in the US population in the year 2004 (10).

Most of the increase in thyroid cancer incidence has been attributed to the diagnosis of small (<2 cm), papillary thyroid cancer with no significant change in the less common variants of thyroid cancer (follicular, medullary and anaplastic) (7).

The increase in thyroid cancer incidence is generally believed to result, to a considerable extent, from increased access to high-resolution imaging (particularly ultrasonography) and increased use of fine-needle aspiration (FNA) biopsy of small nodules, as well as progressively decreasing stringency of histopathologic criteria applied to the diagnosis of papillary cancer during the past 10-15 years. However, whether these factors can account entirely for this continuous trend, or whether other factors also contribute, remains unknown. Ionizing radiation is a well-known risk factor for thyroid cancer; therefore, concerns remain that the rising incidence might, in part, be due to the wider use of medical radiation and increased exposure to radiation as a result of nuclear power incidence such as that in Chernobyl.

Also, the number of patients who have no diagnostic and indeterminate fine needle aspiration biopsy of a thyroid nodule is likely to encompass an even greater number of patients because of the increasing incidence of thyroid nodules that are detected incidentally (11). Although the overall prognosis of most patients with thyroid cancer of follicular cell origin (except for anaplastic thyroid cancer) is excellent, the optimal treatment strategy is controversial (the extent of thyroidectomy, need for prophylactic versus therapeutic lymph node dissection, and routine use and appropriate dose of radioiodine ablation).

Thyroid cancers typically occur in thyroid nodules, which are common and can be detected by palpation and imaging in a large proportions of adults, particularly those of increased age (12-14). Most thyroid nodules are benign, and the clinical challenge is to accurately and rapidly identify those nodules that harbor cancer. Sampling of thyroid nodules using FNA biopsy with subsequent cytological examination of collected cells is the most accurate and widely used diagnostic tool at this time. It provides a definitive diagnosis of malignant or benign nodules in most cases. However, a conclusive diagnosis cannot be obtained by use of FNA cytology for about 25% of all nodules (13, 15-18), which hampers the clinical management of patients with these nodules. New diagnostic approaches for such nodules are needed.

The knowledge of genetic alterations occurring in thyroid cancer has rapidly expanded in the past decade. This improved knowledge has provided new insights into thyroid cancer etiology and has offered novel diagnostic tools and prognostic markers that enabled improved and personalized management of patients with thyroid nodules.

1.1 Common alterations in thyroid cancer

Similar to other cancer types, thyroid cancer initiation and progression occur through gradual accumulation of various genetic and epigenetic alterations, including activating and inactivating somatic mutations, alterations in gene expression patterns, microRNA (miRNA) deregulation and aberrant gene methylation. Among these alterations, most of the data that have accumulated relate to somatic mutations, many of which occur early in the transformation process and are essential for cancer development. Point mutation and chromosomal rearrangements are very frequent in thyroid cancer progression. The former is a result of single nucleotide change within the DNA chain, whereas the latter represents a large-scale genetic abnormality with breakage and fusion of parts of the same or different chromosomes. Importantly, a growing body of evidence suggests that these two distinct mutational mechanisms are associated with specific etiologic factors involved in thyroid carcinogenesis.

1.2 Somatic mutations

Most mutations in thyroid cancer involve the effectors of the MAPK pathway and the PI3K-AKT pathway (Fig. 2). MAPK activation is crucial for tumor initiation. The mutated genes that affect these pathways encode cell-membrane receptor tyrosine kinases RET and NTRK1 and intracellular transducers BRAF and RAS.

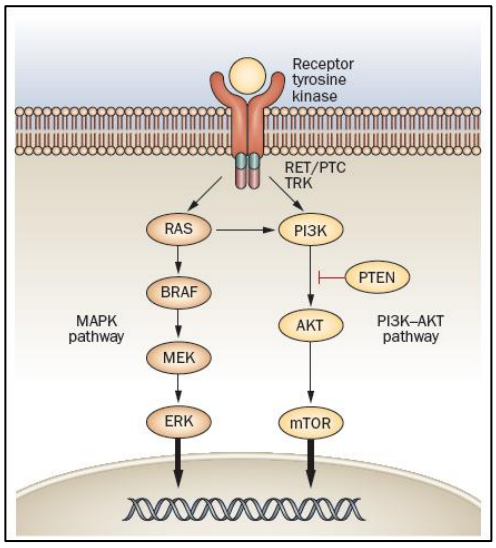

Figure 2. The main signaling pathways involved in thyroid carcinogenesis are the MAPK and PI3K-AKT pathways (Nikiforov, Y.E. & Nikiforova, M. N., Nature Reviews, 2011)

These typically mutually exclusive mutations occur approximately 70% of patients with papillary thyroid carcinomas and are associated with particular clinical, histopathological and biological tumor characteristics (Fig. 3) (19-22).

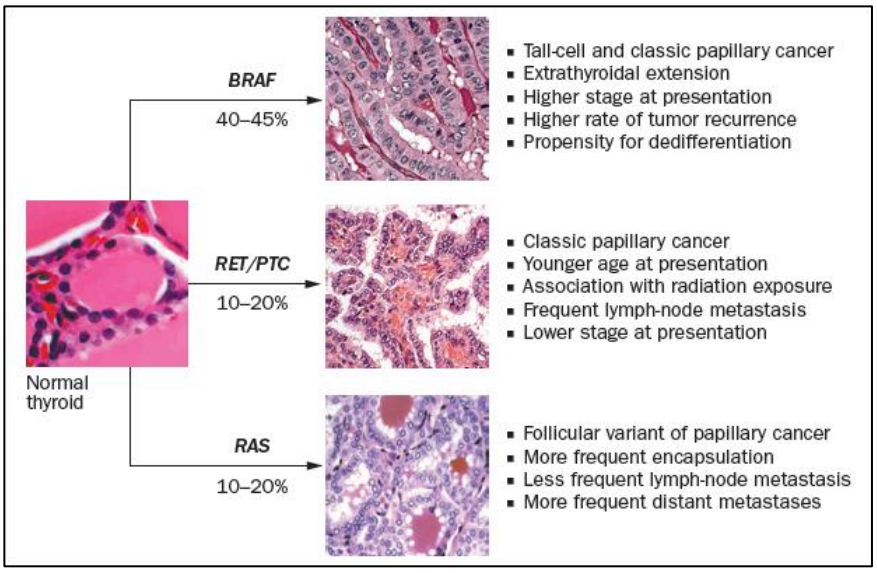

Figure 3. Molecular alterations in papillary thyroid cancer and their average prevalence and association with clinical and histopathological features of tumors. From *Expert Reviews of Molecular Diagnostics* **8** (1), 83-89 (2008)

In follicular thyroid cancer, in addition to mutation of RAS, another common event is PAX8/PPARγ rearrangement. Thyroid cancer progression and dedifferentiation involves a number of additional mutations that affect the PI3K-AKT pathway and other cell signaling pathways.

1.2.1 BRAF

BRAF is a serine-threonine kinase that belongs to the family of RAF proteins, which are intracellular effectors of the MAPK signaling cascade. Upon activation triggered by RAS binding and protein recruitment to the cell membrane, these kinases phosphorylate and activate MEK, which in turn activates ERK and consequent effectors of the MAPK cascade.

Point mutations of the BRAF gene are found in about 45% of thyroid papillary carcinomas (20, 23). Virtually all point mutations involve nucleotide 1799 (generally T>A, Fig. 4) and result in a valine-to-glutamate substitution at residue 600 (V600E) (24, 25).

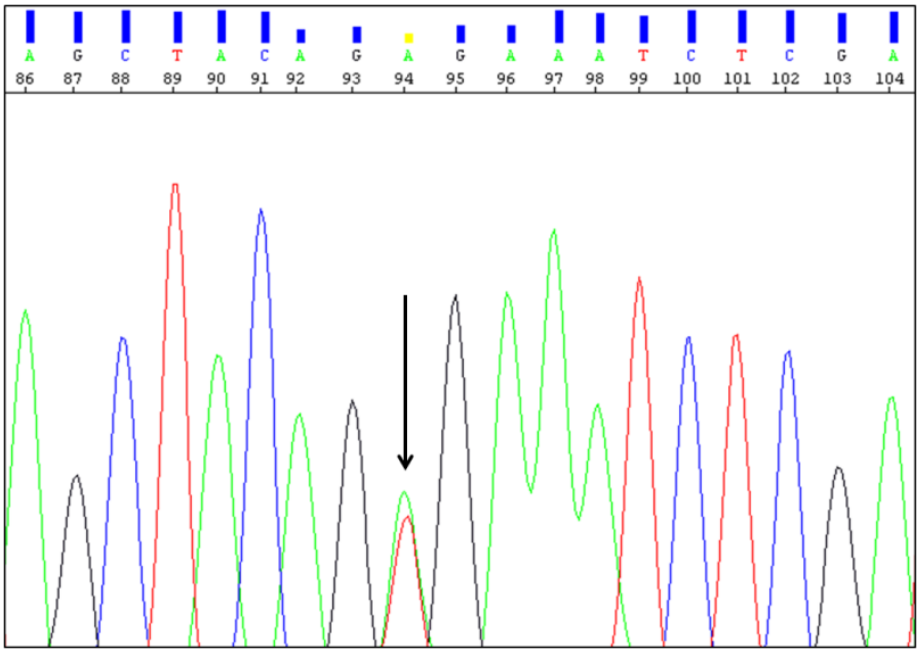

Figure 4. BRAF 1799T>A point mutation.

BRAF V600E mutation leads to constitutive activation of BRAF kinase and the mechanisms of activation have been recently elucidated. In the dephosphorylated, wild type BRAF protein the hydrophobic interactions between the activation loop and the ATP binding site maintains the protein in an inactive conformation. The V600E substitution disrupts these interactions and allows the formation of new interactions that keep the protein in a catalytically competent conformation, resulting in continuous phosphorylation of MEK.

BRAF mutations are highly prevalent in papillary carcinomas with classical histology and in the tall cell variant, but are rare in the follicular variant (19, 24). In many studies, the presence of BRAF mutation has been found to correlate with aggressive tumor characteristics such as extra thyroidal extension, advanced tumor stage at presentation, tumor recurrence, and lymph node or distant metastases (26-28). Importantly, BRAF V600E mutation has been found to be an independent predictor of tumor recurrence even in patients with stage I-II of the disease (29, 30). BRAF mutations have also been associated with the decreased ability of tumors to trap radioiodine and treatment failure of the recurrent disease, which may be due to the deregulation of function of the sodium iodide symporter (NIS) and other genes metabolizing iodide in thyroid follicular cells (29, 31).

Other and rare mechanisms of BRAF activation in papillary thyroid cancer include K601E point mutation, small in-frame insertions or deletion surrounding codon 600 (32-34). In addition to papillary carcinomas, BRAF is found mutated in thyroid anaplastic and poorly differentiated carcinomas, typically in those tumors that also contain areas of well-differentiated papillary carcinoma (27, 28, 35). In those tumors, BRAF mutation is detectable in both well-differentiated and poorly differentiated or anaplastic tumor areas, providing evidence that it occurs early in tumorigenesis.

1.2.2 RET/PTC

The RET proto-oncogene codes for a cell membrane receptor tyrosine kinase. In the thyroid gland, RET is highly expressed in parafollicular C-cells but not in follicular cells, where it can be activated by chromosomal rearrangement known as RET/PTC rearrangement (36, 37). In RET/PTC, the 3’ portion of the RET gene is fused to the 5’ portion of various unrelated

genes. At least 11 types of RET/PTC have been reported to date, all formed by the RET fusion to different partners (38, 39). The two most common rearrangement types, RET/PTC1 and RET/PTC3 account for the vast majority of all rearrangements found in papillary carcinomas. Several studies suggest that the oncogenic effects of RET/PTC require signaling along the MAPK pathway and the presence of the functional BRAF kinase (40-42). Indeed, BRAF silencing in cultured thyroid cells reverses the RET/PTC-induced effects such as ERK phosphorylation, inhibition of thyroid specific gene expression, and increased cell proliferation (41, 42). RET/PTC is found on average in about 20% of adult sporadic papillary carcinomas, although its prevalence is highly variable between different observations (38, 39). In general RET/PTC incidence is higher in tumors from patients with a history of radiation exposure and in pediatric populations. The distribution of RET/PTC rearrangement within each tumor may vary from involving almost all neoplastic cells (clonal RET/PTC) to being detected only in a small fraction of tumor cells (non-clonal RET/PTC) (43, 44). The heterogeneity may be of potential problem for the RET receptor-targeted therapy, since tumors with non-clonal RET/PTC frequently have other genetic alterations and may not respond to RET inhibitors in the same way as tumors harboring the clonal rearrangement.

1.2.3 RAS

The RAS genes (HRAS, KRAS and NRAS) encode highly related G-proteins that are located at the inner surface of the cell membrane and play a central role in the intracellular transduction of signals arising from cell membrane receptors tyrosine kinase and G-protein-coupled receptors. In its inactive state, RAS protein is bound to guanosine diphosphate (GDP). Upon activation, it releases GDP and binds guanosine triphospate (GTP), activating the MAPK and other signaling pathway, such as PI3K/AKT. Normally, the activated RAS-GTP protein becomes quickly inactive due to its intrinsic guanosine triphosphatase (GTPase) activity and the action of cytoplasmic GTPase-activating proteins, which catalyze the conversion of the active GTP form to the inactive GDP-bound form. In many human neoplasms, point mutations occur in the discrete domain of the RAS gene, which result in either an increased affinity for GTP (mutations in codons 12 and 13) or inactivation of the autocatalytic GTPase function

(mutations in codon 61). As a result, the mutant protein becomes permanently switched in the active position and constitutively activates its downstream signaling pathways.

Point mutations of RAS occur with variable frequency in all types of papillary thyroid follicular cell-derived tumors. In papillary carcinomas, RAS mutations are relatively infrequent, as they occur in about 10% of tumors (45, 46). Papillary carcinomas with RAS mutations almost always have the follicular variant histology; this mutation also correlates with significantly less prominent nuclear features of papillary carcinoma, more frequent encapsulation, and low rate of lymph node metastases (19, 47). Some studies have reported the association between RAS mutations and more aggressive behavior of papillary carcinoma and with higher frequency of distant metastases (48). In follicular thyroid carcinomas, RAS mutations are found in 40-50% of tumors (49-51) and may also correlate with tumor dedifferentiation and less favorable prognosis (52, 53). RAS mutations may predispose to tumor dedifferentiation, as they are found with high prevalence in anaplastic (undifferentiated) thyroid carcinomas. This may be due to the effect of mutant RAS to promote chromosomal instability, which has been documented in the *in vitro* settings. RAS mutations, however, are not specific for thyroid malignancy and also occur in benign follicular adenomas (54).

1.2.4 PAX8/PPARγ

PAX8/PPARγ rearrangement results from the translocation t(2;3)(q13;p25) that leads to the fusion between the PAX8 gene, which encodes a paired domain transcription factor, and the peroxisome proliferator-ativated receptor (PPARγ) gene (55). PAX8-PPARγ occurs in about 35% of conventional follicular carcinomas, and with lower prevalence in oncocytic (Hurtle cell) carcinomas (56-58). Tumors harboring PAX8-PPARγ rearrangement tend to present at a younger age, be smaller in size, and more frequently have vascular invasion. The rearrangement causes the over expression of the PPARγ protein which can be detected by immunohistochemistry (55, 59).

The mechanisms of cell transformation induced by PAX8-PPARγ are not fully understood. Some evidence has been presented for inhibition of normal PPARγ function via a dominant-negative effect of the PAX8-PPARγ protein on wild type PPARγ (55, 60). Other studies have

found the activation of known PPARγ target genes in tumors harboring PAX8-PPARγ, arguing against the dominant-negative effect (61). Other possible mechanisms include deregulation of PAX8 function, known to be critical for thyroid cell differentiation, and activation of a set of genes related to neither wild type PPARγ nor wild type PAX8 pathways (61, 62).

PAX8-PPARγ rearrangements and RAS point mutations rarely overlap in the same tumor, suggesting that follicular carcinomas may develop via at least two distinct molecular pathways, initiated by either PAX8-PPARγ or RAS mutation (57).

1.3 Other molecular events

Distinct alterations in gene expression have been observed in papillary carcinomas and other types of thyroid cancers (22, 63-66). These alterations include down regulation of genes responsible for specialized thyroid function (such as thyroid hormone synthesis), up regulation of many genes involved in cell adhesion, motility and cell-cell interaction, and different patterns of deregulation of the expression of genes that encode cytokines and other proteins involved in inflammation and immune response.

Among papillary carcinomas, different mRNA expression profiles have been observed in the classic papillary, follicular and tall-cell variants (64-68). Moreover significant correlations have been observed between BRAF, RAS, RET/PTC and TRK (tyrosine receptor kinases) mutations and specific patterns of gene expression. This information has shed light on the molecular basis for the distinct phenotypic and biological features associated with each mutation type (22, 63). Acquisition of more invasive tumor characteristics and dedifferentiation of BRAF-mutated cancers seems to coincide with profound deregulation of the expression of genes that encode proteins involved in cell adhesion and the inter-cellular junction, which provides evidence for induction of an epithelial-mesenchimal transition along with increased cell motility and invasiveness (69, 70).

Moreover, several miRNAs have been found to be deregulated in thyroid cancer (71-74). Generally miRNA expression profiles of papillary carcinoma are different from those of follicular carcinoma and other thyroid tumors (75). Several specific miRNAs, such as miR-

146b, miR-221 and miR-222, are highly up regulated in papillary carcinomas and many have a pathogenic role in the development of these tumors (72, 74, 75).

Alterations in gene expression owing to aberrant methylation of gene promoter regions or histone modification also occur in thyroid cancer. These epigenetic events can alter the function of tumor suppressor genes and thus contribute to activation of important signaling pathways, such as PI3K-AKT and MAPK cascade. Changes in epigenetic regulation might also result in down regulation of thyroid specific genes during tumor progression and dedifferentiation (76-78). Hypermethylation of the metalloproteinase inhibitor gene TIMP3 and other tumor suppressor genes is frequently observed in thyroid cancers with the BRAF V600E substitution, which may contribute to the aggressive biological behavior of tumors carrying this mutation (79).

1.4 Clinical management of thyroid nodules

The current standard of care in the management of a patient with a thyroid nodule is FNA biopsy with subsequent cytological determination of the suspicion for malignancy. Although FNA biopsy of thyroid nodule is very sensitive in the detection of malignancy, it is indeterminate or suspicious in 20-30% of cases (80, 81). Because clinicians often cannot determine malignancy, either pre- or intra-operatively, patients with suspicious thyroid lesions cannot be optimally managed (82). This often results in two scenarios:

- Patients who ultimately have a benign lesion on final histopathology may be subjected to unnecessary surgery;
- Patients with a malignant thyroid nodule may need to undergo a second operation for completion thyroidectomy only after a diagnosis of cancer is rendered on permanent histological section.

Therefore it is compelling to find further sources of information in order to define optimal surgical approach for patients with uncertain diagnoses pre- and intra-operatively.

1.4.1 Limitations of fine-needle aspiration cytology to distinguish benign from malignant thyroid neoplasm

Since its introduction over three decades ago, FNA cytology has reduced the number of diagnostic thyroidectomies being performed for benign thyroid nodules and increased the number of patients receiving complete initial surgical treatment for malignant thyroid neoplasms (83). There are four general categories of FNA cytological findings: benign (70%) (Fig. 2), malignant (≤10%), indeterminate or suspicious (10-20%), and non diagnostic (≤10%) (84).

Indeterminate results and non diagnostic yields are the main limitation of FNA cytology. Indeterminate FNA results occur because of the overlapping cytological features that are present in both benign and malignant thyroid nodules of follicular cell origin. Suspicious FNA results commonly occur because of focal nuclear atypia and insufficient sampling of enough follicular epithelium. In cases of follicular nuclear atypia, infrequent or sparse areas of nuclear enlargement, nuclear grooves, or intra nuclear inclusions are observed which are suggestive but not conclusive for papillary thyroid cancer. Although the diagnostic accuracy of FNA cytologies interpreted as benign or malignant is high, even in expert hands, interpretation of results can be inaccurate in up to 10% of cases (85, 86). Because clinicians have become reliant on FNA cytology, false negative FNA cytology results may delay treatment and adversely affect patient outcome (86). Moreover, several investigators have found significant discordance rates on secondary review of thyroid FNA cytology that affect the management of patients with thyroid neoplasm (85). Since FNA cytology cannot discriminate between malignant and benign thyroid nodules in up to 30% of cases, and because false negative and discordant FNA cytology result do occur, additional diagnostic tests that would improve the pre-operative accuracy of distinguishing benign from malignant thyroid neoplasm are necessary to improve patient outcome and reduce the cost of providing optimal care for patients with thyroid nodules.

1.4.2 Clinical utility of molecular markers

Molecular markers hold great promise in improving the diagnosis of cancer in patients with thyroid nodules. Such improvement would particularly benefit patients with nodules that are classified as indeterminate for malignancy by FNA cytology. The inability to rule out cancer in these nodule leads to diagnostic lobectomy for most of these patients, although 60-90% of the surgically removed thyroid nodules are found to be benign (87, 88). Those patients that are found to have cancer in their nodules after initial surgery are also treated sub optimally, as they have to undergo a second surgery to complete the thyroidectomy. Both the unnecessary surgeries and the two-step surgical management can be avoided with more accurate pre-operative diagnosis of cancer in thyroid nodules.

Diagnostic use of mutational markers for the analysis of thyroid FNA samples has been explored for single genes and for a panel of mutations. Among single genes, the majority of studies have focused on BRAF mutations. However, despite high specificity for cancer, testing for BRAF mutation alone misses many thyroid cancers that are negative for this mutation. The performance of molecular testing can be improved by including other frequently occurring mutations in the analysis. Use of a panel of mutations including BRAF and RAS point mutations and RET/PTC and PAX8/PPARγ rearrangements, with the possible addition of the TRK rearrangement, for analysis of thyroid FNA samples has been explored (89-92). Studies that evaluated the use of this panel in a setting of the clinical diagnostic laboratory demonstrated that finding any mutation was a strong predictor of malignancy in thyroid nodules irrespective of the cytological diagnosis (89-91). On the basis of the high probability of cancer in nodules positive for mutations, these patients with the possible exception of patients with RAS-positive nodules can be treated by total thyroidectomy as the initial surgical approach (Fig. 5).

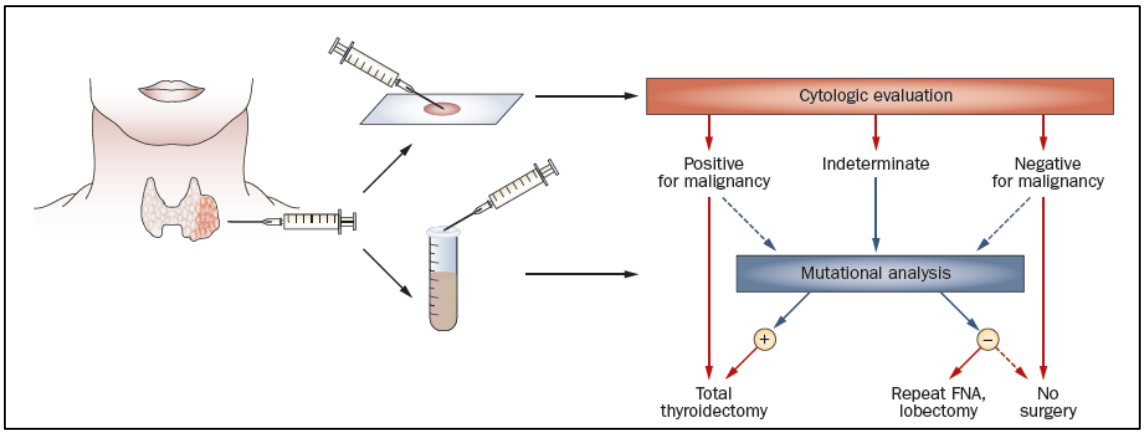

Figure 5. Potential clinical management of patients with thyroid nodules on the basis of a combination of cytological examination and molecular analysis. (Nikiforov, Y.E. & Nikiforova, M. N., Nature Reviews, 2011)

Thus molecular testing can be particularly helpful for nodules with indeterminate cytology. Nodules positive for mutations indicate a high risk of cancer; therefore, patients with these nodules can be treated by total thyroidectomy. Patients with nodules that yield an indeterminate diagnosis on cytology but are negative for mutations might require a repeated FNA and diagnostic lobectomy, although consideration of following up some of these patients annually may be given, particularly for patients with the cytological diagnosis of atypia of undetermined significance/follicular lesion of undetermined significance. Molecular testing of nodules found to be negative for malignancy by cytology decreases the rate of false-negative cytological results, but its cost-effectiveness has not been assessed. Molecular testing of samples classified as malignant by cytology can identify BRAF-positive tumors, which may require more extensive surgery that BRAF-negative tumors, although specific recommendations for surgical management of thyroid cancer based on the mutational status have not been developed yet.

In addition to gene mutations, gene expression profiling has also been explored for the diagnostic assessment of thyroid malignancy.

1.5 Gene expression profiling to distinguish benign from malignant thyroid neoplasm

The recent advances in molecular biology techniques have allowed the use of high throughput approaches for biomarker discovery (93). One such approach that has been used in thyroid cancer research by several groups is microarray analysis because it allows for the correlation of genes expression profile with clinical variables, the classification or definition of different tumor types, and the identification of genes or network of genes involved in carcinogenesis. Microarray studies in thyroid cancer have identified anywhere from 47 to 627 genes that are differentially expressed between benign and malignant thyroid neoplasms depending on the selection criteria and methods of data analysis (66, 94-96). Although some of the reported differentially expressed genes have an established role in thyroid tumorigenesis, the study design and data analysis of the microarray data have been different and thus have not resulted in a consistent group of candidate diagnostic gene markers. A few studies have validated the results of the microarray analysis by qPCR or immunohistochemistry (94, 97). This limitation is important given that some investigators have suggested that microarray analysis may incorrectly identify 30% of genes and gene expression levels (98). Another obstacle to the application of microarray analysis to distinguish benign from malignant thyroid neoplasms is the technical feasibility of using microarray analysis in FNA samples because of the limited amount of total RNA that can be extracted from FNA biopsy samples. Lubitz and associates (99) in 22 ex-vivo thyroid nodule FNA biopsy samples were able to demonstrate that microarray analysis is feasible by using a second complementary cDNA synthesis step and amplification method. They found that FNA biopsy samples clustered with the same tissue samples when using 25 differentially expressed genes identified from a training set.

Although gene expression profiling studies have identified many possible biomarkers with high accuracy, the clinical application of this approach to FNA samples is unclear and remains limited to the use of post-surgery samples.

Moreover, the expression of biomarkers, identified by microarray studies, needs to be validated by quantitative PCR (qPCR).

Here we selected several genes (KIT, SYNGR2, C21orf4, Hs.296031, DDI2, CDH1, LSM7, TC1, NATH), whose expression is known from literature (66, 100-109) to be deregulated in thyroid

malignancy, in order to build a qPCR computational model able to pre-operatively differentiate benign from malignant thyroid lesions.

The proto-oncogene KIT is a type III receptor tyrosine-kinase, cellular homologue of the viral oncogene of the feline sarcoma retrovirus HZ4-FeSV. It plays various roles in haematopoiesis, melanogenesis and spermatogenesis, and in the development of the interstitial cells of Cajal. Its ligand is the stem cell factor (SCF) (110, 111). The role of KIT in human neoplasia is not fully cleared yet. A number of tumor types are associated with activation of KIT through its over expression or through activating mutations (110, 112, 113), while in highly metastatic melanomas, breast cancer and thyroid carcinoma the progression into a malignant phenotype correlates mostly with loss of KIT expression (114, 115). Among the few papers studying KIT status in thyroid cancer, Natali *et al.* in 1995 (116) reported the loss of the receptor during the transformation of normal thyroid epithelium to papillary carcinoma. Similarly, in 2004 Mazzanti *et al.* (66), by using microarray assay, were able to identify out of thousand of genes, KIT as one of the most significant down expressed gene in PTC compared to benign lesions. Other laboratories confirmed this result by using qPCR (68, 117). Moreover, multiple miRNAs, predicted to target KIT, have been reported to be up regulated in PTC (72, 118). These findings indicate that KIT receptor may be involved in the growth control of thyroid epithelium and that this function may be lost in malignant transformation.

Besides KIT, we selected other 8 genes to study (SYNGR2, C21orf4, Hs.296031, DDI2, CDH1, LSM7, TC1, NATH). Among them, SYNGR2, C21orf4, Hs.296031, CDH1 and LSM7 were selected by microarray assay to have high diagnostic accuracy for distinguish thyroid nodules (66). TC1 and NATH were included in this study since their role in thyroid carcinogenesis has been shown in literature (100-102, 109).

SYNGR2 has been characterized as an integral vesicle membrane protein (114) and the only data available indicate its up regulation in fetal mouse ovaries (103). LSM7 has been described in the family of Sm-like proteins, involved in pre-messenger RNA splicing and decapping (104). The interaction of LSM7 with the TACC1 complex may participate in breast cancer oncogenesis (105). C21orf4 encodes a predicted transmembrane protein (Tmem50b) and is one of few genes significantly over-expressed during cerebellar development in a Down syndrome mouse model (106). The role of SYNGR2, LSM7 and C21orf4 in thyroid carcinogenesis has not yet been explored. E-Cadherin (CDH1) expression is reduced in thyroid

carcinomas (107) and its promoter resulted to be hypermethylated in thyroid neoplasm (108). Hs.24183 (now Hs.145049) has been identified as part of the 3'UTR of DDI2 (DNA-damage inducible 1 homolog 2) gene in H. sapiens, but no data exists about its role in thyroid. For Hs.296031 the only information available is about gene sequence and mapping, but no gene and protein function are known. In contrast, the expression of the thyroid cancer-1 (TC1) gene resulted to be related to malignant transformation in thyroid and the potential use of TC1 gene expression as a marker of malignancy in thyroid nodules is also shown in literature (109). NATH (N-acetyl transferase human) is involved in protein acetylation that represents an important post-translational modification regulating oncogenesis, apoptosis and cell cycle. NATH resulted to be over-expressed at the mRNA level in papillary thyroid carcinomas relative to non neoplastic thyroid tissue (100).

1.6 Bayesian Neural Networks: clinical utility

Several attempts to use Bayesian Neural Networks in the clinical setting are described in literature (119-121), more specifically Liu and colleagues (121) have shown the clinical utility of a Bayesian network for differentiating benign from malignant thyroid nodules using sonographic and demographic features.

The procedure uses a Probabilistic Neural Network (PNN) to classify cases into malignant and benign categories, based on the input variables by implementing a nonparametric method for classifying observations into one of the output groups based on the observed variables. Rather than making any assumption about the nature of the distribution of the variables within each group, it constructs a nonparametric estimate of each group's density function at a desired location based on neighbouring observations from that group. Observations are assigned to groups based on the product of three factors:

- The estimated density function in the neighbourhood of the point
- The prior probabilities of belonging to each group
- The costs of misclassifying cases that belong to a given group

The approach to classifying cases can be formulated as a neural network. The basic setup of the network consists of four layers:

- An input layer
- A pattern layer
- A summation layer
- An output layer, also having one binary neuron for each output class that turns on or off depending on whether or not a case is assigned to the corresponding group

Conceptually, the input layer provides the information from the predictor variables by feeding their values (standardized by subtracting the mean and dividing by the standard deviation) to the neurons in the pattern layer. The pattern layer passes the values through an activation function, which uses the input values to estimate the probability density function for each group at a given location. The density estimates are then passed to the summation layer, which combines the information from the training cases with prior probabilities and misclassification costs to derive a score for each group. The scores are then used to turn on the binary neuron in the output layer corresponding to the group with the largest score and turn off all other output neurons.

2. AIM

The aim of the current study was to find a molecular approach able to pre-operatively diagnose benign and malignant thyroid tumors. Towards this goal, we evaluated KIT expression as a pre-operative diagnostic biomarker. We also assessed the expression profile of the following genes: KIT, SYNGR2, C21orf4, Hs.296031, Hs.24183, CDH1, LSM7, TC1 and NATH in order to build a qPCR-based computational model able to improve the pre-operative diagnostic accuracy along with assessing BRAF mutational status.

Moreover, because of the lack of useful pre-operative diagnostic biomarkers, we sought to determine whether the expression profile of these genes on FNA cytological smears, including old archived samples, could be performed on a routine basis so as it improve the diagnostic sensitivity for malignancy in thyroid nodules read as indeterminate or suspicious without adding time and discomfort for the patient of the FNA procedure.

As in other diseases, molecular pathology is playing a relevant role in diagnosis of thyroid cancer. Very recent papers of our laboratory have proposed a new simple method, named manual macrodissection, to perform molecular analysis on cells obtained by FNAC, and have demonstrated the usefulness of the association cytology-molecular biology for papillary thyroid carcinoma and micro-papillary thyroid carcinoma diagnosis. In this study we evaluated KIT expression in a morpho-molecular diagnostic approach to a series of thyroid FNAC, together with the study of BRAF gene mutational status. We further assessed the diagnostic role of KIT expression in thyroid FNAC.

In the last years a new class of techniques known as Bayesian Neural Networks (BNN) have been proposed as a supplement or alternative to standard statistical techniques. For the purpose of predicting medical outcomes, a BNN can be considered a computer intensive classification method and, in addition, BNNs do not require explicit distributional assumption (such as normality).

Here we used 87 FNA cytological samples to build several pre-operative computational models and 6 unknown samples to be tested in order to find the most discriminative one and to evaluate the diagnostic utility in the pre-operative management of thyroid nodules.

We then aimed to assess a potential correlation of the markers in order to investigate their biological importance and to find a link that could give us a better understanding of the molecular mechanisms underlying thyroid cancer development.

3. MATERIALS AND METHODS

3.1 Thyroid specimens

Pre-operative thyroid FNA slides of 93 patients (49 malignant, 38 benign, 6 unknown) collected over 10 years (2000 to 2010) were selected from archived materials in the Section of Cytopathology, Division of Surgical, Molecular and Ultrastructural Pathology. All patients had a thyroidectomy with histopathological examination based on clinical elements or of a cytological diagnosis of malignancy, suspected malignancy, or if indeterminate. For ethical reasons, we used only cases with two or more slides per patient, and the molecular analysis was performed on only one of the available smears. In all cases, the FNA was performed using ultrasonographic guidance.

3.2 Ethical Board

This study was approved by the Internal Review Board (IRB) of the University of Pisa. All patients gave their consent for the participation to the study.

3.3 FNA slides and DNA extraction

The slides were obtained from FNA samples and fixed in ethyl alcohol for Papanicolau staining. Smears were reviewed by a senior cytopathologist. All archival FNA slides were kept in xylene for 1 to 3 days, depending on the time of storage, to detach the slide coverslips. They were then hydrated in a graded series of ethanol, followed by a wash in dH_2O for 1 minute. The slides were finally air-dried. The processing of the slides was performed in a range of few days to a maximum of 10 years after FNA procedure.

DNA extraction was performed using a commercial kit (Nucleospin; Macherey-Nagel, Düren, Germany) mainly following manufacturer's instruction. A modification was added to the first step: 50% of the lysis solution with no Proteinase K was initially poured on the slides to scrape off the cytological stained sample using a single-edged razor blade. Any scraped

tissue was then collected in a microcentrifuge tube containing the other half of the lysis solution with the Proteinase K. The extracted DNA was kept at -20 °C until used.

3.4 BRAF V600E detection

BRAF exon 15 was analyzed by polymerase chain reaction (PCR) followed by direct sequencing. PCR was performed in a 30 µl final volume, containing 2 µl of DNA, 0.05 mM dNTP (Invitrogen, Carlsbad, CA), 2.5ng/µl of each primer (Invitrogen), 1.5 mM $MgCl_2$, 1x PCR Gold Buffer, and 0.75U AmpliTaq Gold (Applied Byosistems, Foster City, CA). PCRs were performed on a 9700 GenAmp PCR System (Applera Corporation, Foster City, CA) with the following cycling conditions at 94 °C for 7 minutes; 40 cycles at 94 °C for 45 seconds, 56 °C for 45 seconds, and 72 °C for 1 minute; and final step at 72 °C for 10 minutes. Primers were:

BRAF 15 F: 5'-TCATAATGCTTGCTCTGATAGGA-3'
BRAF 15 R: 5'-GGCCAAAAATTTAATCAGTGGA-3'

PCR reactions were run on agarose gel to check the presence of the specific amplification products. PCR bands were cut and purified using the Genelute Gel Extraction Kit (St. Louis, MO). Purified products were then sequenced on the ABI PRISM 3100 Genetic Analyzer (Applied Biosystem).

3.5 KIT genotyping

KIT sequence was screened for mutations in exons 9, 11, 13 and 17 by direct sequencing. PCR was performed using standard conditions: initial denaturation 95°C for 7 min; 40 cycles at 95 °C for 45 sec and 56 °C for 45 sec and 72 °C for 45 sec; final step 72 °C for 10 min with AmpliTaq Gold (Applied Biosystems) on 9700 GeneAmp PCR System (Applied Biosystems). Primers for KIT sequencing were selected using Primer3 software:

- exon 9 F: 5'- CCAGGGCTTTTGTTTTCTTC - 3'
- exon 9 R: 5'- TGGTAGACAGAGCCTAAACATCC - 3'

- exon 11 F: 5’- GATCTATTTTTCCCTTTCTC - 3’
- exon 11 R: 5’- AGCCCCTGTTTCATACTGAC - 3’
- exon 13 F: 5’- TCAGTTTGCCAGTTGTGCTT - 3’
- exon 13 R: 5’- AATGTCATGTTTTGATAACCT - 3’
- exon 17 F: 5’- TTCTTTTCTCCTCCAACCTAA - 3’
- exon 17 R: 5’- TGTCAAGCAGAGAATGGGTA - 3’

The PCR products were purified with Multi Screen PCR Plates (Millipore) and the sequencing reactions were performed in 20 µl final volume using Big Dye Terminator kit v3.1 (Applied Biosystems) and 2.5 pmol/µl of each primer, and then purified with Multi Screen PCR Plates (Millipore). The sequence reactions were loaded on ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) and analyzed using the Sequencing Analysis software 3.4 version.

3.6 RNA extraction and cDNA synthesis

RNA extraction was performed by using a commercial kit (High Pure RNA Paraffin kit, Roche) mainly following the manufacturer’s instructions and adding the same modification step as for DNA extraction. The lysis solution was poured on the slide to scrape off the cytological stained sample by using a single edged razor blade. Whole scraped material was then collected in a microcentrifuge tube and processed for RNA extraction. The quantity/quality of extracted RNA was estimated with Nanodrop 1000 spectrophotometer by using 1 µl of undiluted RNA solution. RNA was treated with DNase I recombinant, RNase-free (Roche). RNA was reverse transcribed in a final volume of 20 µl, containing 5X RT buffer, 10 mM dNTPs, 50 ng/µl Random Primers, 0.1M DTT, 40 U/µl RNaseOUT, 50 µM oligo(dT), DEPC-Treated Water, 15 U/µl Cloned AMV reverse transcriptase (Invitrogen, Carlsbad, CA).

3.7 Quantitative Real-Time PCR (qPCR)

The level of KIT, SYNGR2, C21orf4, Hs.296031, DDI2, CDH1, LSM7, TC1, NATH expression was analyzed by quantitative Real-Time PCR (qPCR) on the Rotor-Gene 6000 real time rotary analyzer (Corbett, Life Science, Australia) following the manufacturing instructions.

Endogenous reference gene (B2M, beta 2 microglobulin) was used to normalize each gene expression level. PCR products were previously sequenced on an Applied Biosystems 3130xl Genetic Analyzer (Foster City, CA) to confirm gene sequence. PCR was performed in 25 µl final volume, containing 5 µl of cDNA, 12.5 µl of MESA GREEN qPCR MasterMix Plus (EUROGENTEC, San Diego, CA), 40 pmol of each primer (Invitrogen, Carlsbad, CA) *per* reaction with the following cycling conditions: initial denaturation 95°C for 5 min; 40 cycles at 95 °C for 15 sec and 61 °C for 40 sec and 72 °C for 40 sec; final step 25 °C for 1 min. Primers were selected using Primer3 software:

KIT F: 5'- GCACCTGCTGCTGAAATGTATGACATAAT - 3'
KIT R: 5'- TTTGCTAAGTTGGAGTAAATATGATTGG - 3'

SYNGR2 F: 5'- ATCTTCTCCTGGGGTGTGCT - 3'
SYNGR2 R: 5'- AGGGTGGCTGTTGGTAGTTG - 3'

C21orf4 F: 5'- GACAACAGTGGCTGTGTTTTAAG - 3'
C21orf4 R: 5'- GCATTGGATACAGCATTTATCAT - 3'

Hs.296031 F: 5'- TGCCAAGGAGCTTTATAGAA - 3'
Hs.296031 R: 5'- ATGACGGCATGTACCAACCA - 3'

DDI2 F: 5'- TGCAGTTCCCAAACTTACCC- 3'
DDI2 R: 5'- CAGCAACATATCTCGGAGCA- 3'

CDH1 F: 5'- GCATTGCCACATACACTCTC- 3'
CDH1 R: 5'- AGCACCTTCCATGACAGAC- 3'

LSM7 F: 5'-GACGATCCGGGTAAAGTTCCA - 3'
LSM7 R: 5'- AGGTTGAGGAGTGGGTCGAA - 3'

TC1 F: 5'- AAATCTTCTGACTAATGCTAAAACG - 3'
TC1 R: 5'- TTATTGTTGCATGACATTTGC - 3'

NATH F: 5’-AAGAAACCAAAGGGGAACTT - 3’

NATH R: 5’- TAATAGGCCCAGTTTTCAGG - 3’

B2M F: 5’- CATTCCTGAAGCTGACAGCATTC - 3’

B2M R: 5’- TGCTGGATGACGTGAGTAAACC - 3’

A first PCR run was performed on control sample expressing the markers and run on 2% agarose gel. The PCR product was excised from the gel, purified by using GenElute™ PCR Clean-Up (Sigma-Aldrich) and measured spectrophotometrically at 260 and 280 nm. The purified product was diluted in a 10-fold series to create the standards for a ten-point standard curve that was run in triplicate. Standard curves were generated for both KIT and B2M and showed a good linearity with consistent correlation coefficient ($R^2 = 0.999$). Ct was determined by the Rotor-Gene 6000 software and exported for analysis after background subtraction. Threshold was set by standard curve and then imported in all the runs for data analysis. PCR efficiencies resulted similar for the marker genes and B2M in each experiment and ranged between 98-102%. The experiment was run in duplicate for each sample.

For each cDNA sample the ratio between the gene of interest expression value and B2M expression value was calculated. The expression ratio mean values and standard deviations of malignant and benign groups were calculated. To verify primers specificities, melting curve analysis was performed. Fluorescent data were acquired during the extension phase. After 40 cycles a melting curve for each gene was generated by slowly increasing (0.1°C/s) the temperature from 60°C to 95°C, while the fluorescence was measured. For each experiment a no-template reaction was included as a negative control. The expression of all the markers was ultimately represented as the ratio of absolute quantification by standard curve of the expression of the markers and B2M expression.

3.8 BRAF genotyping and gene expression data of 112 melanoma metastases

DNA and total RNA was isolated from 112 metastatic tumor samples from as many patients treated at the Surgery Branch, National Institutes of Health, Bethesda, MD. Tumor samples were from snap frozen biopsies.

BRAF sequencing was performed as described above. For gene expression analysis, cDNAs were fragmented, biotinylated, and hybridized to the GeneChip Human Gene 1.0 ST Arrays (Affymetrix WT Terminal Labeling Kit). Probe normalization, background correction, Log2 transformation and summarization were performed using Robust Multi-Chip Average (RMA). Gene summary was obtained by averaging the mean of the probe values end expressed as Log2 intensity (122).

TC1 mRNA expression values (Log2 transformed) were extrapolated and evaluated according BRAF mutational status.

3.9 Statistical analyses

3.9.1 *Gene expression analysis*

Mann-Withney test and Student t-test were used to determine differences between mRNA expression levels of KIT, LSM7, C21orf4, DDI2, SYNGR2, TC1, Hs.296031 and CDH1, NATH, respectively. Also, the TC1 expression level in BRAF V600E and wild type malignant samples was tested by Student t-test. All the analyses were performed by using StatGraphics Centurion (V. 15, StatPoint, Inc.).

3.9.2 ROC analysis

To determine the diagnostic accuracy of the molecular computational model, we calculated the area under the curve (AUC) of the receiver operating characteristic (ROC) curve for each gene individually and in combination by using logistic regression analysis (Medcalc 11, Medcalc Software, Stata Software).

In this analysis the true positive rate (Sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points. Each point on the ROC plot represents a sensitivity/specificity pair corresponding to a particular decision threshold. A test with perfect discrimination (no overlap in the two distributions) has a ROC plot that passes through the upper left corner (100% sensitivity, 100% specificity). Therefore the closer the ROC plot is to the upper left corner, the higher the overall accuracy of the test (123).

3.9.3 BNN classifier

Several computational models (Neural Network Bayesian Classifiers) were built in order to find the best combination of markers able to discriminate benign from malignant thyroid samples. This procedure uses a Probabilistic Neural Network (PNN) to classify cases into malignant and benign categories, based on 10 input variables (KIT, LSM7, C21orf4, DDI2, SYNGR2, TC-1, Hs.296031, CDH1, NATH expressions and BRAF mutational status) by implementing a nonparametric method for classifying observations into one of benign and malignant groups based on the observed expression variables.

3.9.4 Molecular diagnostic accuracy

Fisher's test was used to compare samples correctly classified by the BNN model according to their probability score (> 90% and <90%). Diagnostic accuracy gain was then calculated after applying molecular tests (BRAF, KIT and BNN model).

3.9.5 *Correlation analysis*

In order to evaluate the biological importance of the markers analyzed, a multiple-variable analysis was performed to analyze the correlation between the markers.
These analyses were all performed by using Statgraphics Centurion (V. 15, StatPoint, Inc.).

4. RESULTS

4.1 Expression and genotyping of KIT receptor in benign and malignant thyroid lesions

KIT expression was analyzed by qPCR in the first set of 82 FNAC, histologically diagnosed as 36 benign and 46 malignant thyroid nodules (Table 1). Overall, KIT expression was detected in 59% of PTC (27/46) and in 100% of BN (36/36). The mean of KIT expression values (KIT/B2M ratio) was calculated for both benign (1.72) and malignant (0.138) groups and the difference resulted highly significant ($p<0.0001$).

The sequencing of exons 9, 11, 13, 17 of the KIT gene resulted wild type for all the samples analyzed.

Table 1. Histological and cytological diagnoses of 82 thyroid nodules

Histological Diagnosis	***Cytological Diagnosis***		
	PTC	**SPTC**	**IFP**
PTC: 46 cases	**30 (65%)**	**11 (24%)**	**5 (11%)**
	BN	**IFP**	
BN: 36 cases	**17 (47%)**	**19 (53%)**	

PTC: papillary thyroid carcinoma. SPTC: suspicious for PTC. IFP: indeterminate follicular proliferation. BN: benign nodule.

4.2 KIT expression and biological behaviour of thyroid nodules

The value of KIT expression (KIT/B2M ratio) ranged between 0 and 9.34 (Table 2). To evaluate a possible relationship with the biological behaviour of lesions, KIT expression values (ev) were arbitrarily organized in four classes (Table 3):

- Class I: KIT ev = 0;
- Class II: KIT ev > 0 and ≤ 0.5;

- Class III: KIT ev > 0.5 and ≤ 3;
- Class IV: KIT ev > 3;

The percentage of malignant and of benign cases was calculated in each class and its statistical significance was determined (p-value).

Table 2. Morphological and molecular diagnosis in 82 thyroid nodules

case	HD	CD	KIT Class	KIT ev	BRAF
1	PTC	PTC	I	0	V600E
2	PTC	PTC	I	0	V600E
3	PTC	PTC	I	0	V600E
4	PTC	PTC	I	0	V600E
5	PTC	PTC	I	0	V600E
6	PTC	PTC	I	0	V600E
7	PTC	PTC	I	0	V600E
8	PTC	PTC	I	0	V600E
9	PTC	PTC	I	0	V600E
10	PTC	PTC	I	0	V600E
11	PTC	PTC	I	0	V600E
12	PTC	PTC	I	0	V600E
13	PTC	PTC	I	0	WT
14	PTC	PTC	I	0	WT
15	PTC	PTC	I	0	WT
16	PTC	PTC	II	0.5	V600E
17	PTC	PTC	II	0.448	V600E
18	PTC	PTC	II	0.105	V600E
19	PTC	PTC	II	0.07	V600E
20	PTC	PTC	II	0.0533	V600E
21	PTC	PTC	II	0.049	V600E
22	PTC	PTC	II	0.031	V600E
23	PTC	PTC	II	0.022	WT
24	PTC	PTC	II	0.013	V600E
25	PTC	PTC	II	0.0126	WT
26	PTC	PTC	II	0.01	WT
27	PTC	PTC	II	0.004	WT
28	PTC	PTC	II	0.0003	WT
29	PTC	PTC	II	0.0002	V600E
30	PTC	PTC	III	1.0506	WT
31	PTC	SPTC	I	0	WT
32	PTC	SPTC	I	0	WT
33	PTC	SPTC	I	0	WT
34	PTC	SPTC	I	0	WT
35	PTC	SPTC	II	0.44	V600E
36	PTC	SPTC	II	0.42	V600E
37	PTC	SPTC	II	0.022	V600E
38	PTC	SPTC	II	0.011	V600E
39	PTC	SPTC	II	0.42	WT
40	PTC	SPTC	II	0.37	WT
41	PTC	SPTC	III	1.27	WT
42	PTC	IFP	II	0.275	WT
43	PTC	IFP	II	0.097	WT
44	PTC	IFP	II	0.07	WT

45	PTC	IFP	II	0.023	WT
46	PTC	IFP	III	0.57	WT
47	BN	BN	II	0.3	WT
48	BN	BN	II	0.14	WT
49	BN	BN	II	0.07	WT
50	BN	BN	II	0.1079	WT
51	BN	BN	II	0.363	WT
52	BN	BN	III	0.561	WT
53	BN	BN	III	0.6457	WT
54	BN	BN	III	0.858	WT
55	BN	BN	III	0.96	WT
56	BN	BN	III	0.97	WT
57	BN	BN	III	1.051	WT
58	BN	BN	III	1.3823	WT
59	BN	BN	III	1.39	WT
60	BN	BN	III	2.47	WT
61	BN	BN	III	2.5083	WT
62	BN	BN	III	2.6333	WT
63	BN	BN	IV	7.24	WT
64	BN	IFP	II	0.032	WT
65	BN	IFP	II	0.04	WT
66	BN	IFP	II	0.0463	WT
67	BN	IFP	II	0.0599	WT
68	BN	IFP	II	0.1143	WT
69	BN	IFP	II	0.1957	WT
70	BN	IFP	II	0.217	WT
71	BN	IFP	II	0.2619	WT
72	BN	IFP	III	0.7302	WT
73	BN	IFP	III	0.84	WT
74	BN	IFP	III	0.948	WT
75	BN	IFP	III	1.09	WT
76	BN	IFP	III	1.3727	WT
77	BN	IFP	III	1.64	WT
78	BN	IFP	III	1.67	WT
79	BN	IFP	IV	3.4	WT
80	BN	IFP	IV	7.69	WT
81	BN	IFP	IV	8.73	WT
82	BN	IFP	IV	9.34	WT

HD: histological diagnosis. CD: cytological diagnosis. KIT ev: KIT expression value. PTC: papillary thyroid carcinoma. SPTC: suspicious for PTC. BN: benign nodule. IFP: indeterminate follicular proliferation. WT: wild type.

In class I the percentage of malignancy is 100% (19/19 cases), whereas in class IV the percentage of benignity is 100% (5/5). In class II the percentage of malignant cases is higher than benign cases: 65% (24/37) vs. 35% (13/37). On the other hand, class III has a higher percentage of benign cases than malignant ones: 86% (18/21) vs. 14% (3/21). Difference between malignant and benign lesions is statistically highly significant in classes I, IV and III ($p<0.0001$, $p=0.0091$, $p<0.0001$). P-value in class II is 0.14.

Table 3. Classes of KIT expression value

Class	KIT ev	PTC		BN		p value
		n	%	n	%	
I	0	19	100	0	0	< 0.0001
II	> 0 - ≤ 0.5	24	65	13	35	= 0.1400
III	> 0.5 - ≤ 3.0	3	14	18	86	< 0.0001
IV	> 3	0	0	5	100	= 0.0091
		46		36		

KIT ev: KIT expression value. PTC: papillary thyroid carcinoma. BN: benign nodule.

Figure 6 reports the results of fitting a logistic regression model. The p-value of the diagnostic model is less than 0.05, showing a statistically significant relationship between the variables at the 95% confidence level. Specificity and sensitivity of the diagnostic performance of the model were evaluated by ROC analysis and the AUC was 0.881, with C.I. 95% 0.79-0.94 and $p = 0.001$, indicating that the model has a statistically significant efficacy in discriminating malignant from benign lesions (Fig. 7).

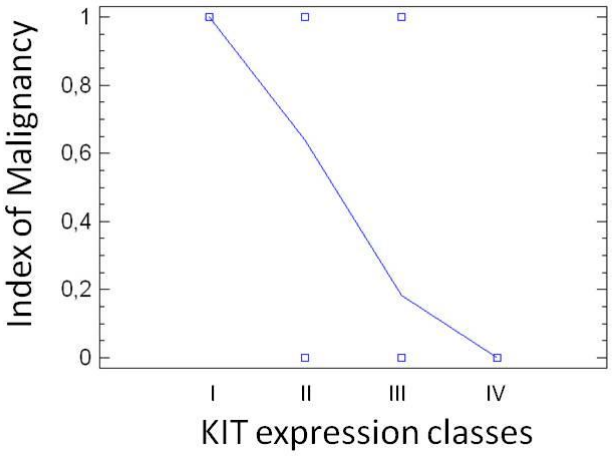

Figure 6. Logistic regression. Logistic regression model describing the relationship between the risk of malignancy and the classes of KIT expression. The p-value of the model resulted to be less than 0.05

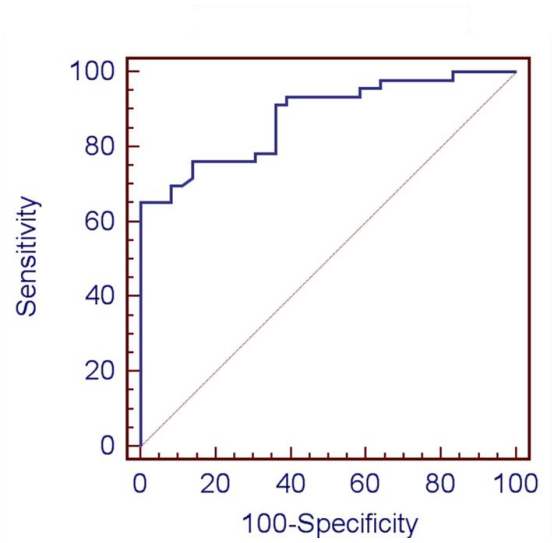

Figure 7. ROC analysis for KIT expression. The true positive rate (Sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points. Each point on the ROC plot represents a sensitivity/specificity pair corresponding to a particular decision threshold. (AUC=0.9, C.I 95% 0.79-.94; **p=0.001**)

We also evaluated the predicted malignancy probability of samples belonging to any of the KIT expression classes (Fig. 8).

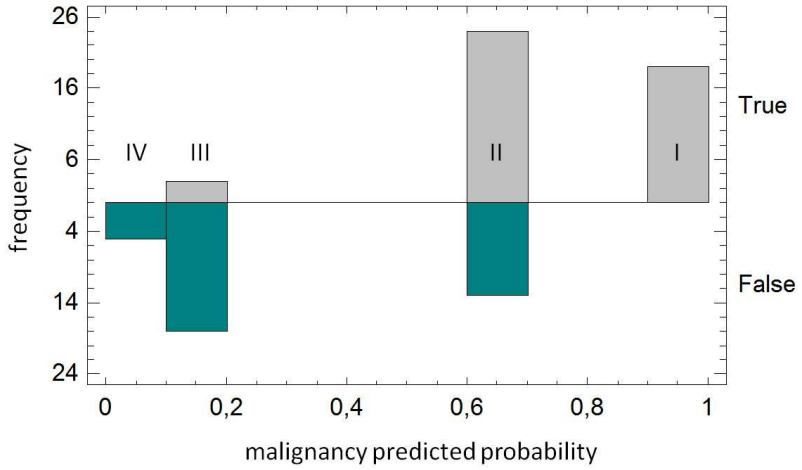

Figure 8. The plot describes the malignancy predicted probability of samples falling in any of the four classes of KIT expression (histograms). The prediction of malignancy is 100% true in class I and 100% false in class IV, where all samples are actually malignant and benign lesions respectively. In class II the prediction is mostly in favor of a malignant status while in class III most of the predictions of malignancy are false

4.3 KIT expression in FNAC

The cytological diagnoses of the 46 histologically confirmed PTCs were distributed through the 4 classes as following (Table 4):

- 30 cases cytologically diagnosed as PTC: 15 (50%) were in class I, 14 (47%) in class II, 1 (3%) in class III and none (0%) in class IV.
- 11 cases cytologically diagnosed as SPTC: 4 (36%) in class I, 6 (55%) in class II, 1 (9%) in class III and none (0%) in class 4.
- 5 cases cytologically diagnosed as IFP: 4 (80%) were in class II, 1 (20%) in class III, whereas no case was present in class I and IV.

The cytological diagnoses of the 36 histologically confirmed BN were distributed through the 4 classes as following:

- 17 cases cytologically diagnosed as benign: 5 (29%) were in class II, 11 (65%) in class III, 1 (6%) in class IV, whereas no case was present in class I.
- 19 cases cytologically diagnosed as IFP: 8 (42%) were in class II, 7 (37%) in class III and 4 (21%) in class IV, whereas no case was present in class I.

Table 4. Distribution of cytological diagnosis in the KIT expression classes

HD	CD	class I		class II		class III		class IV	
		n	%	n	%	n	%	n	%
PTC	PTC: 30	15	50	14	47	1	3	0	0
	SPTC: 11	4	36	6	55	1	9	0	0
	IFP: 5	0	0	4	80	1	20	0	0
total	46	19		24		3		0	
BN	BN: 17	0	0	5	29	11	65	1	6
	IFP: 19	0	0	8	42	7	37	4	21
total	36			13		18		5	

HD: histological diagnosis. CD: cytological diagnosis. PTC: papillary thyroid carcinoma. SPTC: suspicious for PTC. IFP: indeterminate follicular proliferation. BN: benign nodule. n: number of cases

4.4 BRAF V600E genotyping and KIT expression values of cytological samples

The BRAF V600E mutation (Table 2) was found in 54% of the 46 malignant samples (25/46). 48% BRAF V600E mutated samples were in KIT class I (12/25), as shown in Table 5. Class II contained the residual 52% of BRAF mutated cases (13/25). Class III and IV had no BRAF mutated cases. BRAF V600E was significantly more present in class I and II ($p = 0.00026$). No benign samples were BRAF mutated.

4.5 KIT/BRAF combined molecular analysis in thyroid nodule FNAC

Table 5 shows that four cases of SPTC are in KIT class I, whereas four more cases harbor a BRAF V600E mutation. All these 8 cases can be reasonably considered as PTC. At the same time, the four cases with a cytological diagnosis of IFP that are in KIT class IV can reasonably considered as benign nodules.

Table 5. BRAF mutational status according to KIT expression classes and to morphological diagnosis

HD	CD	class I		class II		class III		class IV	
		V600E	WT	V600E	WT	V600E	WT	V600E	WT
PTC: 46	PTC: 30	**12**	3	**9**	5	0	1	0	0
	SPTC: 11	0	4	**4**	2	0	1	0	0
	IFP: 5	0	0	0	4	0	1	0	0
BN: 36	BN: 17	0	0	0	5	0	11	0	1
	IFP: 19	0	0	0	8	0	7	0	4

HD: histological diagnosis. CD: cytological diagnosis. WT: wild type. PTC: papillary thyroid carcinoma. SPTC: suspicious for PTC. IFP: indeterminate follicular proliferation.

4.6 Role of molecular diagnosis (BRAF/KIT) in increasing the diagnostic accuracy of FNAC

As shown in Table 6, if the 8 cases of SPTC in KIT class I or hosting a BRAF V600E mutation are moved to the diagnostic group of PTC, the total number of PTC rises from 30 (65%) to 38 (83%), with an advantage in diagnostic accuracy of malignancy of 18%.

On the other hand, if the 4 cases of IFP in KIT class IV are moved to the diagnostic group of BN, the total number of BN rises from 17 (47%) to 21 (58%), with an advantage in diagnostic accuracy of benignity of 11%.

Finally, if we consider both PTC diagnosis and BN diagnosis, the whole diagnostic accuracy gain is of 15% with a statistically significant p-value of 0.03.

Table 6. Role of molecular diagnosis in increasing the diagnostic accuracy of FNAC

CD	**KIT class I**	**BRAF V600E**	**KIT class IV**
SPTC	4	4	
IFP			4

	CD	MD	DA
PTC	30/46: 65%	38/46: 83%	+ 18%
BN	17/36: 47%	21/36: 58%	+ 11%
PTC + BN	47/82: 57%	59/82: 72%	+ 15%

CD: cytological diagnosis. PTC: papillary thyroid carcinoma. SPTC: suspicious for PTC. BN: benign nodule. MD: molecular diagnosis. DA: diagnostic accuracy

4.7 Gene expression levels of the 9 markers

After studying KIT gene expression in the first set of samples, few samples were removed from the cohort because of the paucity of biological material. We then added other samples and the final cohort comprised a total of 93 patients (49 malignant, 38 benign, 6 unknown). We tested the gene expression of KIT, LSM7, C21orf4, DDI2, SYNGR2, TC-1, Hs.296031, CDH1, and NATH on the new set of samples. We found KIT, CDH1, LSM7, C21orf4, DDI2 mRNA expression levels significantly higher in benign thyroid tumors compared to the malignant ones, $p(KIT)<0.0001$; $p(CDH1)=0.004$; $p(LSM7)=0.03$; $p(C21orf4)=0.01$; $p(DDI2)=0.0001$. However, there was no significant difference in NATH, SYNGR2, TC1, Hs.296031 mRNA expression. Among these markers, only TC1 was over-expressed in malignant lesions compared to the benign ones (Fig. 9A).

4.8 BRAF status

Among the 49 malignant samples, 28 carried the V600E mutation on BRAF exon 15. The diagnostic accuracy of BRAF mutational status was calculated as the ratio between the number of patients carrying V600E mutation over the total of malignant samples (28/49=57%). All the benign samples were wild type.

4.9 ROC analyses

We employed receiver-operated characteristics (ROC) curve analyses to determine model robustness for predicting malignancy in thyroid samples using the expression of each gene individually (Fig. 9B, Table 7). Among the markers, KIT showed the highest AUC (0.9) meaning that it is the most powerful marker. We also performed a ROC analysis for the statistical significant markers (KIT, CDH1, LSM7, C21orf4, DDI2) and BRAF status in combination, the AUC resulted to be 0.8824, the sensitivity 91% and specificity 63% (Fig. 9C). Although the AUC resulted quite similar to the KIT one, the predictive power increases when the markers were combined together.

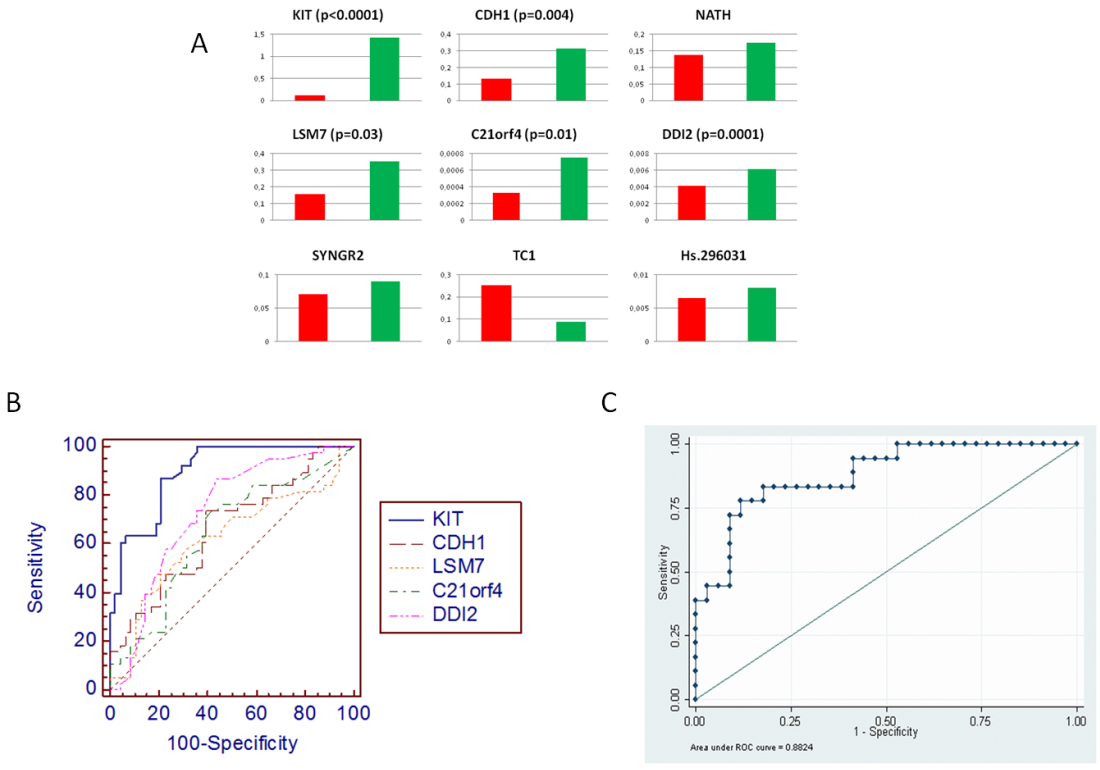

Figure 9. Expression means of 49 malignant (red) and 38 benign (green) samples for each marker (A). ROC analysis for KIT, CDH1, LSM7, C21orf4, DDI2 separately. Among the markers, KIT resulted to be the most powerful in discriminating benign from malignant thyroid tumors (AUC=0.9) (B). ROC analysis for KIT, CDH1, LSM7, C21orf4, DDI2, and BRAF status in combination (AUC=0.88) (C).

Table 7. ROC analysis for each marker individually

	Sensitivity	Specificity	AUC[a]	SE[b]	95% CI[c]	p value
KIT*	79.6	86.8	0,900	0,0313	0,817-0.954	0,0001
CDH1*	61.2	73.7	0.700	0,0586	0,559-0.766	0,0041
NATH	57.8	57.9	0,553	0,0658	0,440-0.662	0,4209
LSM7*	69.4	57.9	0.625	0.0633	0,515-0.727	0,0477
C21orf4*	58.3	73.7	0.644	0.0607	0,533-0.744	0,0180
DDI2*	56.2	86.8	0.729	0.0551	0.622-0.819	0,0001
SYNGR2	47.9	78.9	0.608	0.0613	0.497-0.712	0.0780
TC1	85.0	38.2	0,581	0,0679	0,460-0,695	0,2336
Hs.296031	77.8	32.4	0,490	0,0671	0,375-0,605	0,8761

[a]AUC (area under the curve), [b]SE (standard error), [c]CI (confidence interval), *p<0.05

4.10 Principal Component Analysis and Clustering

We then performed Principal Component Analysis in order to visualize the discriminative power of all the markers according to malignant and benign status and to BRAF mutational status (Fig. 10).

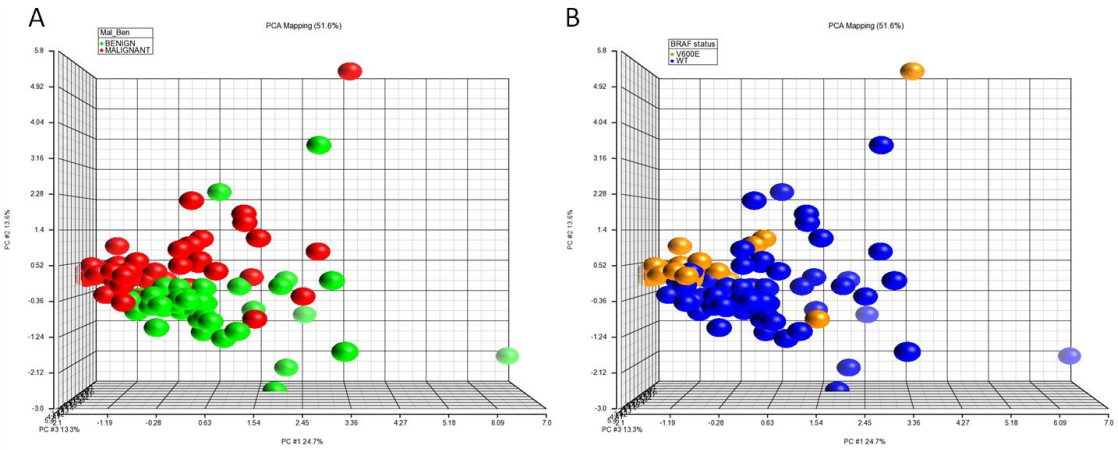

Figure 10. Principal Component analysis referring to malignant and benign status (A) and BRAF mutational status (B)

Overall, a separation between malignant and benign samples as well as between BRAF V600E and wild type samples was found, though not absolute.

A similar finding was observed when clustering samples according to the 9 expression markers (Fig. 11A).

We then sought to determine whether a differential expression profile of the 9 markers was present within the malignant group according to BRAF mutational status. Quite interestingly, malignant tumors bearing the V600E mutation were observed to cluster separately from wild type samples expressing the markers in a higher extent (Fig. 11B), and, more importantly, the behaviour of the malignant wild type samples recapitulated the one of benign samples.

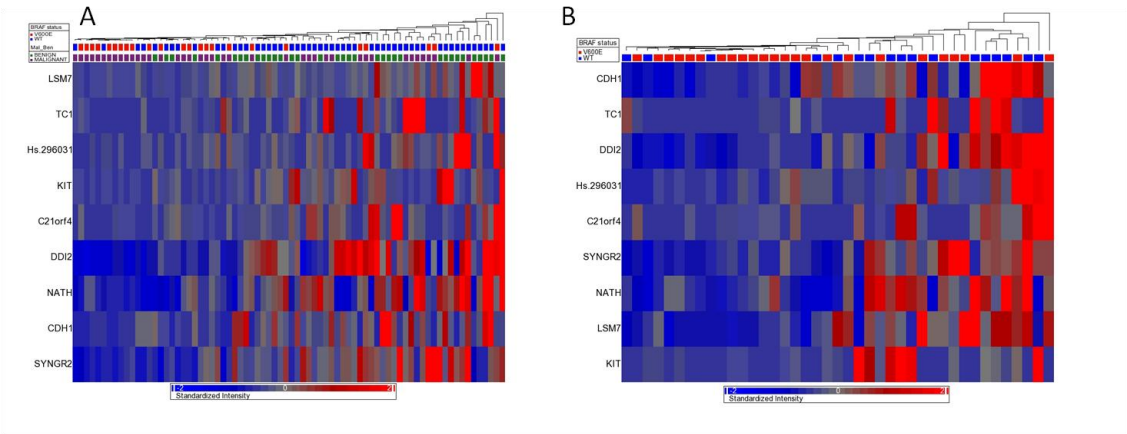

Figure 11. (A) Self-organizing map showing the clustering among benign and malignant samples according to the expression of the 9 markers, the top legend refers to BRAF mutational status and malignant and benign status. (B) Self-organizing map showing the clustering among the malignant samples according to the expression of the 9 markers, the top legend refers to BRAF mutational status

By looking at the clustering of malignant samples according to BRAF mutational status, we observed that TC1 expression tended to be higher in BRAF wild type samples and we sought to determine the statistical significance of such trend.

TC1 expression resulted significantly (p=0.001) higher in BRAF wild type samples as compared to the V600E ones (Fig. 12A), thus, oncogenic BRAF may play a role in modulating TC1 expression.

We then sought to test whether the tendency of BRAF V600E samples to down regulate TC1 was restricted to thyroid tumors or conversely present also in melanoma. Towards this goal,

we extrapolated TC1 expression values (Log2 transformed) from microarray data of 112 melanoma metastases and evaluated its differential expression according to BRAF mutational status without finding any statistically significant association (p=0.8) (Fig. 12B). Thus, the modulation of TC1 expression by BRAF V600E mutation might be a thyroid-specific event.

We further compared the TC1 expression of thyroid benign samples to the one of BRAF wild type and V600E malignant samples separately. A statistically significant TC1 higher expression (p=0.005) was found in BRAF wild type but not V600E malignant samples as compared to the benign tumors (Fig. 12C), thus the behavior of thyroid tumors bearing BRAF V600E mutation recapitulates the one of benign samples in terms of TC1 expression.

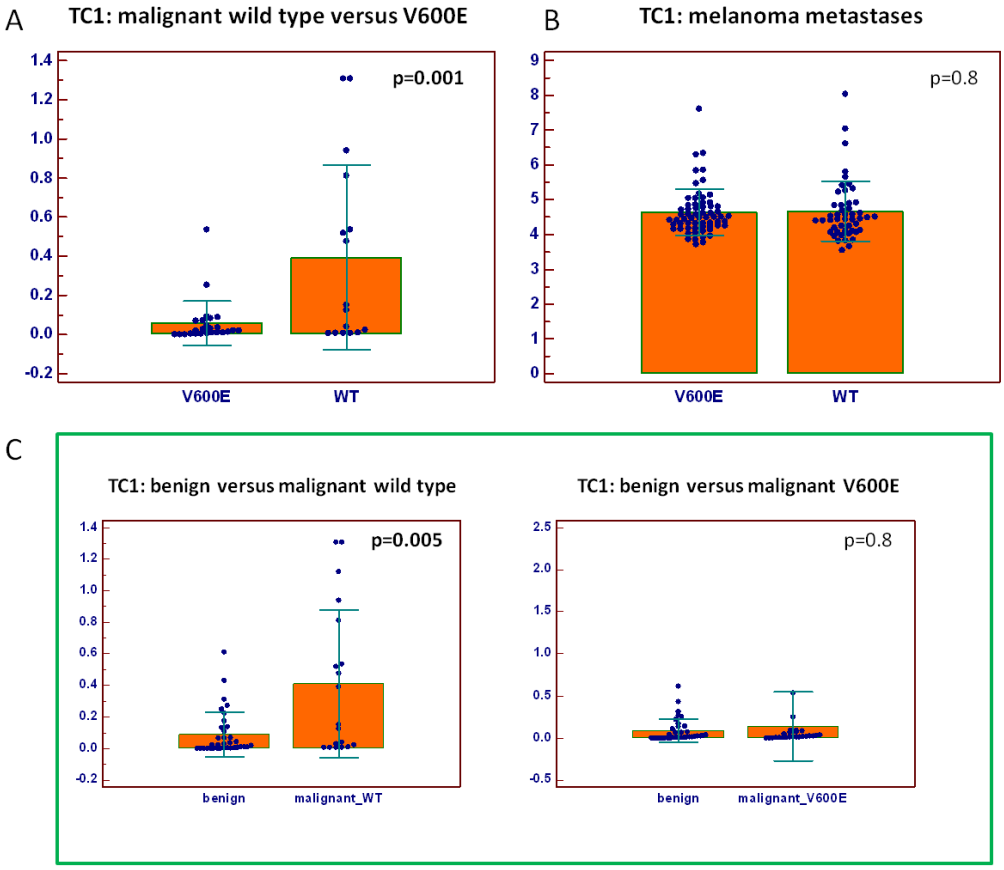

Figure 12. TC1 expression in BRAF wild type versus V600E thyroid samples (A) and melanoma metastases (B). TC1 expression in benign versus malignant BRAF wild type samples and in benign versus malignant BRAF V600E samples in thyroid (C)

We also tested the differential expression of each other marker according to BRAF mutational status in the malignant group (Fig. 13). Among all the markers and besides TC1, only KIT and CDH1 expression resulted significantly different among BRAF wild type and V600E samples.

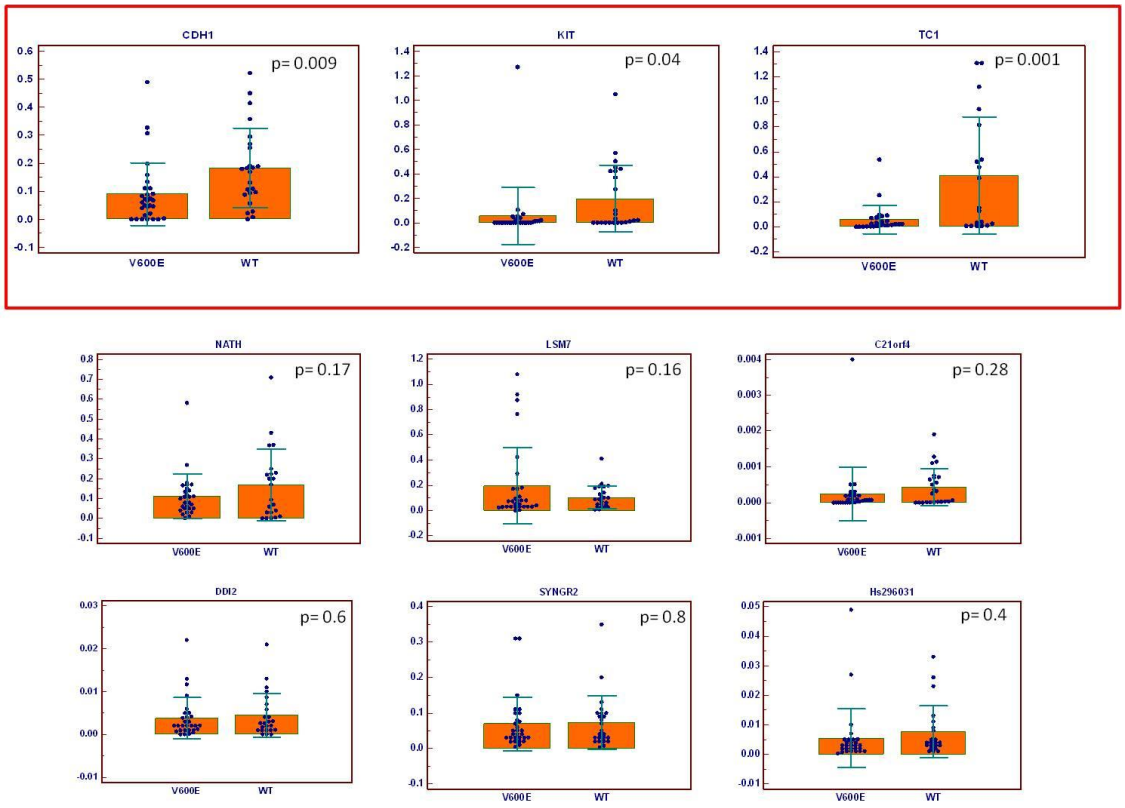

Figure 13. Differential expression of each single marker according to BRAF mutational status in the malignant samples

We finally tested whether KIT, TC1 and CDH1 expression could discriminate between BRAF wild type and V600E samples. The discriminant analysis showed a predictive power of 80% (p=0.0009)

4.11 Neural Networks

The markers expression data were used to build Bayesian Neural Networks (BNN) in order to estimate the probability of thyroid malignancy.

We built several BNNs in order to find the most predictive one. This procedure uses a Probabilistic Neural Network (PNN) to classify cases into malignant and benign categories, based on 9 input variables (KIT, LSM7, C21orf4, DDI2, SYNGR2, TC1, Hs.296031, CDH1, and NATH), by implementing a nonparametric method for classifying observations into one of benign and malignant groups based on the observed expression variables.

The Neural Network Bayesian Classifier made up of all markers has a predictive power of 80%, while the classifier made up of KIT, CDH1, LSM7, C21orf4, DDI2, TC1 and Hs.296031 resulted to have a discriminative power of 87.7%.

This analysis was then conducted on 6 unknown samples in order to confirm the accuracy of the model. The pathological diagnosis for each sample was kept blinded until after the analysis was completed. When the blind was broken, we found that 5 of the 6 unknown samples were diagnosed by the model in concordance with the diagnosis determined by standard pathologic criteria.

We also built a neural network classifier made up of the markers used in the most predictive model (KIT, CDH1, LSM7, C21orf4, DDI2, TC1 and Hs.296031) plus BRAF status. This classifier resulted to have a predictive power of 88.8%, and, more importantly, it resulted to completely discriminate the 6 unknown samples when the blind was broken (Table 8). By applying the BNN model, no classification errors came out when the probability of diagnosis was higher than 90%, thus allowing us to use this model as a correct predictor of samples with a probability score >90% ($p<0.0001$).

Table 8. Probability values of the prediction model for the unknown samples

Unknown samples	Benign probability	Malignant probability	Predicted diagnosis	Pathological diagnosis
A	3,07E-07	1	Malignant	Malignant
B	0,294935	0,705065	Malignant	Malignant
C	0,427773	0,572227	Malignant	Malignant
D	7,09E-11	1	Malignant	Malignant
E	0,00012769	0,999872	Malignant	Malignant
F	0,94438	0,05562	Benign	Benign

4.12 Role of molecular diagnosis (BRAF/KIT/BNN) in increasing the diagnostic accuracy of FNAC

We stratified the samples depending on either the histological and cytological diagnosis (Table 9) and then calculated the gain of diagnostic accuracy obtained by applying BRAF molecular analysis, KIT expression model and BNN model to the indeterminate samples (Table 10).

Among the indeterminate samples (IFP and SPTC) at the cytological level, 11 SPTC were correctly diagnosed as malignant by BRAF test, 4 additional samples were correctly classified by KIT model as 1 malignant and 3 benign, and 9 additional samples were diagnosed by the BNN model as 1 malignant and 8 benign. As shown in Table 10, after the molecular analysis we can move 13 malignant samples to the diagnostic group of PTC and the total number of PTC rises from 30 (61%) to 43 (88%) with an advantage in diagnostic accuracy of malignancy of 27%. Similarly, if we move the 11 IFP samples diagnosed as benign after molecular analysis in the diagnostic group of BN the total number of BN raises from 19 (50%) to 30 (79%) with a gain of diagnostic accuracy of benignity of 29%.

Finally, if we consider both PTC and BN diagnoses, the whole diagnostic accuracy gain is of 28% with a statistically significant p-value of 0.0001.

Table 9. Histological and cytological diagnosis of 87 thyroid nodules

Histological Diagnosis	***Cytological Diagnosis***		
	PTC[a]	**SPTC[b]**	**IFP[c]**
PTC[a]: 49 cases	30 (61%)	14 (29%)	5 (10%)
	BN[d]	**IFP[c]**	
BN[d]: 38 cases	19 (50%)	19 (50%)	

[a]PTC: papillary thyroid carcinoma, [b]SPTC: suspicious papillary thyroid carcinoma, [c]IFP: indeterminate follicular proliferation, [d]BN: benign

Table 10. Role of molecular diagnosis in increasing the diagnostic accuracy of FNAC

CD[a]	**BRAF V600E**	**KIT class 1** (malignancy probability 100%)	**KIT class 4** (benignity probability 100%)	**BNN**[d] (probability score >90%)
SPTC[b]	11(M[i])	1(M[i])		1(M[i])
IFP[c]			3(B[h])	8(B[h])
correctly diagnosed samples	**CD**[e]	**MD**[f]	**DA**[g]	
	49/87: 56%	73/87: 84%	+28%	

[a]CD: cytological diagnosis, [b]SPTC: suspicious papillary thyroid carcinoma, [c]IFP: indeterminate follicular proliferation, [d]BNN: Bayesian neural network, [e]CD: cytological diagnosis, [f]MD: molecular diagnosis, [g]DA: diagnostic accuracy, [h]B: benign, [i]M: malignant

4.13 Correlation analysis

A multiple-variable analysis was performed to analyze the correlation between the markers. The knowledge of the correlation of the biomarkers analyzed could give us a better understanding of the mechanisms underlying thyroid cancer biology. Thus, the statistical correlation may reflect biologically correlation between markers.

The following pairs of variables resulted to be statistically ($p<0.05$) correlated:

KIT and NATH, KIT and C21orf4, KIT and DDI2, CDH1 and NATH, CDH1 and DDI2, NATH and C21orf4, NATH and DDI2, NATH and SYNGR2, NATH and Hs296031, C21orf4 and DDI2, C21orf4 and Hs296031, DDI2 and TC1, DDI2 and Hs296031, SYNGR2 and Hs296031, TC1 and Hs296031 (Fig. 14, Table 11).

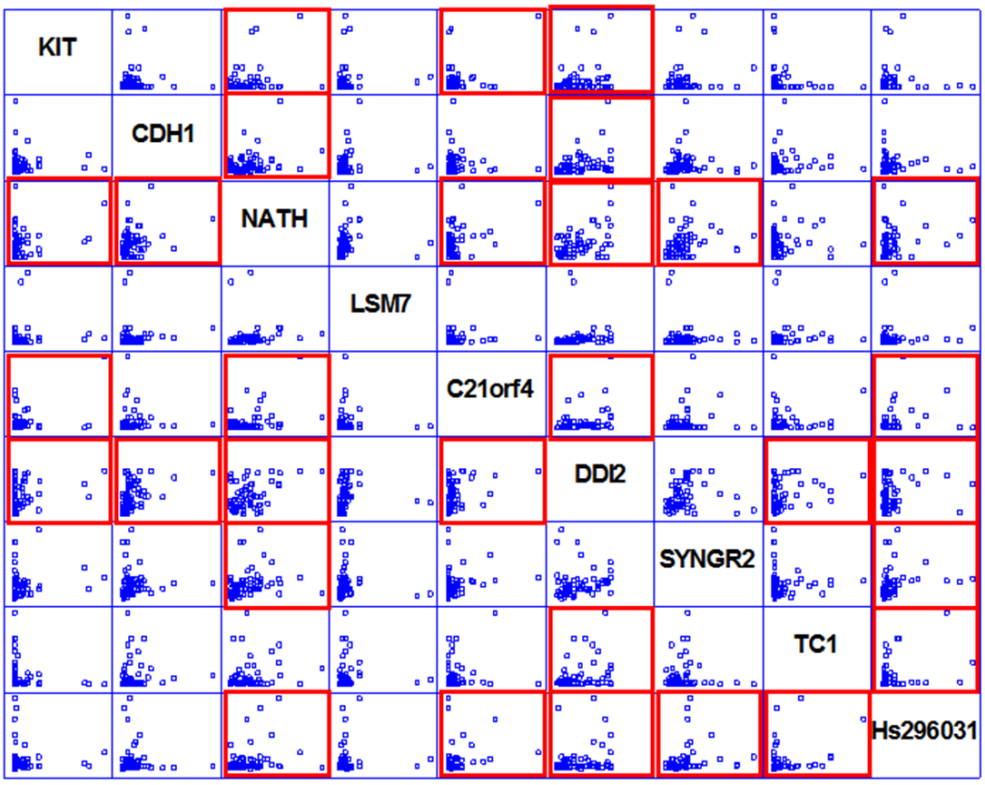

Figure 14. Correlation graph. The indicated variables (KIT, CDH1, NATH, LSM7, C21orf4, DDI2, SYNGR2, TC1, Hs. 296031) are displayed on the vertical axis of every plot in that row and on the horizontal axis of every plot in that column. Each pair of variables is thus shown twice, once above the diagonal and once below it. Significative correlations are marked with red boxes

Table 11. Correlation matrix of the 9 expression markers. In bold are the Pearson correlation R values and in red the statistical significant p-value for the specific variables pair

	KIT	CDH1	NATH	LSM7	C21orf4	DDI2	SYNGR2	TC1	Hs296031
KIT		**0,0410**	**0,2598**	**0,0787**	**0,4374**	**0,2736**	**0,1658**	**-0,1417**	**0,0008**
		0,7250	0,0234	0,4990	0,0001	0,0168	0,1524	0,2220	0,9944
CDH1	**0,0410**		**0,3383**	**0,0948**	**-0,0057**	**0,3966**	**0,0513**	**0,0823**	**-0,0674**
	0,7250		0,0028	0,4155	0,9613	0,0004	0,6601	0,4794	0,5631
NATH	**0,2598**	**0,3383**		**0,1170**	**0,4320**	**0,4503**	**0,2372**	**0,1901**	**0,2814**
	0,0234	0,0028		0,3140	0,0001	0,0000	0,0391	0,1001	0,0138
LSM7	**0,0787**	**0,0948**	**0,1170**		**-0,0225**	**0,1133**	**-0,0571**	**0,0309**	**0,0669**
	0,4990	0,4155	0,3140		0,8471	0,3300	0,6243	0,7913	0,5656
C21orf4	**0,4374**	**-0,0057**	**0,4320**	**-0,0225**		**0,4032**	**0,1646**	**0,1862**	**0,2566**
	0,0001	0,9613	0,0001	0,8471		0,0003	0,1553	0,1073	0,0252
DDI2	**0,2736**	**0,3966**	**0,4503**	**0,1133**	**0,4032**		**0,1934**	**0,3165**	**0,2291**
	0,0168	0,0004	0,0000	0,3300	0,0003		0,0941	0,0053	0,0466
SYNGR2	**0,1658**	**0,0513**	**0,2372**	**-0,0571**	**0,1646**	**0,1934**		**0,0072**	**0,2468**
	0,1524	0,6601	0,0391	0,6243	0,1553	0,0941		0,9509	0,0316
TC1	**-0,1417**	**0,0823**	**0,1901**	**0,0309**	**0,1862**	**0,3165**	**0,0072**		**0,3405**
	0,2220	0,4794	0,1001	0,7913	0,1073	0,0053	0,9509		0,0026
Hs296031	**0,0008**	**-0,0674**	**0,2814**	**0,0669**	**0,2566**	**0,2291**	**0,2468**	**0,3405**	
	0,9944	0,5631	0,0138	0,5656	0,0252	0,0466	0,0316	0,0026	

5. DISCUSSION

Papillary thyroid carcinoma (PTC) is the most common malignancy in thyroid tissue; about 80% of incident thyroid cancers are PTC. Although PTC is usually associated with alterations in the RET/PTC-RAS-BRAF signaling pathway (20, 41), the detailed molecular mechanism is unclear. A few papers mentioning a role for KIT in thyroid malignancies suggested to perform an analysis of KIT expression on thyroid cells obtained by FNAC from benign and malignant thyroid nodules, with the double aim to study a human model of thyroid cancer and, at the same time, to verify if KIT expression analysis could be of any clinical interest. The present study evaluated KIT expression in a group of thyroid FNAC, assessed later on as malignant or benign by post-surgical histological analysis, and describes the silencing of KIT in PTC. These data are in full accordance with a previous paper of Mazzanti *et al*. (66), who were able to identify two classifier models, of 10 and 16 genes respectively, discriminative of PTC and BN, in both of which KIT gene was included. In this analysis KIT resulted heterogeneously expressed in goiters, whereas PTCs were negative. These data give also strength to the report that multiple miRNAs predicted to target KIT (miR-221, miR-222, miR-146) are up-regulated in PTC (72, 118).

The biological significance of loss of KIT in thyroid malignancies is not clear. SCF, the KIT ligand, is not mitogenic in primary cultures of thyrocytes even in conjunction with thyroid-stimulating hormone (124), a result which would indicate that SCF/KIT pathway may control some aspects of the thyrocyte differentiated phenotype rather than cell division. This would agree with the apparently strong selection for loss of KIT expression in neoplastic transformation of thyroid epithelium. This negative selection is in stark contrast with the gain of function due to genetic alterations of tyrosine kinase receptors (TRKs) in other types of cancers, suggesting that TRK signaling pathways may have opposite biological effects in different cell types.

To explore the diagnostic utility of KIT expression in thyroid nodules, its expression values were divided in four arbitrarily defined classes, with class I characterized by the complete silencing of the gene. Class I and IV represented the two most informative groups, with 100% of the samples found malignant or benign respectively. Class III was also very informative including 86% of benign samples and having overall the highest statistical significance. On the other hand, in class II the samples belonging to the malignant group were 66%, which resulted

non significant. The fact that this was the only value of the study that does not reach a statistical significance, but is close to it, suggests that a higher number of cases could modify the statistical results. However, a ROC analysis was performed to measure the diagnostic performance of the model (Fig. 7), showing its good efficacy in predicting the malignant events (AUC=0.9 C.I 95% 0.79-.94; p=0.001).

On the other side, quite interestingly, BRAF V600E mutation, which is a well-known marker for PTC, was found to be statistically more present in the KIT classes I and II, while class III and IV did not contain any sample, supporting therefore the association of low KIT expression levels to a malignant status.

Molecular Pathology is the modern version of Pathology, where the whole of morphology and molecular alterations represents a powerful approach to diagnosis. In this line, this study aimed to verify the diagnostic potential of KIT expression analysis and demonstrated that the combined BRAF mutation and KIT expression approach is able to increase the diagnostic accuracy of FNAC of thyroid nodules of 18% for a diagnosis of malignancy and 9% for a diagnosis of benignity.

Despite several carcinomas showed activating mutations of KIT gene (GISTs, melanomas, haematopoietic and lymphoid tumors), they have not been described in thyroid tumors and this study revealed a wild type sequence of KIT gene in exons 9, 11, 13, and 17.

Finally, as previously published by Jin *et al.* (125), the present study shows that not only DNA but RNA too can be easily extracted from stained smears of FNAC and easily analyzed by qRT-PCR. KIT receptor expression was detectable regardless of the time of specimen collection from the archived material, we were able to successfully use slides prepared 7 years ago and kept in our archives. Moreover all of the smears were independently reviewed by a senior cytopathologist, to assure adequate thyroid cell representation and confirm the cytological diagnosis of the slides in which KIT receptor expression was investigated. The simple method named manual macro-dissection and described elsewhere allows to perform molecular analysis only on selected cell population to be studied.

Many candidate markers of thyroid cancer have been identified in microarray studies that require analytic and clinical validation in a cohort large enough to permit evaluation of the clinical utility of these markers. Quantitative RT-PCR has become a highly reliable technique that allows for precise quantification of gene expression levels identified on microarray

studies from various laboratories (126-128). Moreover, it is a technique that has been in clinical use as a diagnostic test in various fields of medicine. We believe our approach using real-time quantitative RT-PCR in a large study cohort of clinical FNA slide samples gives an accurate estimate of the clinical utility of using our 8-gene assay, as a diagnostic test to distinguish benign from malignant thyroid neoplasm.

Currently, the diagnosis of thyroid nodules relies primarily on cytology. For the majority of patients with PTC, FNA-based cytology can make a diagnosis with high accuracy (129). However, there is a significant proportion of neoplasm in which this FNA-based pre-operative cytological diagnosis fails.

The primary aim of this study was to find a diagnostic accurate pre-operative assay able to distinguish benign from malignant thyroid neoplasm. We found 5 of 9 proposed gene markers (KIT, LSM7, C21orf4, DDI2, CDH1) differentially expressed in malignant and benign thyroid samples with a significant p-value (<0.05). NATH, TC1, Hs.296031, SYNGR2 mRNA were not significantly differentially expressed between benign and malignant samples.

We also performed the receiver operating characteristic (ROC) curves analysis in order to optimize the model for negative and positive predictive value in our thyroid cohort. We found that the ROC analysis for KIT, LSM7, C21orf4, DDI2, and CDH1 had a high diagnostic accuracy. The AUC for each significant marker ranged between 0.625 and 0.900. Therefore all the significant markers, alone and in combination, can be used to distinguish between malignant and benign FNA samples. The ROC analysis for the significant markers and BRAF status in combination showed an AUC of 0.8824, a sensitivity of 91% and specificity of 63%.

When we applied the Principal Component Analysis to the benign and malignant samples we observed an overall separation among them according to the expression of the 9 markers in the study, thus the markers can together discriminate between benign and malignant thyroid tumors (Fig. 10A); such separation also reflects the differential BRAF mutational status (Fig. 10B). Similarly, a different classification of benign and malignant samples as well as of BRAF wild type and V600E samples was observed when clustering the samples according to the expression of the 9 makers (Fig. 11A), though not absolute.

We then tested the differential expression profile of the 9 markers within the malignant group according to BRAF mutational status. Quite interestingly, malignant tumors bearing the V600E mutation were observed to cluster separately from wild type samples expressing the

markers in a higher extent, and, more importantly, the behaviour of the malignant wild type samples recapitulated overall the one of benign samples. This is an important finding; in fact since BRAF V600E has been shown to be associated with higher tumor aggressiveness (130-132), the co-under expression of the 9 markers emerges as sign of more malignant tumor behaviour and should be prospectively evaluated.

From the clustering of malignant samples according to BRAF mutational status we observed that TC1 expression tended to be higher in BRAF wild type samples and we sought to determine the statistical significance of such trend.

TC1 expression resulted significantly ($p=0.001$) higher in BRAF wild type samples as compared to the V600E ones (Fig. 12A), thus, oncogenic BRAF may play a role in modulating TC1 expression. Such tendency can be ascribed to a more aggressive behaviour of tumors that down regulate TC1. The mechanisms underlying TC1 down expression in BRAF V600E malignant samples are currently unknown and further studies are warranted to exploit this phenomenon.

We then tested whether the tendency of BRAF V600E samples to down regulate TC1 was restricted to thyroid tumors or conversely present also in melanoma. Towards this goal, we extrapolated TC1 expression values (Log2 transformed) from microarray data of 112 melanoma metastases and evaluated its differential expression according to BRAF mutational status without finding any statistically significant association ($p=0.8$) (Fig. 12B). Thus, the modulation of TC1 expression by BRAF V600E mutation might be thyroid-specific.

We lastly compared the TC1 expression of benign samples to the one of BRAF wild type and V600E malignant samples separately. A statistically significant TC1 higher expression ($p=0.005$) was found in BRAF wild type but not V600E malignant samples as compared to the benign tumors (Fig. 12C), thus the behavior of thyroid tumors bearing BRAF V600E mutation recapitulates the one of benign samples in terms of TC1 expression.

By looking at the expression of the other markers we found that, besides TC1, also CDH1 and KIT expression significantly discriminate between BRAF wild type and V600E malignant samples (Fig. 13) with a power of 80% ($p=0.0009$). As mentioned before, CDH1 encodes for the E-cadherin and its loss has been shown to be important for disruption of tight epithelial cell-cell contacts and release of invasive tumor cells from the primary tumor, thus its association with BRAF mutational status may reflect an aggressive tumor phenotype. As for CDH1, also KIT down regulation has been related to the progression into a malignant

phenotype and its association with BRAF V600E mutation may reflect this phenomenon. Regarding TC1, it is difficult to compare our data to the literature because studies addressing TC1 gene expression are sparse. The little information available suggests that TC1 is over expressed in thyroid papillary carcinoma (101, 109, 133) and concordant to the literature we found a higher TC1 expression in malignant tumors. However, the expression of TC1 is significantly lower in malignant thyroid cancers bearing BRAF V600E mutation, that suggests a potential influence of BRAF status on TC1 expression, but further functional studies are needed to understand the underlying mechanisms.

In the last few years, a new class of techniques known as Bayesian Neural Networks (BNN) have been proposed as a supplement or alternative to standard statistical techniques (134). BNNs do not require explicit distributional assumptions (such as normality). This advantage has generated considerable interest in the use of neural network techniques for the classification of medical outcomes (134). We developed a Bayesian Artificial Neural Network model based on data collected from FNA samples. Bayesian classification has been applied across the spectrum of medicine, from optimization of pharmacotherapy dosing (135, 136), predicting cancer screening (137) and diagnostic test results (138, 139), to determining injury severity (140), assessing operative risk (141) and predicting surgical outcomes (142-145). We built several Neural Networks and the most predictive one has resulted to be made up of KIT, CDH1, LSM7, C21orf4, DDI2, TC1 and Hs.296031, with a power of 87.7%. The network was then validated on 6 unknown samples. The model determined the accurate diagnosis of 5 of 6 unknown samples tested. Accuracy was based on a comparison to the gold standard pathological diagnosis as determined by clinical pathologists.

It's important to notice that we have put in the model also the not significant markers (TC1, Hs.296031), because their contribution in the discriminative power seems to be relevant. In fact, even though a variable is not significant, its combination with other variables may be significant, because of a certain tendency that makes them to go in the same way.

The classifier built by using also BRAF status resulted to have a predictive marker of 88.8% and to successfully discriminate the unknown samples when the blind was broken (Table 8), thus the gene expression analysis combined to the BRAF mutational analysis could represent a very useful test to pre-operatively discriminate benign and malignant thyroid tumors.

The probability of the prediction of diagnosis for almost all the samples resulted to range between 95% and 100%, thus, although the general prediction value is 88.8%, the discriminative power to assess each sample can reach a value of 100%. These data also strengthen the importance of the 8-markers model as an adjunctive tool for the pre-operative diagnosis of thyroid nodules.

We also stratified the samples depending on either the histological and cytological diagnoses (Table 9). The gain of diagnostic accuracy obtained by applying BRAF molecular analysis, KIT expression model and BNN model was then calculated.

By applying the BNN model, no classification errors came out when the probability of diagnosis was higher than 90%, thus allowing us to use this model as a correct predictor of samples with a probability score >90% ($p<0.0001$).

We then calculated the diagnostic accuracy gained by applying molecular diagnostic tests (Table 10).

Among the uncertain samples (IFP and SPTC) at the cytological level, 11 were correctly diagnosed by BRAF test, 4 additional samples by KIT model and 9 additional samples by the BNN model. It is important to point out that IFP lesions are often very difficult to diagnose even at frozen section and in this study we developed a molecular approach that is able to correctly classify as certain benign 46% (11/24) of IFP lesions. Therefore by the use of molecular approach these patients would have been clinically enrolled to the follow up group instead of sent to surgery. Thus, the combined use of the molecular tests resulted to produce a diagnostic accuracy gain of 28% (Table 10). Basically, what we propose is the use of BRAF molecular analysis (after uncertain cytological diagnosis) to assess the malignancy of thyroid nodules in the first place, then the use of KIT model for the indeterminate nodules and at last the use of the 8-gene model to ultimately assess the diagnosis of the nodules that otherwise would remain suspicious. The combinatorial power of these tools could definitely increase the percentage of thyroid nodules correctly classified while decreasing the ones remained indeterminate.

All these findings strengthen the importance of molecular pathology where the morphology and the molecular alterations represent a powerful approach to diagnosis. In this line, this study aimed to assess the diagnostic potential of the 8-gene expression model as an adjunctive tool in the pre-operative management of thyroid nodules. We demonstrated that

the 8-gene expression model provides an increased diagnostic power to the molecular pathology approach based on BRAF mutation and KIT expression analysis.

We also performed a multiple variable analysis among all the markers analyzed, independently on the diagnostic classification, in order to evaluate a possible functional correlation among the markers (Fig. 10). In literature there is no evidence about the biological correlation among the well studied markers, but it's interesting to note that the unknown marker Hs.296031 statistically correlates with NATH, C21orf4, DDI2, SYNGR2 and TC1. This may reflect also a biological correlation, thus further studies are needed to explore this issue.

6. CONCLUSION

In summary, we have demonstrated that the expression-based classification of thyroid lesions here proposed is highly accurate and may provide a tool to overcome the difficulties in today's pre-operative diagnosis of thyroid suspicious malignancies. We hoped that this test will be a useful adjunct to the pre-operative diagnosis of thyroid nodules.

REFERENCES

1. **Lloyd RV, Erickson LA, Casey MB, Lam KY, Lohse CM, Asa SL, Chan JK, DeLellis RA, Harach HR, Kakudo K, LiVolsi VA, Rosai J, Sebo TJ, Sobrinho-Simoes M, Wenig BM, Lae ME** 2004 Observer variation in the diagnosis of follicular variant of papillary thyroid carcinoma. Am J Surg Pathol 28:1336-1340
2. **Nikiforova MN, Nikiforov YE** 2008 Molecular genetics of thyroid cancer: implications for diagnosis, treatment and prognosis. Expert Rev Mol Diagn 8:83-95
3. **Nikiforov YE, Nikiforova MN** 2011 Molecular genetics and diagnosis of thyroid cancer. Nat Rev Endocrinol 7:569-580
4. **Albores-Saavedra J, Henson DE, Glazer E, Schwartz AM** 2007 Changing patterns in the incidence and survival of thyroid cancer with follicular phenotype--papillary, follicular, and anaplastic: a morphological and epidemiological study. Endocr Pathol 18:1-7
5. **Burgess JR, Tucker P** 2006 Incidence trends for papillary thyroid carcinoma and their correlation with thyroid surgery and thyroid fine-needle aspirate cytology. Thyroid 16:47-53
6. **Colonna M, Grande E, Jonasson JG, Eurocare Working G** 2006 Variation in relative survival of thyroid cancers in Europe: results from the analysis on 21 countries over the period 1983-1994 (EUROCARE-3 study). Eur J Cancer 42:2598-2608
7. **Davies L, Welch HG** 2006 Increasing incidence of thyroid cancer in the United States, 1973-2002. JAMA 295:2164-2167
8. **Ferlay J, Parkin DM, Steliarova-Foucher E** 2010 Estimates of cancer incidence and mortality in Europe in 2008. Eur J Cancer 46:765-781
9. **Liu S, Semenciw R, Ugnat AM, Mao Y** 2001 Increasing thyroid cancer incidence in Canada, 1970-1996: time trends and age-period-cohort effects. Br J Cancer 85:1335-1339
10. **Jemal A, Tiwari RC, Murray T, Ghafoor A, Samuels A, Ward E, Feuer EJ, Thun MJ** 2004 Cancer statistics, 2004. CA Cancer J Clin 54:8-29
11. **Marchetti I, Iervasi G, Mazzanti CM, Lessi F, Tomei S, Naccarato AG, Aretini P, Alberti B, Di Coscio G, Bevilacqua G** 2011 Detection of the BRAF(V600E) Mutation in Fine Needle Aspiration Cytology of Thyroid Papillary Microcarcinoma Cells Selected by

Manual Macrodissection: An Easy Tool to Improve the Preoperative Diagnosis. Thyroid

12. **Frates MC, Benson CB, Doubilet PM, Kunreuther E, Contreras M, Cibas ES, Orcutt J, Moore FD, Jr., Larsen PR, Marqusee E, Alexander EK** 2006 Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. J Clin Endocrinol Metab 91:3411-3417
13. **Gharib H** 2004 Changing trends in thyroid practice: understanding nodular thyroid disease. Endocr Pract 10:31-39
14. **Mazzaferri EL** 1992 Thyroid cancer in thyroid nodules: finding a needle in the haystack. Am J Med 93:359-362
15. **Greaves TS, Olvera M, Florentine BD, Raza AS, Cobb CJ, Tsao-Wei DD, Groshen S, Singer P, Lopresti J, Martin SE** 2000 Follicular lesions of thyroid: a 5-year fine-needle aspiration experience. Cancer 90:335-341
16. **Sclabas GM, Staerkel GA, Shapiro SE, Fornage BD, Sherman SI, Vassillopoulou-Sellin R, Lee JE, Evans DB** 2003 Fine-needle aspiration of the thyroid and correlation with histopathology in a contemporary series of 240 patients. Am J Surg 186:702-709; discussion 709-710
17. **Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Sherman SI, Tuttle RM** 2006 Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 16:109-142
18. **Yassa L, Cibas ES, Benson CB, Frates MC, Doubilet PM, Gawande AA, Moore FD, Jr., Kim BW, Nose V, Marqusee E, Larsen PR, Alexander EK** 2007 Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. Cancer 111:508-516
19. **Adeniran AJ, Zhu Z, Gandhi M, Steward DL, Fidler JP, Giordano TJ, Biddinger PW, Nikiforov YE** 2006 Correlation between genetic alterations and microscopic features, clinical manifestations, and prognostic characteristics of thyroid papillary carcinomas. Am J Surg Pathol 30:216-222
20. **Kimura ET, Nikiforova MN, Zhu Z, Knauf JA, Nikiforov YE, Fagin JA** 2003 High prevalence of BRAF mutations in thyroid cancer: genetic evidence for constitutive activation of the RET/PTC-RAS-BRAF signaling pathway in papillary thyroid carcinoma. Cancer Res 63:1454-1457

21. **Soares P, Trovisco V, Rocha AS, Lima J, Castro P, Preto A, Maximo V, Botelho T, Seruca R, Sobrinho-Simoes M** 2003 BRAF mutations and RET/PTC rearrangements are alternative events in the etiopathogenesis of PTC. Oncogene 22:4578-4580

22. **Frattini M, Ferrario C, Bressan P, Balestra D, De Cecco L, Mondellini P, Bongarzone I, Collini P, Gariboldi M, Pilotti S, Pierotti MA, Greco A** 2004 Alternative mutations of BRAF, RET and NTRK1 are associated with similar but distinct gene expression patterns in papillary thyroid cancer. Oncogene 23:7436-7440

23. **Cohen Y, Xing M, Mambo E, Guo Z, Wu G, Trink B, Beller U, Westra WH, Ladenson PW, Sidransky D** 2003 BRAF mutation in papillary thyroid carcinoma. J Natl Cancer Inst 95:625-627

24. **Xing M** 2005 BRAF mutation in thyroid cancer. Endocr Relat Cancer 12:245-262

25. **Ciampi R, Nikiforov YE** 2005 Alterations of the BRAF gene in thyroid tumors. Endocr Pathol 16:163-172

26. **Wan PT, Garnett MJ, Roe SM, Lee S, Niculescu-Duvaz D, Good VM, Jones CM, Marshall CJ, Springer CJ, Barford D, Marais R** 2004 Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. Cell 116:855-867

27. **Nikiforova MN, Kimura ET, Gandhi M, Biddinger PW, Knauf JA, Basolo F, Zhu Z, Giannini R, Salvatore G, Fusco A, Santoro M, Fagin JA, Nikiforov YE** 2003 BRAF mutations in thyroid tumors are restricted to papillary carcinomas and anaplastic or poorly differentiated carcinomas arising from papillary carcinomas. J Clin Endocrinol Metab 88:5399-5404

28. **Namba H, Nakashima M, Hayashi T, Hayashida N, Maeda S, Rogounovitch TI, Ohtsuru A, Saenko VA, Kanematsu T, Yamashita S** 2003 Clinical implication of hot spot BRAF mutation, V599E, in papillary thyroid cancers. J Clin Endocrinol Metab 88:4393-4397

29. **Xing M, Westra WH, Tufano RP, Cohen Y, Rosenbaum E, Rhoden KJ, Carson KA, Vasko V, Larin A, Tallini G, Tolaney S, Holt EH, Hui P, Umbricht CB, Basaria S, Ewertz M, Tufaro AP, Califano JA, Ringel MD, Zeiger MA, Sidransky D, Ladenson PW** 2005 BRAF mutation predicts a poorer clinical prognosis for papillary thyroid cancer. J Clin Endocrinol Metab 90:6373-6379

30. **Kim TY, Kim WB, Rhee YS, Song JY, Kim JM, Gong G, Lee S, Kim SY, Kim SC, Hong SJ, Shong YK** 2006 The BRAF mutation is useful for prediction of clinical recurrence in

low-risk patients with conventional papillary thyroid carcinoma. Clin Endocrinol (Oxf) 65:364-368

31. **Riesco-Eizaguirre G, Gutierrez-Martinez P, Garcia-Cabezas MA, Nistal M, Santisteban P** 2006 The oncogene BRAF V600E is associated with a high risk of recurrence and less differentiated papillary thyroid carcinoma due to the impairment of Na+/I- targeting to the membrane. Endocr Relat Cancer 13:257-269
32. **Trovisco V, Vieira de Castro I, Soares P, Maximo V, Silva P, Magalhaes J, Abrosimov A, Guiu XM, Sobrinho-Simoes M** 2004 BRAF mutations are associated with some histological types of papillary thyroid carcinoma. J Pathol 202:247-251
33. **Carta C, Moretti S, Passeri L, Barbi F, Avenia N, Cavaliere A, Monacelli M, Macchiarulo A, Santeusanio F, Tartaglia M, Puxeddu E** 2006 Genotyping of an Italian papillary thyroid carcinoma cohort revealed high prevalence of BRAF mutations, absence of RAS mutations and allowed the detection of a new mutation of BRAF oncoprotein (BRAF(V599lns)). Clin Endocrinol (Oxf) 64:105-109
34. **Hou P, Liu D, Xing M** 2007 Functional characterization of the T1799-1801del and A1799-1816ins BRAF mutations in papillary thyroid cancer. Cell Cycle 6:377-379
35. **Begum S, Rosenbaum E, Henrique R, Cohen Y, Sidransky D, Westra WH** 2004 BRAF mutations in anaplastic thyroid carcinoma: implications for tumor origin, diagnosis and treatment. Mod Pathol 17:1359-1363
36. **Fusco A, Grieco M, Santoro M, Berlingieri MT, Pilotti S, Pierotti MA, Della Porta G, Vecchio G** 1987 A new oncogene in human thyroid papillary carcinomas and their lymph-nodal metastases. Nature 328:170-172
37. **Grieco M, Santoro M, Berlingieri MT, Melillo RM, Donghi R, Bongarzone I, Pierotti MA, Della Porta G, Fusco A, Vecchio G** 1990 PTC is a novel rearranged form of the ret proto-oncogene and is frequently detected in vivo in human thyroid papillary carcinomas. Cell 60:557-563
38. **Nikiforov YE** 2002 RET/PTC rearrangement in thyroid tumors. Endocr Pathol 13:3-16
39. **Tallini G, Asa SL** 2001 RET oncogene activation in papillary thyroid carcinoma. Adv Anat Pathol 8:345-354
40. **Knauf JA, Kuroda H, Basu S, Fagin JA** 2003 RET/PTC-induced dedifferentiation of thyroid cells is mediated through Y1062 signaling through SHC-RAS-MAP kinase. Oncogene 22:4406-4412

41. **Melillo RM, Castellone MD, Guarino V, De Falco V, Cirafici AM, Salvatore G, Caiazzo F, Basolo F, Giannini R, Kruhoffer M, Orntoft T, Fusco A, Santoro M** 2005 The RET/PTC-RAS-BRAF linear signaling cascade mediates the motile and mitogenic phenotype of thyroid cancer cells. J Clin Invest 115:1068-1081

42. **Mitsutake N, Knauf JA, Mitsutake S, Mesa C, Jr., Zhang L, Fagin JA** 2005 Conditional BRAFV600E expression induces DNA synthesis, apoptosis, dedifferentiation, and chromosomal instability in thyroid PCCL3 cells. Cancer Res 65:2465-2473

43. **Zhu Z, Ciampi R, Nikiforova MN, Gandhi M, Nikiforov YE** 2006 Prevalence of RET/PTC rearrangements in thyroid papillary carcinomas: effects of the detection methods and genetic heterogeneity. J Clin Endocrinol Metab 91:3603-3610

44. **Unger K, Zitzelsberger H, Salvatore G, Santoro M, Bogdanova T, Braselmann H, Kastner P, Zurnadzhy L, Tronko N, Hutzler P, Thomas G** 2004 Heterogeneity in the distribution of RET/PTC rearrangements within individual post-Chernobyl papillary thyroid carcinomas. J Clin Endocrinol Metab 89:4272-4279

45. **Namba H, Rubin SA, Fagin JA** 1990 Point mutations of ras oncogenes are an early event in thyroid tumorigenesis. Mol Endocrinol 4:1474-1479

46. **Ezzat S, Zheng L, Kolenda J, Safarian A, Freeman JL, Asa SL** 1996 Prevalence of activating ras mutations in morphologically characterized thyroid nodules. Thyroid 6:409-416

47. **Zhu Z, Gandhi M, Nikiforova MN, Fischer AH, Nikiforov YE** 2003 Molecular profile and clinical-pathologic features of the follicular variant of papillary thyroid carcinoma. An unusually high prevalence of ras mutations. Am J Clin Pathol 120:71-77

48. **Hara H, Fulton N, Yashiro T, Ito K, DeGroot LJ, Kaplan EL** 1994 N-ras mutation: an independent prognostic factor for aggressiveness of papillary thyroid carcinoma. Surgery 116:1010-1016

49. **Lemoine NR, Mayall ES, Wyllie FS, Williams ED, Goyns M, Stringer B, Wynford-Thomas D** 1989 High frequency of ras oncogene activation in all stages of human thyroid tumorigenesis. Oncogene 4:159-164

50. **Suarez HG, du Villard JA, Severino M, Caillou B, Schlumberger M, Tubiana M, Parmentier C, Monier R** 1990 Presence of mutations in all three ras genes in human thyroid tumors. Oncogene 5:565-570

51. **Motoi N, Sakamoto A, Yamochi T, Horiuchi H, Motoi T, Machinami R** 2000 Role of ras mutation in the progression of thyroid carcinoma of follicular epithelial origin. Pathol Res Pract 196:1-7

52. **Basolo F, Pisaturo F, Pollina LE, Fontanini G, Elisei R, Molinaro E, Iacconi P, Miccoli P, Pacini F** 2000 N-ras mutation in poorly differentiated thyroid carcinomas: correlation with bone metastases and inverse correlation to thyroglobulin expression. Thyroid 10:19-23

53. **Garcia-Rostan G, Zhao H, Camp RL, Pollan M, Herrero A, Pardo J, Wu R, Carcangiu ML, Costa J, Tallini G** 2003 ras mutations are associated with aggressive tumor phenotypes and poor prognosis in thyroid cancer. J Clin Oncol 21:3226-3235

54. **Nikiforov YE** 2008 Thyroid carcinoma: molecular pathways and therapeutic targets. Mod Pathol 21 Suppl 2:S37-43

55. **Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM, Fletcher JA** 2000 PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected]. Science 289:1357-1360

56. **French CA, Alexander EK, Cibas ES, Nose V, Laguette J, Faquin W, Garber J, Moore F, Jr., Fletcher JA, Larsen PR, Kroll TG** 2003 Genetic and biological subgroups of low-stage follicular thyroid cancer. Am J Pathol 162:1053-1060

57. **Nikiforova MN, Lynch RA, Biddinger PW, Alexander EK, Dorn GW, 2nd, Tallini G, Kroll TG, Nikiforov YE** 2003 RAS point mutations and PAX8-PPAR gamma rearrangement in thyroid tumors: evidence for distinct molecular pathways in thyroid follicular carcinoma. J Clin Endocrinol Metab 88:2318-2326

58. **Dwight T, Thoppe SR, Foukakis T, Lui WO, Wallin G, Hoog A, Frisk T, Larsson C, Zedenius J** 2003 Involvement of the PAX8/peroxisome proliferator-activated receptor gamma rearrangement in follicular thyroid tumors. J Clin Endocrinol Metab 88:4440-4445

59. **Nikiforova MN, Biddinger PW, Caudill CM, Kroll TG, Nikiforov YE** 2002 PAX8-PPARgamma rearrangement in thyroid tumors: RT-PCR and immunohistochemical analyses. Am J Surg Pathol 26:1016-1023

60. **Gregory Powell J, Wang X, Allard BL, Sahin M, Wang XL, Hay ID, Hiddinga HJ, Deshpande SS, Kroll TG, Grebe SK, Eberhardt NL, McIver B** 2004 The PAX8/PPARgamma fusion oncoprotein transforms immortalized human thyrocytes

through a mechanism probably involving wild-type PPARgamma inhibition. Oncogene 23:3634-3641

61. **Giordano TJ, Au AY, Kuick R, Thomas DG, Rhodes DR, Wilhelm KG, Jr., Vinco M, Misek DE, Sanders D, Zhu Z, Ciampi R, Hanash S, Chinnaiyan A, Clifton-Bligh RJ, Robinson BG, Nikiforov YE, Koenig RJ** 2006 Delineation, functional validation, and bioinformatic evaluation of gene expression in thyroid follicular carcinomas with the PAX8-PPARG translocation. Clin Cancer Res 12:1983-1993

62. **Reddi HV, McIver B, Grebe SK, Eberhardt NL** 2007 The paired box-8/peroxisome proliferator-activated receptor-gamma oncogene in thyroid tumorigenesis. Endocrinology 148:932-935

63. **Giordano TJ, Kuick R, Thomas DG, Misek DE, Vinco M, Sanders D, Zhu Z, Ciampi R, Roh M, Shedden K, Gauger P, Doherty G, Thompson NW, Hanash S, Koenig RJ, Nikiforov YE** 2005 Molecular classification of papillary thyroid carcinoma: distinct BRAF, RAS, and RET/PTC mutation-specific gene expression profiles discovered by DNA microarray analysis. Oncogene 24:6646-6656

64. **Chevillard S, Ugolin N, Vielh P, Ory K, Levalois C, Elliott D, Clayman GL, El-Naggar AK** 2004 Gene expression profiling of differentiated thyroid neoplasms: diagnostic and clinical implications. Clin Cancer Res 10:6586-6597

65. **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

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

67. **Lubitz CC, Fahey TJ, 3rd** 2005 The differentiation of benign and malignant thyroid nodules. Adv Surg 39:355-377

68. **Prasad NB, Somervell H, Tufano RP, Dackiw AP, Marohn MR, Califano JA, Wang Y, Westra WH, Clark DP, Umbricht CB, Libutti SK, Zeiger MA** 2008 Identification of genes differentially expressed in benign versus malignant thyroid tumors. Clin Cancer Res 14:3327-3337

69. **Knauf JA, Sartor MA, Medvedovic M, Lundsmith E, Ryder M, Salzano M, Nikiforov YE, Giordano TJ, Ghossein RA, Fagin JA** 2011 Progression of BRAF-induced thyroid cancer is associated with epithelial-mesenchymal transition requiring concomitant MAP kinase and TGFbeta signaling. Oncogene 30:3153-3162

70. **Vasko V, Espinosa AV, Scouten W, He H, Auer H, Liyanarachchi S, Larin A, Savchenko V, Francis GL, de la Chapelle A, Saji M, Ringel MD** 2007 Gene expression and functional evidence of epithelial-to-mesenchymal transition in papillary thyroid carcinoma invasion. Proc Natl Acad Sci U S A 104:2803-2808

71. **Chen YT, Kitabayashi N, Zhou XK, Fahey TJ, 3rd, Scognamiglio T** 2008 MicroRNA analysis as a potential diagnostic tool for papillary thyroid carcinoma. Mod Pathol 21:1139-1146

72. **He H, Jazdzewski K, Li W, Liyanarachchi S, Nagy R, Volinia S, Calin GA, Liu CG, Franssila K, Suster S, Kloos RT, Croce CM, de la Chapelle A** 2005 The role of microRNA genes in papillary thyroid carcinoma. Proc Natl Acad Sci U S A 102:19075-19080

73. **Nikiforova MN, Chiosea SI, Nikiforov YE** 2009 MicroRNA expression profiles in thyroid tumors. Endocr Pathol 20:85-91

74. **Pallante P, Visone R, Ferracin M, Ferraro A, Berlingieri MT, Troncone G, Chiappetta G, Liu CG, Santoro M, Negrini M, Croce CM, Fusco A** 2006 MicroRNA deregulation in human thyroid papillary carcinomas. Endocr Relat Cancer 13:497-508

75. **Nikiforova MN, Tseng GC, Steward D, Diorio D, Nikiforov YE** 2008 MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility. J Clin Endocrinol Metab 93:1600-1608

76. **Xing M** 2007 Gene methylation in thyroid tumorigenesis. Endocrinology 148:948-953

77. **Zuo H, Gandhi M, Edreira MM, Hochbaum D, Nimgaonkar VL, Zhang P, Dipaola J, Evdokimova V, Altschuler DL, Nikiforov YE** 2010 Downregulation of Rap1GAP through epigenetic silencing and loss of heterozygosity promotes invasion and progression of thyroid tumors. Cancer Res 70:1389-1397

78. **Russo D, Damante G, Puxeddu E, Durante C, Filetti S** 2011 Epigenetics of thyroid cancer and novel therapeutic targets. J Mol Endocrinol 46:R73-81

79. **Hu S, Liu D, Tufano RP, Carson KA, Rosenbaum E, Cohen Y, Holt EH, Kiseljak-Vassiliades K, Rhoden KJ, Tolaney S, Condouris S, Tallini G, Westra WH, Umbricht**

CB, Zeiger MA, Califano JA, Vasko V, Xing M 2006 Association of aberrant methylation of tumor suppressor genes with tumor aggressiveness and BRAF mutation in papillary thyroid cancer. Int J Cancer 119:2322-2329

80. **Baloch ZW, Fleisher S, LiVolsi VA, Gupta PK** 2002 Diagnosis of "follicular neoplasm": a gray zone in thyroid fine-needle aspiration cytology. Diagn Cytopathol 26:41-44
81. **Yoder BJ, Redman R, Massoll NA** 2006 Validation of a five-tier cytodiagnostic system for thyroid fine needle aspiration biopsies using cytohistologic correlation. Thyroid 16:781-786
82. **Chen H, Zeiger MA, Clark DP, Westra WH, Udelsman R** 1997 Papillary carcinoma of the thyroid: can operative management be based solely on fine-needle aspiration? J Am Coll Surg 184:605-610
83. **Gharib H, Goellner JR** 1993 Fine-needle aspiration biopsy of the thyroid: an appraisal. Ann Intern Med 118:282-289
84. **Shibru D, Chung KW, Kebebew E** 2008 Recent developments in the clinical application of thyroid cancer biomarkers. Curr Opin Oncol 20:13-18
85. **Baloch ZW, Hendreen S, Gupta PK, LiVolsi VA, Mandel SJ, Weber R, Fraker D** 2001 Interinstitutional review of thyroid fine-needle aspirations: impact on clinical management of thyroid nodules. Diagn Cytopathol 25:231-234
86. **Yeh MW, Demircan O, Ituarte P, Clark OH** 2004 False-negative fine-needle aspiration cytology results delay treatment and adversely affect outcome in patients with thyroid carcinoma. Thyroid 14:207-215
87. **Mazzaferri EL** 1993 Management of a solitary thyroid nodule. N Engl J Med 328:553-559
88. **Baloch ZW, LiVolsi VA, Asa SL, Rosai J, Merino MJ, Randolph G, Vielh P, DeMay RM, Sidawy MK, Frable WJ** 2008 Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. Diagn Cytopathol 36:425-437
89. **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

90. **Cantara S, Capezzone M, Marchisotta S, Capuano S, Busonero G, Toti P, Di Santo A, Caruso G, Carli AF, Brilli L, Montanaro A, Pacini F** 2010 Impact of proto-oncogene mutation detection in cytological specimens from thyroid nodules improves the diagnostic accuracy of cytology. J Clin Endocrinol Metab 95:1365-1369

91. **Ohori NP, Nikiforova MN, Schoedel KE, LeBeau SO, Hodak SP, Seethala RR, Carty SE, Ogilvie JB, Yip L, Nikiforov YE** 2010 Contribution of molecular testing to thyroid fine-needle aspiration cytology of "follicular lesion of undetermined significance/atypia of undetermined significance". Cancer Cytopathol 118:17-23

92. **Moses W, Weng J, Sansano I, Peng M, Khanafshar E, Ljung BM, Duh QY, Clark OH, Kebebew E** 2010 Molecular testing for somatic mutations improves the accuracy of thyroid fine-needle aspiration biopsy. World J Surg 34:2589-2594

93. **Bedognetti D, Balwit JM, Wang E, Disis ML, Britten CM, Delogu LG, Tomei S, Fox BA, Gajewski TF, Marincola FM, Butterfield LH** 2011 SITC/iSBTc Cancer Immunotherapy Biomarkers Resource Document: online resources and useful tools - a compass in the land of biomarker discovery. J Transl Med 9:155

94. **Barden CB, Shister KW, Zhu B, Guiter G, Greenblatt DY, Zeiger MA, Fahey TJ, 3rd** 2003 Classification of follicular thyroid tumors by molecular signature: results of gene profiling. Clin Cancer Res 9:1792-1800

95. **Finley DJ, Zhu B, Barden CB, Fahey TJ, 3rd** 2004 Discrimination of benign and malignant thyroid nodules by molecular profiling. Ann Surg 240:425-436; discussion 436-427

96. **Weber F, Shen L, Aldred MA, Morrison CD, Frilling A, Saji M, Schuppert F, Broelsch CE, Ringel MD, Eng C** 2005 Genetic classification of benign and malignant thyroid follicular neoplasia based on a three-gene combination. J Clin Endocrinol Metab 90:2512-2521

97. **Aldred MA, Huang Y, Liyanarachchi S, Pellegata NS, Gimm O, Jhiang S, Davuluri RV, de la Chapelle A, Eng C** 2004 Papillary and follicular thyroid carcinomas show distinctly different microarray expression profiles and can be distinguished by a minimum of five genes. J Clin Oncol 22:3531-3539

98. **Kothapalli R, Yoder SJ, Mane S, Loughran TP, Jr.** 2002 Microarray results: how accurate are they? BMC Bioinformatics 3:22

99. **Lubitz CC, Ugras SK, Kazam JJ, Zhu B, Scognamiglio T, Chen YT, Fahey TJ, 3rd** 2006 Microarray analysis of thyroid nodule fine-needle aspirates accurately classifies benign and malignant lesions. J Mol Diagn 8:490-498; quiz 528

100. **Arnesen T, Betts MJ, Pendino F, Liberles DA, Anderson D, Caro J, Kong X, Varhaug JE, Lillehaug JR** 2006 Characterization of hARD2, a processed hARD1 gene duplicate, encoding a human protein N-alpha-acetyltransferase. BMC Biochem 7:13

101. **Chua EL, Young L, Wu WM, Turtle JR, Dong Q** 2000 Cloning of TC-1 (C8orf4), a novel gene found to be overexpressed in thyroid cancer. Genomics 69:342-347

102. **Fluge O, Bruland O, Akslen LA, Varhaug JE, Lillehaug JR** 2002 NATH, a novel gene overexpressed in papillary thyroid carcinomas. Oncogene 21:5056-5068

103. **Olesen C, Nyeng P, Kalisz M, Jensen TH, Moller M, Tommerup N, Byskov AG** 2007 Global gene expression analysis in fetal mouse ovaries with and without meiosis and comparison of selected genes with meiosis in the testis. Cell Tissue Res 328:207-221

104. **Tharun S, He W, Mayes AE, Lennertz P, Beggs JD, Parker R** 2000 Yeast Sm-like proteins function in mRNA decapping and decay. Nature 404:515-518

105. **Conte N, Charafe-Jauffret E, Delaval B, Adelaide J, Ginestier C, Geneix J, Isnardon D, Jacquemier J, Birnbaum D** 2002 Carcinogenesis and translational controls: TACC1 is down-regulated in human cancers and associates with mRNA regulators. Oncogene 21:5619-5630

106. **Moldrich RX, Laine J, Visel A, Beart PM, Laffaire J, Rossier J, Potier MC** 2008 Transmembrane protein 50b (C21orf4), a candidate for Down syndrome neurophenotypes, encodes an intracellular membrane protein expressed in the rodent brain. Neuroscience 154:1255-1266

107. **Hardy RG, Vicente-Duenas C, Gonzalez-Herrero I, Anderson C, Flores T, Hughes S, Tselepis C, Ross JA, Sanchez-Garcia I** 2007 Snail family transcription factors are implicated in thyroid carcinogenesis. Am J Pathol 171:1037-1046

108. **Hoque MO, Rosenbaum E, Westra WH, Xing M, Ladenson P, Zeiger MA, Sidransky D, Umbricht CB** 2005 Quantitative assessment of promoter methylation profiles in thyroid neoplasms. J Clin Endocrinol Metab 90:4011-4018

109. **Sunde M, McGrath KC, Young L, Matthews JM, Chua EL, Mackay JP, Death AK** 2004 TC-1 is a novel tumorigenic and natively disordered protein associated with thyroid cancer. Cancer Res 64:2766-2773

110. **de Silva CM, Reid R** 2003 Gastrointestinal stromal tumors (GIST): C-kit mutations, CD117 expression, differential diagnosis and targeted cancer therapy with Imatinib. Pathol Oncol Res 9:13-19

111. **Ronnstrand L** 2004 Signal transduction via the stem cell factor receptor/c-Kit. Cell Mol Life Sci 61:2535-2548

112. **D'Amato G, Steinert DM, McAuliffe JC, Trent JC** 2005 Update on the biology and therapy of gastrointestinal stromal tumors. Cancer Control 12:44-56

113. **McIntyre A, Summersgill B, Grygalewicz B, Gillis AJ, Stoop J, van Gurp RJ, Dennis N, Fisher C, Huddart R, Cooper C, Clark J, Oosterhuis JW, Looijenga LH, Shipley J** 2005 Amplification and overexpression of the KIT gene is associated with progression in the seminoma subtype of testicular germ cell tumors of adolescents and adults. Cancer Res 65:8085-8089

114. **All-Ericsson C, Girnita L, Muller-Brunotte A, Brodin B, Seregard S, Ostman A, Larsson O** 2004 c-Kit-dependent growth of uveal melanoma cells: a potential therapeutic target? Invest Ophthalmol Vis Sci 45:2075-2082

115. **Ulivi P, Zoli W, Medri L, Amadori D, Saragoni L, Barbanti F, Calistri D, Silvestrini R** 2004 c-kit and SCF expression in normal and tumor breast tissue. Breast Cancer Res Treat 83:33-42

116. **Natali PG, Berlingieri MT, Nicotra MR, Fusco A, Santoro E, Bigotti A, Vecchio G** 1995 Transformation of thyroid epithelium is associated with loss of c-kit receptor. Cancer Res 55:1787-1791

117. **Rosen J, He M, Umbricht C, Alexander HR, Dackiw AP, Zeiger MA, Libutti SK** 2005 A six-gene model for differentiating benign from malignant thyroid tumors on the basis of gene expression. Surgery 138:1050-1056; discussion 1056-1057

118. **Mazeh H, Mizrahi I, Halle D, Ilyayev N, Stojadinovic A, Trink B, Mitrani-Rosenbaum S, Roistacher M, Ariel I, Eid A, Freund HR, Nissan A** 2011 Development of a microRNA-based molecular assay for the detection of papillary thyroid carcinoma in aspiration biopsy samples. Thyroid 21:111-118

119. **Chang HH, Ramoni MF** 2009 Transcriptional network classifiers. BMC Bioinformatics 10 Suppl 9:S1

120. **Stojadinovic A, Peoples GE, Libutti SK, Henry LR, Eberhardt J, Howard RS, Gur D, Elster EA, Nissan A** 2009 Development of a clinical decision model for thyroid nodules. BMC Surg 9:12

121. **Liu YI, Kamaya A, Desser TS, Rubin DL** 2011 A bayesian network for differentiating benign from malignant thyroid nodules using sonographic and demographic features. AJR Am J Roentgenol 196:W598-605

122. **Jin P, Han TH, Ren J, Saunders S, Wang E, Marincola FM, Stroncek DF** 2010 Molecular signatures of maturing dendritic cells: implications for testing the quality of dendritic cell therapies. J Transl Med 8:4

123. **Zweig MH, Campbell G** 1993 Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 39:561-577

124. **Zsebo KM, Williams DA, Geissler EN, Broudy VC, Martin FH, Atkins HL, Hsu RY, Birkett NC, Okino KH, Murdock DC, et al.** 1990 Stem cell factor is encoded at the Sl locus of the mouse and is the ligand for the c-kit tyrosine kinase receptor. Cell 63:213-224

125. **Jin L, Lloyd RV, Nassar A, Lappinga PJ, Sebo TJ, Swartz K, Seys AR, Erickson-Johnson MR, Roth CW, Evers BR, Oliveira AM, Zhang J** 2011 HMGA2 expression analysis in cytological and paraffin-embedded tissue specimens of thyroid tumors by relative quantitative RT-PCR. Diagn Mol Pathol 20:71-80

126. **Ginzinger DG** 2002 Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. Exp Hematol 30:503-512

127. **Ohlsson L, Hammarstrom ML, Israelsson A, Naslund L, Oberg A, Lindmark G, Hammarstrom S** 2006 Biomarker selection for detection of occult tumour cells in lymph nodes of colorectal cancer patients using real-time quantitative RT-PCR. Br J Cancer 95:218-225

128. **Schroder CP, Ruiters MH, de Jong S, Tiebosch AT, Wesseling J, Veenstra R, de Vries J, Hoekstra HJ, de Leij LF, de Vries EG** 2003 Detection of micrometastatic breast cancer by means of real time quantitative RT-PCR and immunostaining in perioperative blood samples and sentinel nodes. Int J Cancer 106:611-618

129. **Segev DL, Clark DP, Zeiger MA, Umbricht C** 2003 Beyond the suspicious thyroid fine needle aspirate. A review. Acta Cytol 47:709-722

130. **Kim SJ, Lee KE, Myong JP, Park JH, Jeon YK, Min HS, Park SY, Jung KC, Koo do H, Youn YK** 2012 BRAF V600E mutation is associated with tumor aggressiveness in papillary thyroid cancer. World J Surg 36:310-317

131. **Kebebew E, Weng J, Bauer J, Ranvier G, Clark OH, Duh QY, Shibru D, Bastian B, Griffin A** 2007 The prevalence and prognostic value of BRAF mutation in thyroid cancer. Ann Surg 246:466-470; discussion 470-461

132. **Lee X, Gao M, Ji Y, Yu Y, Feng Y, Li Y, Zhang Y, Cheng W, Zhao W** 2009 Analysis of differential BRAF(V600E) mutational status in high aggressive papillary thyroid microcarcinoma. Ann Surg Oncol 16:240-245

133. **de Melo Martins PC, Parise Junior O, Pereira Hors C, Villela Miguel RE, da Costa Andrade VC, Garicochea B** 2007 C8orf4/TC-1 (thyroid cancer-1) gene expression in thyroid cancer and goiter. ORL J Otorhinolaryngol Relat Spec 69:127-130

134. **Sargent DJ** 2001 Comparison of artificial neural networks with other statistical approaches: results from medical data sets. Cancer 91:1636-1642

135. **Rodvold KA, Pryka RD, Kuehl PG, Blum RA, Donahue P** 1990 Bayesian forecasting of serum gentamicin concentrations in intensive care patients. Clin Pharmacokinet 18:409-418

136. **Wakefield J, Racine-Poon A** 1995 An application of Bayesian population pharmacokinetic/pharmacodynamic models to dose recommendation. Stat Med 14:971-986

137. **Burnside ES, Rubin DL, Fine JP, Shachter RD, Sisney GA, Leung WK** 2006 Bayesian network to predict breast cancer risk of mammographic microcalcifications and reduce number of benign biopsy results: initial experience. Radiology 240:666-673

138. **Christiansen CL, Wang F, Barton MB, Kreuter W, Elmore JG, Gelfand AE, Fletcher SW** 2000 Predicting the cumulative risk of false-positive mammograms. J Natl Cancer Inst 92:1657-1666

139. **Edwards FH, Schaefer PS, Cohen AJ, Bellamy RF, Thompson L, Graeber GM, Barry MJ** 1989 Use of artificial intelligence for the preoperative diagnosis of pulmonary lesions. Ann Thorac Surg 48:556-559

140. **Burd RS, Ouyang M, Madigan D** 2008 Bayesian logistic injury severity score: a method for predicting mortality using international classification of disease-9 codes. Acad Emerg Med 15:466-475

141. **Fazio VW, Tekkis PP, Remzi F, Lavery IC** 2004 Assessment of operative risk in colorectal cancer surgery: the Cleveland Clinic Foundation colorectal cancer model. Dis Colon Rectum 47:2015-2024

142. **Biagioli B, Scolletta S, Cevenini G, Barbini E, Giomarelli P, Barbini P** 2006 A multivariate Bayesian model for assessing morbidity after coronary artery surgery. Crit Care 10:R94

143. **Edwards FH, Peterson RF, Bridges C, Ceithaml EL** 1995 1988: use of a Bayesian statistical model for risk assessment in coronary artery surgery. Updated in 1995. Ann Thorac Surg 59:1611-1612

144. **Hoot N, Aronsky D** 2005 Using Bayesian networks to predict survival of liver transplant patients. AMIA Annu Symp Proc:345-349

145. **Lenihan CR, O'Kelly P, Mohan P, Little D, Walshe JJ, Kieran NE, Conlon PJ** 2008 MDRD-estimated GFR at one year post-renal transplant is a predictor of long-term graft function. Ren Fail 30:345-352

ACKNOWLEDGMENTS

First and foremost I wish to thank prof. Generoso Bevilacqua to allow me to be part of the Bios School. This experience meant a lot to me and made me grow up either scientifically and personally.

To be honest, I have to say that my time in Pisa was made enjoyable in large part due to the many friends that became a part of my life. I met a special group that has been source of friendships as well as good advice and collaboration.
Thanks to all of you: Francesca, Antonio, Claudia, Alessandro, Valerio, Maria Pia, Ivana, Sara, Federica, Mohammad, Mazhar, and all the colleagues from the other groups! Some of those people (you know who you are!) became my family and I will always be grateful to them for making me realize that, no matter what, life is just easy when you have someone like them to count on!

Thank to Katia for her advice, supervision and crucial contribution. Her support and friendship have been invaluable on both an academic and a personal level, for that I am extremely grateful.

Thank to Dr. Marchetti, his kind support has been of great value in this study.

I thank Paolo for the generosity with which he helped me with statistics, and for his willingness to muse on questions I faced. He is always ready to help with a smile.

A special thank goes to my tutor Chiara. She encouraged me to not only grow as a biologist but also as an independent thinker. She provided me unflinching encouragement and support in various ways. Her truly scientist intuition makes her a constant oasis of ideas, which exceptionally inspired and enriched my growth as a student, a researcher and a scientist. Her passion in science has triggered and nourished my intellectual maturity that I will benefit from, for a long time to come. But most importantly she made my biggest dream come true! Thanks Chiara!

Lastly, I would like to thank my enlarged family (brother and sister in law included) for all their love (masked sometimes by fake rudeness!) and encouragement. Thank to my parents who

raised me with the freedom to make my own decision and supported me in all my pursuits. And thank to Riccardo and Silvia, because the special link we share from our born made them my best friend all my life and I love them dearly.