**UNIVERSITY OF CATANIA**

---

**PhD**
**in Translational Biomedicine**
**XXXV Cycle**

# Molecular Landscape in a Cohort of Thyroid Neoplasms from the Volcanic Area of Sicily

Dott.ssa Federica Martorana

**PhD Coordinator**
Prof. Carlo Vancheri

**PhD Tutor**
Prof. Paolo Vigneri

**PhD Co-Tutor**
Dott. Michele Massimino

# INDEX

## 1. ABSTRACT

Thyroid cancer is the most frequent tumor of the endocrine system. An increased incidence of papillary thyroid cancers (PTC) has been observed in volcanic areas worldwide, due to a chronic exposure to low-level heavy metals pollution. Additionally, subjects with a history of childhood tumors, who received chemo-radiotherapy, display a higher risk to develop PTC later in life. Despite the molecular landscape of PTC has been largely explored, it remains unknown if tumors diagnosed in patients from a volcanic area present distinct genomic features. Similarly, molecular characteristics of PTC arising in individuals with an history of pediatric malignancy have been only partly elucidated, with most of the data regarding radiation-related tumors. Lastly, follicular adenomas represent benign thyroid neoplasms, whose molecular background have been scarcely investigated.

In this study, we performed a next generation sequencing analysis in a cohort of PTC and thyroid adenomas diagnosed in subjects from the volcanic area of Mount Etna, using a custom panel. Of the 22 PTC samples analyzed, 7 were from patients exposed to chemotherapy and/or radiotherapy for a pediatric cancer.

In our population, de novo PTC presented a high incidence of *BRAF* and *ZFHX3* mutations. Some cases harbored a concomitant alteration in *BRAF* and *H/KRAS*, which are mutually exclusive according to the existing literature. However, all the *KRAS* mutations involved non-coding regions of the gene and their functional significance remains controversial. As expected, 3 follicular variant PTC were RAS-driven.

Secondary PTC displayed two mutational patterns. On the one side, radiation-induced tumors showed an hypermutated profile, with a non-canonical *BRAF* fusion and an enrichment in copy number variations (CNV). On the other hand, PTC occurring in

patients exposed to chemotherapy alone presented a stable genomic profile, with few mutations, not reflecting the classical BRAF/RAS dualism. Of note, *ZFHX3* was altered in all the secondary PTC.

Lastly, adenomas showed a discrete number of alterations, mainly involving *KRAS*, *ZFHX3* and *KMT2C*, besides some unexpected CNV.

Further bioinformatic analyses about protein-protein networks, enriched biological functions and pathways showed several differences in the studied populations. While de novo PTC were mainly enriched with MAPK signaling-related functions and pathways, secondary tumors also presented an enrichment in gene expression, TP53 activation/regulation and DNA repair. Lastly, molecular alterations in thyroid adenomas also converged on the MAPK signaling, but with some unique features, such as enrichment in biological process related to cell adhesion, motility and migration and FGFR-related pathways.

Overall, our study suggest that the molecular landscape of PTC developing in subjects from a volcanic area is only partly concordant with pre-existing literature evidence. Moreover, individuals previously exposed to radiotherapy, but not those exposed to chemotherapy only, seem to have a greater risk to develop a hypermutated, and hence more aggressive, PTC. Lastly, the profile of thyroid adenomas was enriched with SNV and CNV converging on biological pathways that were substantially different to those observed in PTC. These findings may have significant clinical implications and thus deserve further validation in larger and prospective cohorts.

# 2. INTRODUCTION

## 2.1. Thyroid Cancers: General Features

Worldwide, thyroid cancers account for 2.5% of all malignancies and about 90% of tumors affecting the endocrine system (1). During 2022, in Italy, 12200 new cases of thyroid cancers have been reported (8700 in women and 3500 in men), representing 4% of all the oncological diagnoses (2). These tumors are three to four times more common in women than men, with a median age of onset between 25 and 65 years (3). Over the last decades, the incidence of thyroid malignancies has increased steadily by almost 5% per year in both sexes, to lastly reach a plateau (4, 5) (**Figure 1.1**). Improved diagnostic techniques are probably the main cause of the increased incidence of thyroid tumors, but environmental factors (e.g. exposure to radiation or pollution) and lifestyle habits may also play a role (6, 7).

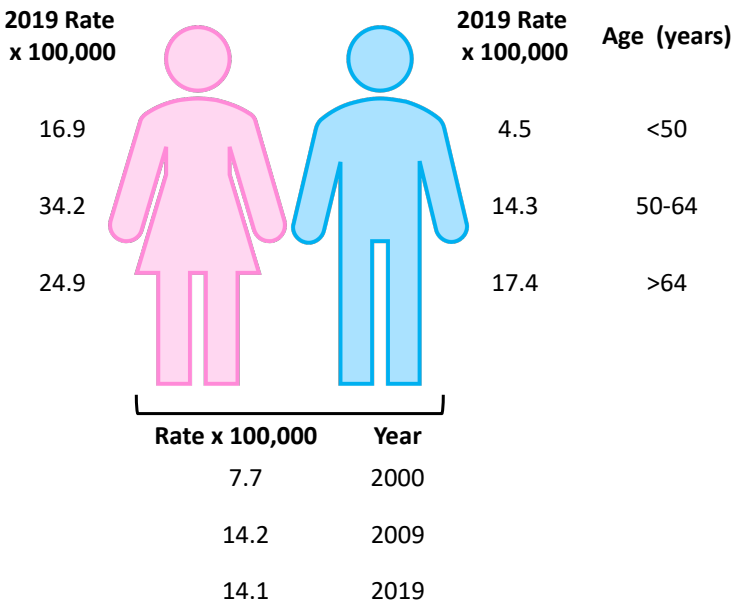



**Figure 1.1. Thyroid cancer epidemiology according to age, gender and year**
Data from the American Surveillance, Epidemiology, and End Results (SEER) database.

The overall thyroid cancer-related mortality is low, with 0.7 and 0.5 deaths per 100 000 person-years for women and men, respectively. In contrast with the increasing incidence, the mortality rate of these tumors has not significantly changed over time (8). The main prognostic factors are represented by tumor histology, with the best survival rates in patients with well-differentiated cancers and the worst for those with anaplastic tumors, and by age at diagnosis, with younger individuals displaying the best outcomes (9, 10).

Thyroid gland includes two different types of epithelial cells, with distinct embryological origins and function. Follicular cells, the main cellular component, are organized in functional units called follicles and produce thyroxine and triiodothyronine. Parafollicular-cells, also known as clear (C) cells, are a minor component with neuroendocrine origin and secrete the hormone calcitonin (11).

From these cellular populations stem different types of thyroid cancers (**Figure 1.2**). The vast majority of thyroid malignancies (>95%) derive from follicular cells and include differentiated (DTC), poorly differentiated (PDTC) and anaplastic (ATC) thyroid carcinomas. According to the most recent World Health Organization (WHO) classification, DTC are further classified into papillary (PTC), follicular (FTC), invasive encapsulated follicular variant papillary (IEFV-PTC), differentiated high grade (DHTC) and oncocytic (previously known as Hürtle cell) carcinomas (12, 13).

Papillary thyroid tumors represent the most common subtype (accounting for about 70-80% of thyroid cancers) and are further classified into several variants, including follicular, tall cell, columnar cell, hobnail cell, solid, diffuse sclerosing, Warthin-like and oncocytic subtype (14).

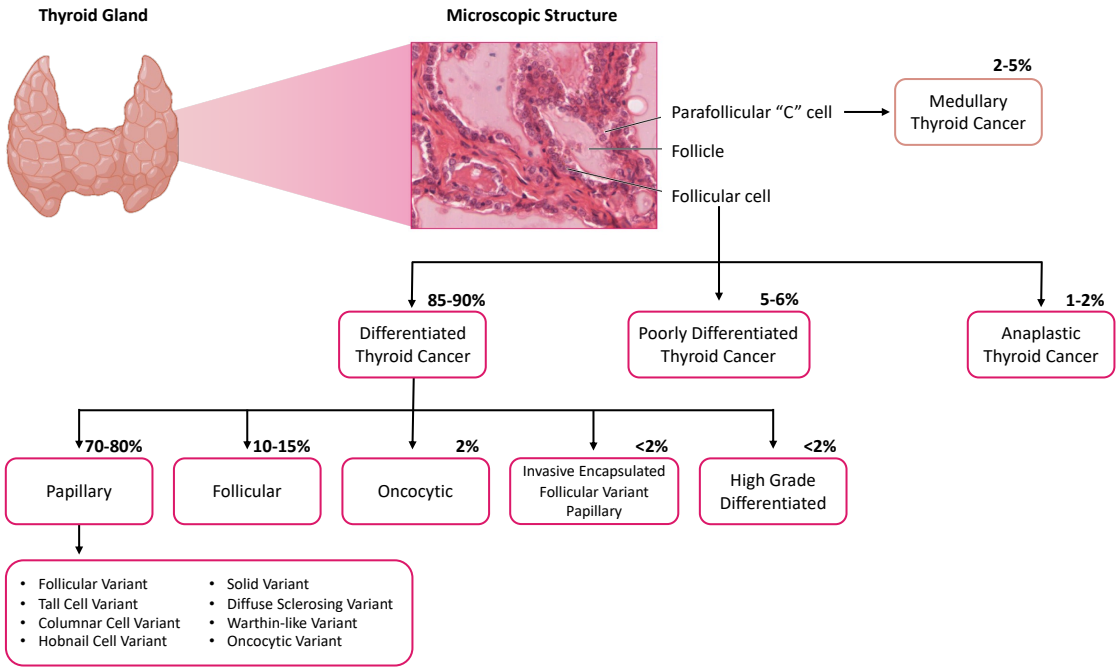



**Figure 1.2. Cellular components of thyroid epithelium and corresponding tumor types with correlated incidence rates**

Poorly differentiated tumors, accounting for 5-6% of all thyroid cancers, represent a controversially defined entity, characterized by increased proliferation rate compared to DTC but maintenance of thyroid differentiation. Additionally they typically display capsular and vascular invasion and focal tumor necrosis (15, 16).

Anaplastic tumors represents less than 2% of all thyroid cancers and are constituted by undifferentiated cells that have lost any similarity with the normal thyroid tissue (17).

Para-follicular cells give origin to medullary thyroid cancer (MTC), which represents 2-5% of all thyroid cancers (18). Consistently with their cell of origin, MTC display a neuroendocrine differentiation. They can be either sporadic, in 75% of cases, or hereditary, in the remaining 25% of cases, the latter defined as part of Multiple Endocrine Neoplasia type 2 (MEN2) syndrome (19).

From the clinical standpoint, PTC are the thyroid tumors with the best prognosis, rarely metastasizing to distant sites. Follicular thyroid cancer and Hurtle cell carcinoma are also

indolent but present a slightly higher rate of distant metastases, especially to the lungs and the bones. Poorly differentiated thyroid cancers are more aggressive, with frequent metastatic spread (12). Standard upfront treatment for DTC and PDTC involves surgery followed by adjuvant hormone replacement and radioactive iodine ($^{131}$I-based RAI) therapy for high-risk diseases (20). The majority of patients who present distant metastases at diagnosis, or develop them subsequently, are eligible to RAI, with a 40% chance of achieving a complete and durable response (20, 21). However, the remaining 60% display primary or acquired resistance to RAI thus needing further treatments, such as targeted agents (22). Currently approved targeted therapies for the treatment of RAI-refractory advanced DTC and PDTC include the small molecules multi-tyrosine kinase inhibitors lenvatinib and sorafenib (23).

Differently from DTC, ATC have an extremely aggressive and usually lethal clinical course with rapid growth, a propension for local invasion and a high rate of distant dissemination (mainly to the lungs, bones and brain) (24). Given the lack of follicular differentiation, ATC do not retain sensitivity to RAI, due to suppression of the sodium iodide symporter (NIS) expression and/or function (25). Therefore, radiotherapy and chemotherapy are the only viable therapeutic options, even though durable response are exceedingly rare . However, in selected cases presenting a $BRAF^{V600E}$ mutation the combination of BRAF inhibitor dabrafenib and MEK inhibitor trametinib can be used (23, 26). **Table 1.1** summarizes pathological and clinical features of thyroid cancers of follicular origin.

In case of MTC, surgery followed by hormone replacement is the standard treatment for localized disease, while locally advanced or metastatic MTC can be treated with targeted agents (cabozantib or vandetanib) or, less frequently, chemotherapy (27).

**Table 1.2. Pathological and Clinical Features of Thyroid Cancers of Follicular Origin**

<table>
<tr><th>Histotype</th><th>Differentiation<br>(growth pattern)</th><th>Grade<br>(mitosis, necrosis)</th><th>Prognosis</th><th>Targeted Therapies</th></tr>
<tr><td>Papillary</td><td rowspan="4">Good<br>(papillae, follicles)</td><td rowspan="3">Low</td><td rowspan="3">Excellent</td><td rowspan="4">Lenvatinib, Sorafenib</td></tr>
<tr><td>Follicular</td></tr>
<tr><td>Oncocytic</td></tr>
<tr><td>Differentiated,<br>high grade</td><td rowspan="3">High</td><td rowspan="2">Intermediate</td></tr>
<tr><td>Poorly<br>Differentiated</td><td>Poor<br>(solid, trabecular, insular)</td></tr>
<tr><td>Anaplastic</td><td>Absent<br>(undifferentiated)</td><td>Dismal</td><td>Dabrafenib + Trametinib<br>(for BRAF mutant)</td></tr>
</table>

Modified from Baloch Z.W. et al., 2022 (14).

### 2.1.1. Thyroid Cancers in Volcanic Areas

A potential correlation between a volcanic environment and an increased risk of thyroid cancer has been hypothesized since 1980s, when Kung et al noticed a high incidence of these neoplasms in Iceland and Hawaii (28). In the following decades, further observations showed high rates of thyroid cancer in other volcanic areas, including Vanuatu, French Polynesia, New Caledonia and Loyalty islands (29-31).

This evidence was then confirmed observational studies, carried out among individual from eastern Sicily (province of Catania) exposed to the volcanic environment of Mount Etna (**Figure 1.3**) (32-35). According to these studies, in the Sicilian volcanic area papillary thyroid cancers, but not follicular or medullary histotypes, display an increased incidence.


**Figure 1.3. Map of Sicily and incidence of thyroid cancer, as reported by Pellegriti G. et al., 2009**

The black area represents the basement of Mount Etna, the gray area represents the province of Catania. Adapted from Pellegriti G. et al., 2009 (32).

Despite these robust evidence, not all the volcanic environment are associated with the same high rates of thyroid cancer. Indeed, different composition of toxic compounds in gas, ash and lava emissions may explain the different risk of thyroid neoplasms in volcanic areas worldwide (Vigneri 2017) (6). Additionally, the presence of aquifer with water enriched in heavy metals - as happens in the area of Mount Etna - may represent a vehicle for contamination (36). In fact, this cause an increased metal concentration in vegetables and a significant biocontamination of the local population (**Figure 1.4**) (34, 37-39).

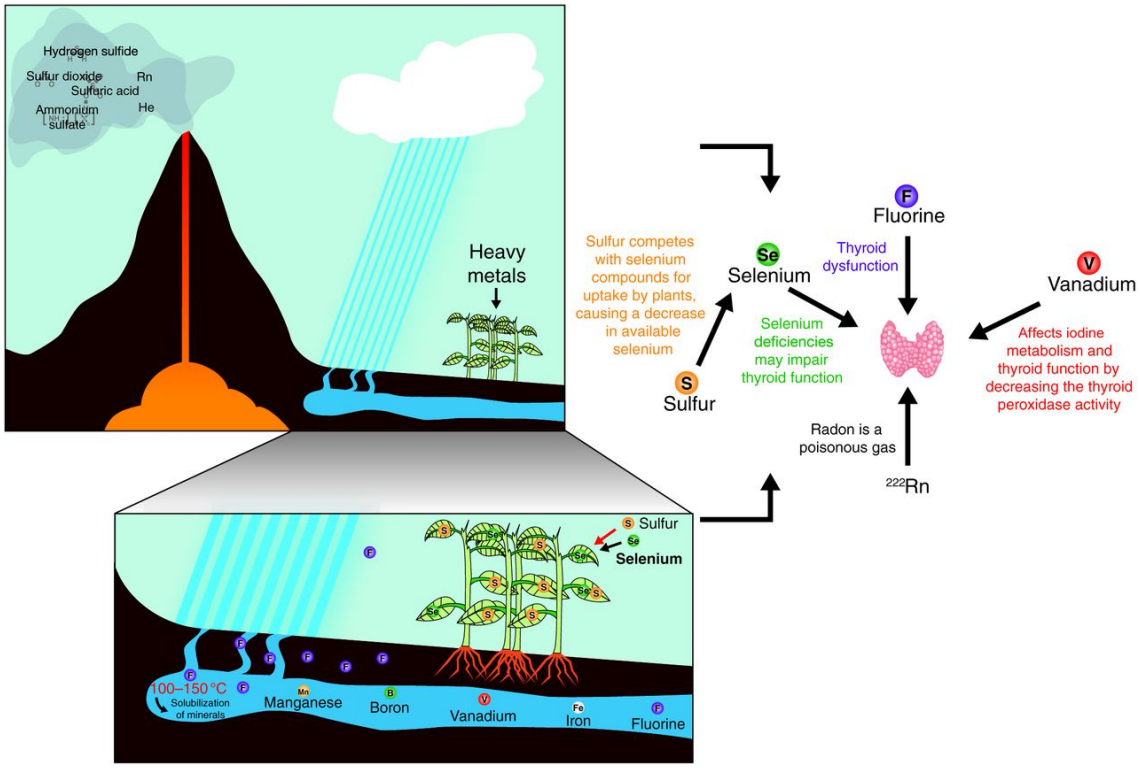



**Figure 1.4. Correlation between volcanic environment, heavy metals pollution in the air and in the acquifer and the development of thyroid cancer**
Adapted from Marcello M.A. et al, 2014 (39).

Heavy metals [e.g. Arsenic (As), Cadmium (Cd), Copper (Cu), Iron (Fe), Lead (Pb), Manganese (Mn), Mercury (Hg), Nickel (Ni), Selenium (Se), Tungsten (W), Zinc (Zn)] may exert a cancerogenic role by causing genetic and epigenetic changes in specific cell types, promoting their malignant transformation. Additionally, several metals can accumulate into the thyroid gland and interfere with hormone synthesis, secretion, metabolism and receptor interaction (6). Metal may also cause negative feedback that increases TSH secretion, that may in turn elicit thyroid cell hyperplasia and eventually thyroid carcinogenesis (40).

More recently, *in vivo* and *in vitro* studies suggested that thyroid stem cells are affected by chronic exposure to low-level heavy metal pollution (41). More in detail, long-lasting exposure to slightly increased metal concentrations may have a profound effect on

thyroid precursor cells *in vivo*. These cells then differentiate into mature thyrocytes with transformed features (42). The biological mechanisms leading to malignant transformation remain on partly understood (43). It has been demonstrated that chronic exposure to low-level increased metals promotes ERK1/2 phosphorylation, leading to a pathway activation that eventually determines immature thyroid cell dysregulated proliferation (41).

Regardless of the underlying cancerogenic mechanisms, it is unknown whether thyroid cancers observed in patients living in a volcanic area and chronically exposed to heavy metals display distinct biological features compared to those arising in subjects without this environmental risk factor.

### 2.1.2. Secondary Thyroid Cancers

It is well known that survivors from pediatric cancers have a 5 to 18-fold increased risk to develop a secondary thyroid malignancy later in life, with a cumulative incidence of about 1.5% after 30 from the primary cancer diagnosis (44-47). In these subjects, the latency between primary tumor and the development of a thyroid malignancy is variable, ranging from 0.6 to 38 years according to different observational studies (48-56). Differentiated thyroid cancers, especially papillary and follicular histotypes, are mainly observed in this population (57).

There are both epidemiological and clinical correlations linking pediatric cancers and increased incidence of secondary thyroid carcinomas. Treatment-related factors have been identified, including radiotherapy, chemotherapy and hematopoietic stem cell transplant (48). Patient-related factors can also be involved, such as female gender, younger age at diagnosis, longer follow-up since primary cancer, older survivor's age and

primary cancer diagnosis. Lastly, but not less important, genetic susceptibility may have a role in individuals with germline mutations and inherited syndromes (48).

**Treatment-related factors:** Among survivors of pediatric malignancies, previous radiotherapy involving the thyroid gland is the main risk factor for secondary differentiated thyroid cancers, particularly for PTC. It has been demonstrated that radiations exert their effect through a non-linear dose-response curve, with a steadely increased risk up to 20 Gy, followed by a rapid decline (**Figure 1.5**) (58, 59).

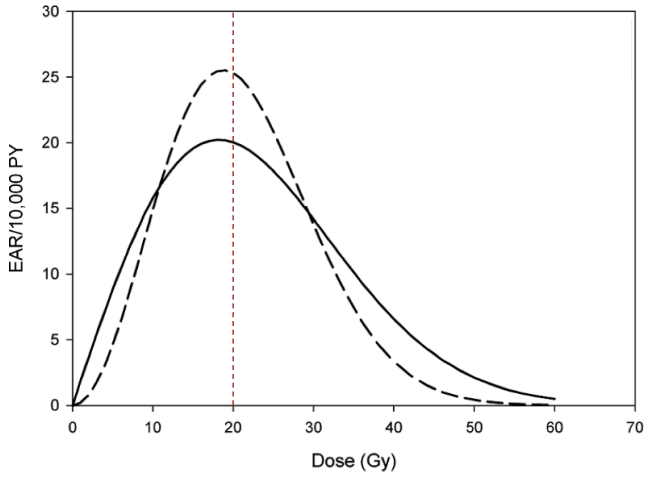



**Figure 1.5**. **Dose-response curves of thyroid cancer risk per radiation dose exposure**
The curves represent the excess absolute risk (EAR) per 10,000 person-years (PY) of thyroid cancer as a function of mean radiation dose to the thyroid in Grays (Gy) (continue line), and adjusted risk for age, sex and type of first cancer (dashed line). Modified from Batthi et al. 2010 (58).

Of, most of the studies describing radiation-related thyroid carcinomas among childhood cancer survivors involved outdated techniques (60). Modern radiotherapy, such as three dimensional-conformal radiation therapy (3D-CRT), intensity modulated RT (IMRT) and proton-beam therapy have to be investigated yet and are likely less toxic (61, 62).

The impact of chemotherapy during childhood on the development of thyroid tumors is still controversial (63). Several studies have showed an increased cancer risk among

patients treated with alkylating agents and anthracyclines, while others did not demonstrate a exceeding risk of thyroid carcinoma following exposure to chemotherapy (58, 63, 64). Furthermore, a report from Veiga et al. on the combination of chemotherapy and radiation recently demonstrated that there is no synergistic association between them (65).

Autologous and allogenic hematopoietic stem cell transplants also seem related to a higher incidence of a subsequent thyroid cancer (66). According to a cohort study involving hematopoietic cell transplant survivors, several factors may cooperate to the increased risk, including young age at the transplant, pre-transplant conditioning with total-body irradiation and chronic graft-versus-host disease. Of note, the majority of patients of this cohort have developed a follicular-type carcinoma (51).

**Patient-related factors:** The incidence of secondary thyroid tumors after childhood malignancies is not the same for every primary cancer subtype. Neuroblastoma, Hodgkin and non-Hodgkin lymphomas, brain tumors and leukemia are those associated with highest risk (**Table 1.2**) (67-69). However, these findings remain controversial.

**Tab.1.2. Risk for secondary thyroid carcinomas by childhood cancer subtype**

| Primary Childhood Cancer | SIR (95% CI) | EAR per 100.000 person-years |
|---|---|---|
| Neuroblastoma | 143.7 (29.6-419.8) | 60.8 |
| Hodgkin lymphoma | 52.5 (24.0-99.6) | 62.7 |
| Non-Hodgkin lymphoma | 40.4 (14.8-88) | 31.0 |
| Glioma | 6.8 (1.9-18) | 7.2 |
| Embryonal CNS tumors | 30 (8.2-77) | 25 |
| Other CNS tumors | 7.7 (0.2-4.3) | 8.8 |
| Leukemia | 18.8 (8.60-35.7) | 11.3 |
| Retinoblastoma | 5.3 (0.1-29.6) | 3.8 |
| Renal tumors | 13.9 (2.9-40.5) | 11.1 |
| Bone sarcomas | 10.8 (1.3-38.9) | 14.3 |
| Soft-tissue sarcomas | 4.1 (0,1-22.6) | 3.6 |
| Epithelial tumors | 3.3 (0.1-18.4) | 3.4 |

**Table Legend:** SIR standardized incidence ratio; CI confidence interval; EAR excess absolute risk, CNS central nervous system. Adapted from Wijnen et al., 2016 (57).

**Genetic susceptibility:** The growing interest for secondary malignancies among childhood cancer survivors led to investigate their correlation with genetic predisposition (48). Individual susceptibility to develop multiple primary neoplasms may be related to hereditary syndromes, such as instance Li-Fraumeni syndrome (due to *TP53* mutations), hereditary breast and ovarian cancer (due to *BRCA1/2* mutations), familial adenomatous polyposis (due to *APC* mutations), Lynch syndrome (due to mismatch repair genes mutations) and familial retinoblastoma (due to *RB1* mutations) (70-73). Moreover, different studies correlated other gene variants with higher risk of secondary tumors (74). It has been reported that some variants of *PRDM1* predispose Hodgkin lymphoma survivors to radiation-induced thyroid carcinomas (75). In addition, *TP53*, *ATM* and *FOXE1* mutations seem contribute to the development of radiation-induced thyroid carcinomas among Chernobyl survivors and childhood cancer survivors

which (50, 74, 76, 77). On the other hand, Akulevich et al. reported that *ATM* Asp1853Asn and *XRCC1* Arg399Gln may have a protective role on the development of secondary thyroid cancers (74). Lastly, an inverse correlation between telomere content and radiation-related thyroid cancer has been observed, with shorter telomeres probably involved in radiogenic PTC pathogenesis (57).

### 2.1.3. Molecular Landscape of Papillary Thyroid Cancers

In 2014, The Cancer Genome Atlas (TCGA) initiative provided a comprehensive genomic characterization of almost 500 PTC (78). Thanks to this considerable effort, the molecular landscape of PTC has been largely unrevealed. In general, PTC biology is mainly driven by alterations of mitogen activated protein kinase (MAPK) signaling and can be classified into two categories according to the presence of mutually exclusive driver mutations involving *BRAF* (more frequently V600E substitutions) or *RAS* (78) (**Figure 1.6**).

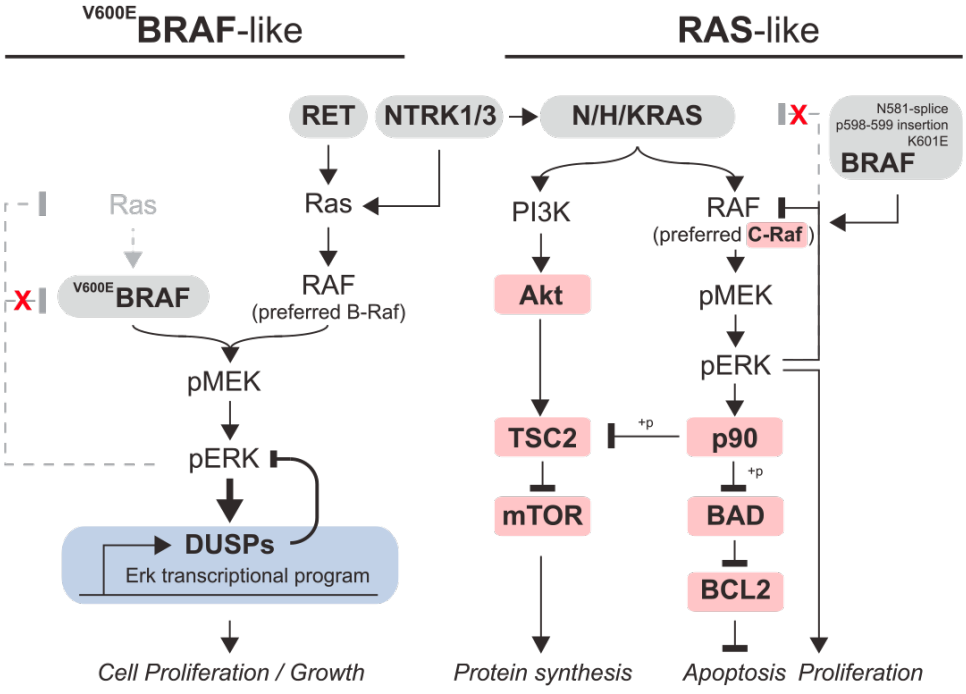


**Figure 1.6. Molecular classification of PTC according to TCGA. Adapted from (78)**

The *BRAF*-mutant tumors are the most frequent, usually present a PTC conventional histology and predominantly signals through MAPK pathway, resulting in loss of differentiation, tumor progression and inhibition of apoptosis (79, 80). Conversely, RAS-mutated tumors, which account for 5% of all PTC, are mostly follicular variants and have two main signaling pathways, MAPK and PI3K (81, 82).

However, *BRAF* and *RAS* mutations are not the only driver events occurring in PTCs. Several single nucleotide variations (SNV), insertions and deletions (INDEL) have been reported to play a role in PTC pathogenesis and progression (83). Among significantly mutated genes found in PTCs by TCGA group there are *EIF1AX*, *PPM1D* and *CHK2*. Overall, alterations in other cancer-related genes that can be gathered in four dinstict functional groups (78):

- Genes involved in epigenetic regulation (*ARID1B*, *EZH1*, *KMT2A*, *KMT2C*);
- Genes involved in intracellular signaling pathways (PI3K/AKT, PPARγ and WNT pathway);
- Tumor suppressor genes (*TP53*, *RB1*, *NF1*, *NF2*, *MEN1*, *PTEN*, *ZFHX3* and *BDP1*);
- Thyroid related genes (*TG*, *TSHR*).

Beside SNV and INDEL, gene fusions leading to the activation of MAPK pathway also display a pivotal role in tumorigenesis and progression of PTC. The most common fusions involve *RET* and more in detail *RET*/*PTC1* (60% of RET-rearranged PTCs), *RET*/*PTC3* (30%), and more rarely *RET*/*PTC2*. These events determine the loss of RET transmembrane domain and the subsequent cytosolic localization of the protein (84). Additionally, fusions of *PPARG*, *ALK* and *NTRK* have been identified as rare but possible events (85).

Copy number variations (CNVs) have been seldom observed in PTC, with exception of chromosome 22q deletion (a region including *NF2* and *CHEK2*), and occurred more frequently in follicular variant PTC (86).

Overall, the genomic portrait of PTC provided by TCGA shows an usually quiet and stable disease, at least in its initial, well-differentiated stages (15, 78).

However, the information provided by TCGA refer to a de novo PTC population, since only a slight minority of the cases presented a previous tumor and radiation exposure.

Most of the available studies report information about the molecular profile of radiation-induced thyroid cancers, either in the post-radiotherapy setting or in Chernobyl accident survivors (87-89). According to these evidence, radiation-induced thyroid tumors present alterations in many genes involved in different pathways, such as cellular response to oxidative stress and irradiation, response to hypoxia, regulation of p53 function, immune response, MAPK, EGFR, RAC/CDC42, hedgehog, TGF/BMP, calcium signaling and WNT pathway (90, 91). Recently, a comprehensive genomic, transcriptomic and epigenomic characterization of papillary thyroid cancers in a large cohort of subjects from the Chernobyl area (n=440) has been published. The analyzed tumors showed a dose-dependent enrichment in fusions (involving MAPK-related genes) and a high rate of DNA double-strand breaks. Transcriptomic and epigenomic profile was consistent with those of the sporadic counterpart and the retrieved alterations was not correlated to the radiation dose (89). This study confirms previous evidence, reporting a higher rate of fusions radio-induced thyroid cancers.

Indeed, *RET* rearrangements account for up to 80% of known molecular alterations in radiation-induced PTC and seem to share the same characteristic of those found in sporadic tumors (92, 93). *BRAF* fusions may also occur in radiation-associated PTC, either

supporting conservation of its kinase domain and its expression or through alternative mechanisms (94, 95). Furthermore, pediatric PTC and radiation-induced PTCs may harbor *NTRK3* or *NTRK1* (96-98). These fusion proteins activate downstream signaling through PI3K, MAPK and PLC-γ, controlling cell-cycle progression, proliferation, apoptosis and survival (99).

Despite secondary thyroid carcinomas have attracted an increasing interest over the years, these entities have been poorly investigated under a biological standpoint, with most of the studies involving subjects exposed to radiation only. Hence additional investigations aimed at characterizing the genetic features of thyroid cancer in childhood cancers survivors are still needed.

### 2.2. Thyroid Adenomas: General Features

The term follicular adenoma (thereinafter adenoma) define a benign thyroid lesion of follicular origin without capsular and vascular invasion (100). It is a common entity, affecting up to 5% of the population according to an old autoptic series (101), but its incidence is even higher (up to 10%) when considering only small-sized nodules below 15 mm (102). As for thyroid carcinomas of follicular origin, exposure to radiation represents a risk factor for the development of follicular adenomas, as demonstrated by their increased incidence in the area of Chernobyl (103).

Until 2022, follicular adenoma was the only recognized thyroid benign lesions of follicular origin. However, the last WHO classification of thyroid neoplasms introduced the terms follicular adenoma with papillary architecture (previously known as papillary adenomatous/hyperplastic nodule) and oncocytic adenoma (13, 14). The first one presents an intrafollicular papillary growth and is often associated with autonomous

hyperfunctioning (i.e. a warm or hot nodule at nuclear imaging) (104). Additionally, follicular adenomas with papillary architecture can be associated with McCune-Albright syndrome (due to germline *GNAS* mutations), and Carney complex (due to germline mutations in *PRKAR1A*) (105, 106). On the other hand, oncocytic adenomas are defined by the presence of at least 75% of cells with oncocytic features (i.e. abundant eosinophilic granular cytoplasm) and without PTC nuclear characteristics (107).

Given their benign nature adenomas lack of metastatic potential and they can be monitored conservatively with no further interventions. Follow-up is required to identify falsely benign nodules at the time of biopsy and those that eventually evolve from benign into malignant over time (108).

Current guidelines do not recommend surgical excision of adenomas, unless the nodule is larger than 4 cm, causes symptoms or the resection is required for diagnostic purposes (109, 110).

### 2.3. Molecular Landscape of Thyroid Adenomas

Compared to the extensively studied genomic profile of thyroid cancers, few reports exist about the molecular features of thyroid adenomas.

Duan et al. performed a Next Generation Sequencing (NGS) analysis on a cohort of follicular-derived thyroid neoplasms, including 48 adenomas, using custom panel to detected single nucleotide variations (SNV) in 18 genes and rearrangements in 4 additional genes. Recurrently mutated genes retrieved in adenomas but not in borderline malignant or malignant neoplasms included *EIF1AX* and *TSHR* (111). Mutations of *NRAS* and *HRAS* were present in all the examined histological variants. Similarly, Madsen et al. reported the results of NGS sequencing using a 48-genes

commercially available panel in a cohort of thyroid neoplasms of follicular origin, including 33 adenomas. The analysis showed a very low mutational load, with most of the adenomas not presenting any mutations. Recurrent alterations included SNV of *HRAS*, *NRAS* and *TP53* (112).

Several study aimed at the identification of molecular similarities and differences between benign adenomas and their malignant counterpart. Older reports, dated back to mid-2000, seem to suggest that adenomas and thyroid cancers display distinct molecular hallmarks and that this difference could be exploited for differential diagnosis between the two entities (113-115). By contrast, more recent analyses showed partly overlapping features between benign adenomas and FTC, with alterations involving mainly RAS-related genes and *PPARG* fusions (116-118).

Overall, the molecular features of adenomas recapitulate both their benign nature and their follicular origin. However, further characterization using modern high-throughput techniques may be useful to better understand the biology of this entity and to shed light into its evolution toward malignant lesions.

## 3. STUDY AIMS

The ***overall goal*** of the present work is to provide a comprehensive molecular characterization of a cohort of thyroid neoplasms (benign and malignant) from patients exposed to a volcanic environment, using a custom NGS panel, and to compare our findings with pre-existing literature data. Besides describing the retrieved molecular alterations, we aim to examine the altered biological processes and pathways using several bioinformatic tools.

Our population of thyroid malignancies is divided into two groups. The first one comprises de novo PTC, i.e. tumors arising without known pre-existing risk factors. The second encompasses tumors occurring in subjects previously exposed to radiotherapy and/or chemotherapy for other hematological or solid cancers (i.e. secondary PTCs). Hence, ***the first aim*** of our study is to compare the molecular features of these populations. Indeed, we hypothesize that secondary PTC may exhibit a distinct biological profile, due to pre-existing exposure to cancerogenic agents.

Moreover, we carried out the same sequencing on a cohort of surgically excised thyroid adenomas. Therefore, a ***second aim*** of this work is to highlight potential molecular drivers of malignancy in benign thyroid nodules. We speculate that a comprehensive analysis of altered genes, biological processes and pathways in adenomas may provide new insight into the clonal evolution from benign to malignant thyroid lesions.

# 4. MATERIALS AND METHODS

## 4.1. Samples Collection

The analyzed samples were obtained from thyroid surgeries performed between 2012 and 2022 for the presence of benign, suspicious malignant or malignant nodules. Patients with de novo and secondary PTC underwent surgery at the Hospitals ARNAS "Garibaldi Nesima" or A.O.U. Policlinico "G. Rodolico - S. Marco", in Catania. All the patients diagnosed with thyroid adenomas underwent surgery at A.O.U. Policlinico "G. Rodolico - S. Marco", in Catania. All of them signed informed consent for the study, which was approved by the Ethics Committee 1.

Formalin-fixed, paraffin embedded (FFPE) samples were internally reviewed in order to confirm the previous histological diagnosis or update it according to the last WHO classification of thyroid tumors.

A dedicated pathologist prepared 8 sections 5-micron thick, using a standard microtome. Tumor content for each sample was determined from hematoxylin and eosin-stained slides.

DNA and RNA extraction and the subsequent molecular analyses were performed at Center for Experimental Oncology and Hematology (COES) of A.O.U. "G. Rodolico - S. Marco", Catania.

## 4.2. DNA and RNA Isolation

Formalin-fixed paraffin-embedded tissue was used to isolate both DNA and RNA. DNA was extracted using QS GeneRead DNA FFPE treatment Kit followed by automatic platform QiIAsymphony (both from Qiagen), while MagMax FFPE ultra kit (ThermoFisher Scientific) was used for RNA isolation. Both nucleic acids were quantified by fluorometer

Qubit 3.0 using Qubit™ dsDNA and RNA HS kit (all from ThermoFisher Scientific) including standards solutions to perform the calibration of the instrument.

### 4.3. Custom NGS Panel Design and Libraries preparation

Ion Ampliseq Designer tool (https://www.ampliseq.com/login/login.action) was used to generate 3 primer pools targeting coding DNA sequencing (CDS) of following 25 genes:

- *APC*, *ARID1B*, *ATM*, *BDP1*, *BRAF*, *CDH4*, *CHEK2*, *EIF1AX*, *EZH1*, *HRAS*, *KRAS*, *NRAS*, *MEN1*, *KMT2A*, *KMT2C*, *NF1*, *PPMID*, *PTEN*, *RB1*, *RET*, *SPOP*, *TG*, *TP53*, *TSHR*, *ZFHX3*.

Additional information about genes included in the DNA panel are provided in **Supplementary Table.1**.

The same tool was used to generate 2 primer pools recognizing fusion regions involving the following 7 genes:

- *RET*, *BRAF*, *NTRK1*, *NTRK3*, *THADA*, *PPARG* and *ALK*.

For DNA and RNA libraries preparation, targeted regions were amplified using Ion Ampliseq library kit plus (ThermoFisher Scientific) with 30 ng of DNA (10 ng per primer pool) and 200ng of RNA (100ng per primer pool) which was previously converted in cDNA by Vilo IV (ThermoFisher Scientific Scientific). Both libraries were barcoded using the Ion Xpress barcode adapter kit (ThermoFisher Scientific) (**Figure 3.1**).

Libraries were then quantified by qPCR with the Library TaqMan Quantification kit (ThermoFisher Scientific) and diluted in an equimolar concentration.

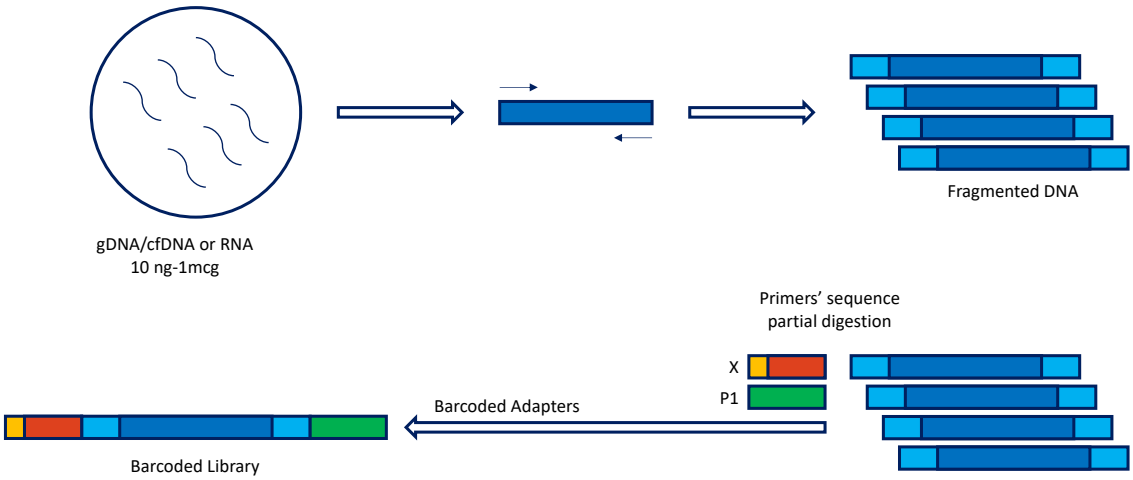



**Figure 3.1. Barcoded Library Preparation**

Next, libraries enrichment, IonChip loading, and sequencing have been performed employing Ion 510™ & Ion 520™ & Ion 530™ Kit-Chef using automatic platforms IonChef and S5 Ion Plus sequencer (all from ThermoFisher Scientific) (**Figure. 3.2**)

### 4.4. Data Analysis

Sequencing raw data were aligned using hg19 (GRCh37) DNA reference genome and the Binary Alignment Map (BAM) files generated by Torrent Suite Software v.5.12.3 (ThermoFisher Scientific) (**Figure. 3.2**). The same bioinformatic platform was also used for quality control according to the chip density loading and number of mapped reads. Each BAM file was then analyzed employing Ion Reporter v.5.18.4.0 for Single Nucleotide Variations (SNV), Copy Number Variations (CNV) and fusion genes calling. For SNVs identification, each variant was filtered according to 5% VAF, read depth>100, Phred quality score >40 and a p value <0.001.

To identify somatic CNV, a reference baseline consisting of ten male subjects without chromosome imbalance, has been generated, loaded in Ion Reporter Software and a metric value. To confirm that a specific region can be used for CNV analysis, MAPD (Median of the Absolute values of all Pairwise Differences) was applied and set to <0.45. For fusion genes, targeted breakpoint regions have been considered “true” when >20.000 reads per primers pool have been obtained from expression control genes.
Tile plots have been generated for SNVs, CNAs and fusion genes to show the molecular landscape for each analyzed population.

### 4.5. Clinical Variant and in Silico Mutation Prediction Tools

The SNV have been classified as pathogenetic according to database Catalog of Somatic Mutation in Cancer (COSMIC). For variants showing unknown significance (VUS) or not reported in COSMIC, the in silico mutation prediction tool POLYmorphism PHENotyping (Polyphen-2) was used, in order to classify these variants as pathogenetic or benign according to their impact on the structure and function of the proteins (119).

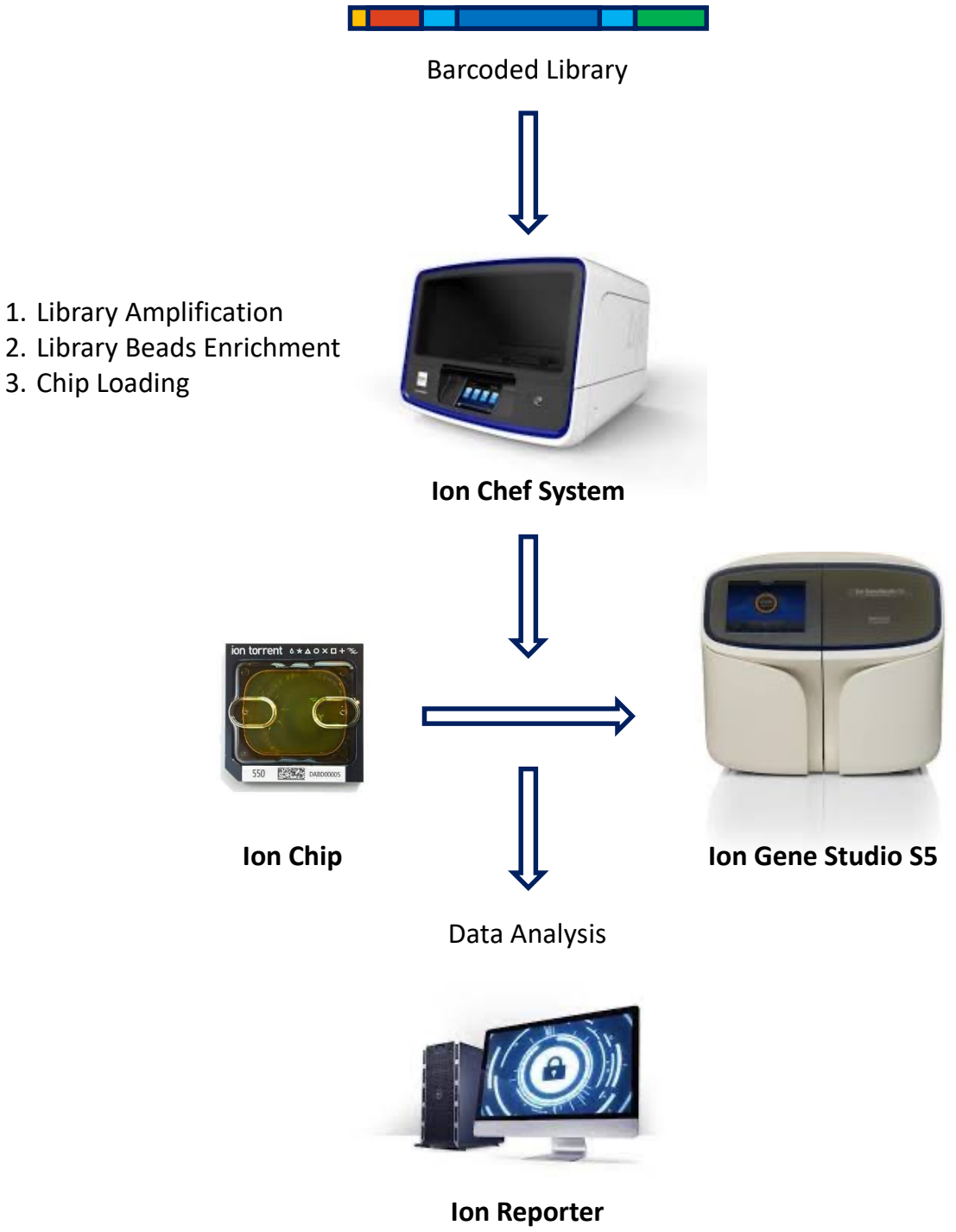



**Figure 3.2. Automated template preparation, chip loading, sequencing and data analysis**

### 4.6. Protein-Protein Networks and UMAP Plot Generation

To generate the protein-protein interaction network (PPI) and cluster analysis, STRING database v11.5 was used, limiting species to “Homo sapiens” and the confidence score to >0.4 (120). Markow cluster algorithm (MCL) was applied for cluster generation. The online EnrichR tool used to classify the Gene Ontology biological process and Reactome pathways of the PPI clusters constructed by STRING (121).

To construct bidimensional Uniform Manifold Approximation and Projection (UMAP) scatter plots for both biological process and pathways, the Appyter tool was used (122).

# 5. RESULTS

## 5.1. Population Characteristics

### 5.1.1. De Novo Papillary Thyroid Cancers

In total, 15 specimens from primary PTCs were analyzed. **Table 5.1** summarizes population's characteristics. Most of the patients were female (60%), with a mean age of 53.7 years (min 14, max 72). Twelve patients had a conventional PTC histology and 3 an infiltrative follicular variant (FV) according to 2022 WHO classification (13). All the patients were from Eastern Sicily and lived in the volcanic area of Mount Etna.

**Table 5.1. Characteristics of Patients with De Novo Papillary Thyroid Cancers**

| Patient ID | Gender | Age at Diagnosis | Histological Variant |
|---|---|---|---|
| P01 | Male | 59 | PTC |
| P02 | Male | 56 | PTC |
| P03 | Female | 27 | PTC |
| P04 | Female | 54 | PTC |
| P05 | Female | 69 | PTC |
| P06 | Female | 54 | PTC |
| P07 | Female | 69 | PTC |
| P08 | Female | 70 | PTC |
| P09 | Female | 54 | PTC |
| P10 | Male | 67 | PTC |
| P11 | Male | 68 | PTC |
| P12 | Male | 72 | PTC |
| P13 | Female | 29 | FV PTC |
| P14 | Male | 44 | FV PTC |
| P15 | Female | 14 | FV PTC |

**Table Legend**: PTC papillary thyroid cancer; FV follicular variant.

### 5.1.2. Secondary Papillary Thyroid Cancers

Seven specimens from PTC occurring in subjects previously exposed to radiotherapy and/or chemotherapy for a pediatric tumor were included in the study (**Table 5.2**). All but one patient were female and all of them were young adults at the time of PTC diagnosis, with an average age of 31.4 years (min 25, max 38). The histological variant was conventional PTC in all cases. Hodgkin lymphoma was the most common previous malignancy, occurring in 4 cases, followed by two cases of acute leukemia (one myeloid and one lymphoblastic) and one brain tumor (astrocytoma). The mean age of onset of previous malignancy was 12.3 years (min 6, max 16). The average latency between the primary pediatric tumor and the development of thyroid cancer was 18.6 years (min 12, max 27). Four patients have been exposed to radiotherapy and all but one received chemotherapy. Even in this case, all the patients were from Eastern Sicily and lived in the volcanic area of Mount Etna.

**Table 5.2. Characteristics of Patients with Secondary Papillary Thyroid Cancers**

| Patient ID | Gender | Age at Diagnosis | Histological Variant | Previous Malignancy | Age at Prev Malignancy | Prev RT | Prev CHT |
|---|---|---|---|---|---|---|---|
| S01 | Female | 37 | PTC | Astrocytoma | 8 | Yes | No |
| S02 | Female | 28 | PTC | Hodgkin Lymphoma | 13 | Yes | Yes |
| S03 | Female | 25 | PTC | Acute Myeloid Leukemia | 10 | No | Yes |
| S04 | Male | 38 | PTC | Hodgkin Lymphoma | 19 | No | Yes |
| S05 | Female | 33 | PTC | Hodgkin Lymphoma | 16 | No | Yes |
| S06 | Female | 26 | PTC | Hodgkin Lymphoma | 14 | Yes | Yes |
| S07 | Female | 33 | PTC | Acute Lymphoblastic Leukemia | 6 | Yes | Yes |

**Table Legend**: Prev previous; PTC papillary thyroid cancer; RT radiotherapy; CHT chemotherapy.

### 5.1.3. Thyroid Adenomas

Sixteen samples of thyroid adenomas were included in the analysis. As reported in **Table 5.3**, 14 of them (87.5%) derived from female patients. Mean age was 52.5 years (min 31, max 75). Again, all the subjects lived in the volcanic area of Mount Etna in Eastern Sicily.

**Table 5.3. Characteristics of Patients with Thyroid Adenomas**

| Patient ID | Gender | Age at Diagnosis |
|---|---|---|
| A01 | Female | 66 |
| A02 | Female | 40 |
| A03 | Female | 53 |
| A04 | Female | 57 |
| A05 | Female | 40 |
| A06 | Female | 52 |
| A07 | Female | 41 |
| A08 | Female | 39 |
| A09 | Female | 75 |
| A10 | Female | 62 |
| A11 | Famale | 31 |
| A12 | Female | 37 |
| A13 | Female | 66 |
| A14 | Female | 42 |
| A15 | Male | 56 |
| A16 | Male | 68 |

## 5.2. Molecular Analysis

### 5.2.1. De Novo Papillary Thyroid Cancers

In the de novo PTC population, we identified 51 known alterations in 16 genes (**Table 5.4, Figure 5.1**). Of these, the vast majority (94%) were SNV, with only two fusions (4%) and one CNV (2%). The average number of alterations per patient was 3.4.

The proto-oncogene *BRAF* displayed the highest number of SNV (n=9), all of which determined the classical substitution of a valine by glutamate at codon 600 (V600E). *KRAS* was the second most frequently mutated gene (n=7) and all the retrieved SNV involved non-coding regions. *KMT2C*, codifying for a histone methyltransferase, and *ZFHX3*, encoding for a zinc-finger transcription factor, harbored 5 SNV each. Other genes with recurrent mutations were *HRAS* and *RET* (3 SNV each), followed by *APC*, *ATM*, *MEN1*, *NF1*, *TG* and *TP53* (2 SNV each). The remaining genes (*CDH4*, *KMT2A*, *NRAS* and *TSHR*) were mutated once only in our de novo PTC population.

The only CNV observed was a copy number gain of *BRAF* gene (chromosome 7), occurring in a male patient diagnosed with PTC (P11).(**Supplementary Figure 1**).

Both the identified rearrangements involved *RET* and *NCOA4* (also known as *PTC3*). **Supplementary Figures 8** and **Supplementary Figures 9** show the details of the fusion regions. Of note, one of these fusions was observed in a patient with FV PTC (P15).

**Table 5.4. Molecular Alterations of De Novo Papillary Thyroid Cancers**

| Gene Name | Alteration n° (%) | SNV n° (%) | AA substitution | CNV n° (%) | CNV | Fusion n° (%) | Fusion |
|---|---|---|---|---|---|---|---|
| ***APC*** | 2 (4%) | 2 (4%) | p.Arg230Cys<br>p.Gly2502Ser | NR | - | NA | - |
| ***ATM*** | 2 (4%) | 2 (4%) | p.Pro1054Arg | NR | - | NA | - |
| ***BRAF*** | 10 (20%) | 9 (19%) | p.Val600Glu | 1 (100%) | Gain | NR | - |
| ***CDH4*** | 1 (2%) | 1 (2%) | p.Val520Leu | NR | - | NA | - |
| ***HRAS*** | 3 (6%) | 3 (6%) | p.Gly13Arg | NR | - | NA | - |
| ***KMT2A*** | 1 (2%) | 1 (2%) | p.Gly909Asp | NR | - | NA | - |
| ***KMT2C*** | 5 (10%) | 5 (10%) | p.Phe3226Leu<br>p.Arg284Gln<br>p.Asp348Asn | NR | - | NA | - |
| ***KRAS*** | 7 (14%) | 7 (16%) | NCR | NR | - | NA | - |
| ***MEN1*** | 2 (4%) | 2 (4%) | p.Arg337Cys<br>p.Arg176Gln | NR | - | NA | - |
| ***NF1*** | 2 (4%) | 2 (4%) | p.Ser302Asn<br>NCR | NR | - | NA | - |
| ***NRAS*** | 1 (2%) | 1 (2%) | p.Gln61Arg | NR | - | NA | - |
| ***RET*** | 5 (8%) | 3 (6%) | p.Met1009Thr<br>p.Arg982Cys | NR | - | 2 (100%) | *NCOA4/RET* |
| ***TG*** | 2 (4%) | 2 (4%) | p.Gly67Ser<br>p.Thr621Met | NR | - | NA | - |
| ***TP53*** | 2 (4%) | 2 (4%) | p.Arg273Leu<br>p.Arg158His | NR | - | NA | - |
| ***TSHR*** | 1 (2%) | 1 (2%) | p.Arg109Gln | NR | - | NA | - |
| ***ZFHX3*** | 5 (10%) | 5 (11%) | p.Ser3349Leu<br>p.Gly585Ser<br>p.Thr428Pro<br>p.Asp265Gly | NR | - | NA | - |
| **Total (%)** | **51 (100%)** | **48 (94%)** | - | **1 (2%)** | - | **2 (4 %)** | - |

**Table Legend**: SNV single nucleotide variation; AA amino acid; NCR non-coding region; CNV copy number alteration; NR not retrieved; NA not assessed

No significant differences emerged between female and male subjects, with a mean alterations number of 3.8 and 3.4, respectively. In terms of alteration types, the *BRAF* copy number gain occurred in a man, while both the *NCOA4/RET* fusions occurred in women. However, the small number of patients does not allow to drive any conclusion about gender differences in our population.

Similarly, comparisons between PTC and FV PTC should be interpreted cautiously. However, it should be noted that none of the 3 FV PTC showed *BRAF* SNV, while these samples were rather enriched in *KRAS*, *HRAS* and *NRAS* mutations.

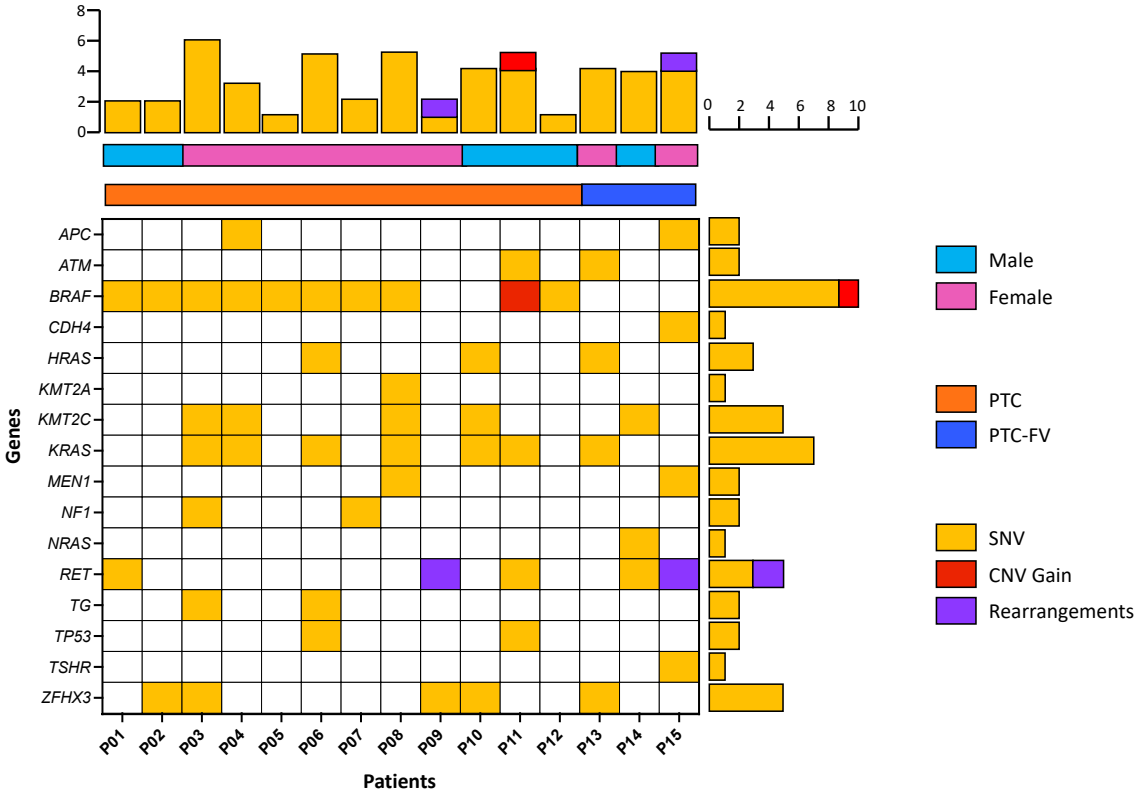



**Figure 5.1. Tile Plot of Molecular Alterations in De Novo Papillary Thyroid Cancers**
The tile plot shows the mutational landscape of the examined de novo papillary thyroid cancer samples. The color of each tile represents a different alteration type. Patients' gender and tumor's histological variant are displayed in the horizontal bars. The numbered lines indicate the total alterations retrieved in each sample (vertical) and for each gene (horizontal). **Figure Legend:** PTC papillary thyroid cancer; PTC FV papillary thyroid cancer follicular variant; SNV single nucleotide variation; CNV copy number variation.

### 5.2.2. Secondary Papillary Thyroid Cancers

Overall, 65 alterations in 24 genes were identified by sequencing the 7 samples from secondary PTC (**Table 5.5**, **Figure 5.2**). Of these, 90% were SNV (n=59), 6% CNV (n=4), 1 deletion (2%) and 1 fusion (2%). The average number of alterations per patient was 9.3. *ZFHX3* was the gene with the highest SNV number in this population (n=6), resulting in 6 different aminoacidic substitutions. *KMT2C* and *ATM* harbored 5 mutations each, followed by *APC* (n=4). Other genes repeatedly mutated were *BRAF* (2 canonical V600E SNV and one non canonical A320T), *HRAS*, *KMT2A*, *KRAS*, *NF1* and *RET* (n=3). *ARID1B*,

*BDP1*, *CDH4*, *EZH1*, *MEN1*, *SPOP*, *TG* and *TP53* were mutated twice, while *CHEK2*, *PPM1D*, *PTEN*, *RB1* and *TSHR* harbored one SNV only. Besides being the most mutated genes, *ZFHX3* also harbored the only INDEL identified in this population (p.Gln3203_Gln3204del). Hence, *ZFHX3* was altered in all the 7 subjects.

As shown in **Supplementary Figure 2** and **Supplementary Figure 3**, copy number gains were retrieved in *ATM* (chromosome 11), *TP53* (chromosome 17) and *ZFHX3* (chromosome 16), while a copy number loss was observed in *EIF1AX* (chromosome X), codifying for an eukaryotic translation initiation factor.

One fusion between *AGK* and *BRAF* was also present (**Supplementary figure 10**).

Given only one male patient was present in this cohort, inter-gender disparities cannot be scored.

Looking at previous exposure to anticancer treatment, it can be noticed that subjects with a history of radiation exposure with or without chemotherapy display a significantly higher number of molecular alterations, compared to those exposed to chemotherapy alone. Indeed, taken together the SNV found in patients S01, S02, S06 and S07 accounted for 86% of the total SNV number in this population. Additionally, the 4 CNV, the *ZFHX3* deletion and the *BRAF* fusion were all retrieved in these samples. However, while S01, S06 and S07 may be considered hypermutated subjects, given the number of alterations displayed (mean=18.3); S02 harbors only 2 SNV and 1 deletion.

Apart from *ZFHX3*, recurrently altered genes in the radiotherapy-exposed population are *ATM* and *KMT2C,* followed by *APC*, *HRAS*, *KTM2A*, *NF1* and *RET*. Subjects previously exposed to chemotherapy only (S03, S04, S05) had a significantly lower number of alteration (mean=2.6), which were all SNV.

Lastly, the type of previous malignancy did not seem to influence the mutational profile of these patients (data not shown).

### Table 5.5. Molecular Alterations of Secondary Papillary Thyroid Cancers

| Gene Name | Alteration n° (%) | SNV n° (%) | AA substit1ution | INDEL n° (%) | INDEL | CNV n° (%) | CNV | Fusion n° (%) | Fusion |
|---|---|---|---|---|---|---|---|---|---|
| ***APC*** | 4 (5%) | 4 (7%) | p.Ser2129Leu<br>p.Gly2502Ser<br>p.Asp1659Gly<br>p.Leu1129Ser | NR | - | NR | - | NA | - |
| ***ARID1B*** | 2 (3%) | 2 (3%) | p.Asp1351Asn | NR | - | NR | - | NA | - |
| ***ATM*** | 6 (8%) | 5 (9%) | p.Asp1853Asn<br>p.Arg2763Ter | NR | - | 1(25%) | Gain | NA | - |
| ***BDP1*** | 2 (3%) | 2 (3%) | p.Gln1676Glu<br>p.Arg400Gln | NR | - | NR | - | NA | - |
| ***BRAF*** | 4 (5%) | 3 (5%) | p.Ala320Thr<br>p.Val600Glu | NR | - | NR | - | 1(100%) | *AGK/BRAF* |
| ***CDH4*** | 2 (3%) | 2 (3%) | p.Ser195Phe<br>p.Asp262Asn | NR | - | NR | - | NA | - |
| ***CHEK2*** | 1 (2%) | 1 (2%) | p.Ala392Val | NR | - | NR | - | NA | - |
| ***EIF1AX*** | 1 (2%) | NR | - | NR | - | 1(25%) | Loss | NA | - |
| ***EZH1*** | 2 (3%) | 2 (3%) | p.Asp181Asn<br>p.Arg609His | NR | - | NR | - | NA | - |
| ***HRAS*** | 3 (5%) | 3 (5%) | p.Ser106Leu<br>p.Arg128Gln | NR | - | NR | - | NA | - |
| ***KMT2A*** | 3 (5%) | 3 (5%) | p.Arg302Gln<br>p.Glu1060Asp<br>p.Glu3451Lys | NR | - | NR | - | NA | - |
| ***KMT2C*** | 5 (7%) | 5 (9%) | p.Arg3077Cys<br>p.Phe3226Leu<br>p.Arg1705Cys<br>p.Arg4690Gln | NR | - | NR | - | NA | - |
| ***KRAS*** | 3 (5%) | 3 (5%) | NCR | NR | - | NR | - | NA | - |
| ***MEN1*** | 2 (3%) | 2 (3%) | p.Arg211Cys<br>p.Asp185Gly | NR | - | NR | - | NR | - |
| ***NF1*** | 3 (5%) | 3 (5%) | p.Arg659Trp<br>p.Ala422Thr<br>p.Pro2613Leu | NR | - | NR | - | NA | - |
| ***PPM1D*** | 1 (2%) | 1 (2%) | p.Arg458Gln | NR | - | NR | - | NA | - |
| ***PTEN*** | 1 (2%) | 1 (2%) | p.Tyr178Cys | NR | - | NR | - | NA | - |
| ***RB1*** | 1 (2%) | 1 (2%) | p.Arg787Gln | NR | - | NR | - | NA | - |
| ***RET*** | 3 (5%) | 3 (5%) | p.Arg330Gln<br>p.Arg982Cys<br>p.Glu595Lys | NR | - | NR | - | NR | - |
| ***SPOP*** | 2 (3%) | 2 (3%) | p.Arg198Gln<br>p.Arg138Cys | NR | - | NR | - | NA | - |
| ***TG*** | 2 (3%) | 2 (3%) | p.Asp2001Asn<br>p.Asp2001Asn | NR | - | NR | - | NA | - |
| ***TP53*** | 3 (5%) | 2 (3%) | p.Gln317Lys<br>p.Cys124Tyr | NR | - | 1(25%) | Gain | NA | - |
| ***TSHR*** | 1 (2%) | 1 (2%) | p.Ala459Val | NR | - | NR | - | NA | - |
| ***ZFHX3*** | 8 (12%) | 6(10%) | p.Ala2741Pro<br>p.Ala2741Pro<br>p.Gly585Ser<br>p.Asp265Gly<br>p.Ala2433Val<br>p.Asp265Gly | 1(100%) | p.Gln3203_Gln3204del | 1(25%) | Gain | NA | - |
| **Total (%)** | **65 (100%)** | **59 (90%)** | - | **1 (2%)** | | **4 (6%)** | - | **1 (2%)** | - |

**Table Legend**: SNV single nucleotide variation; AA amino acid; NCR non-coding region; INDEL insertion/deletion; CNV copy number alteration; NR not retrieved; NA not assessed

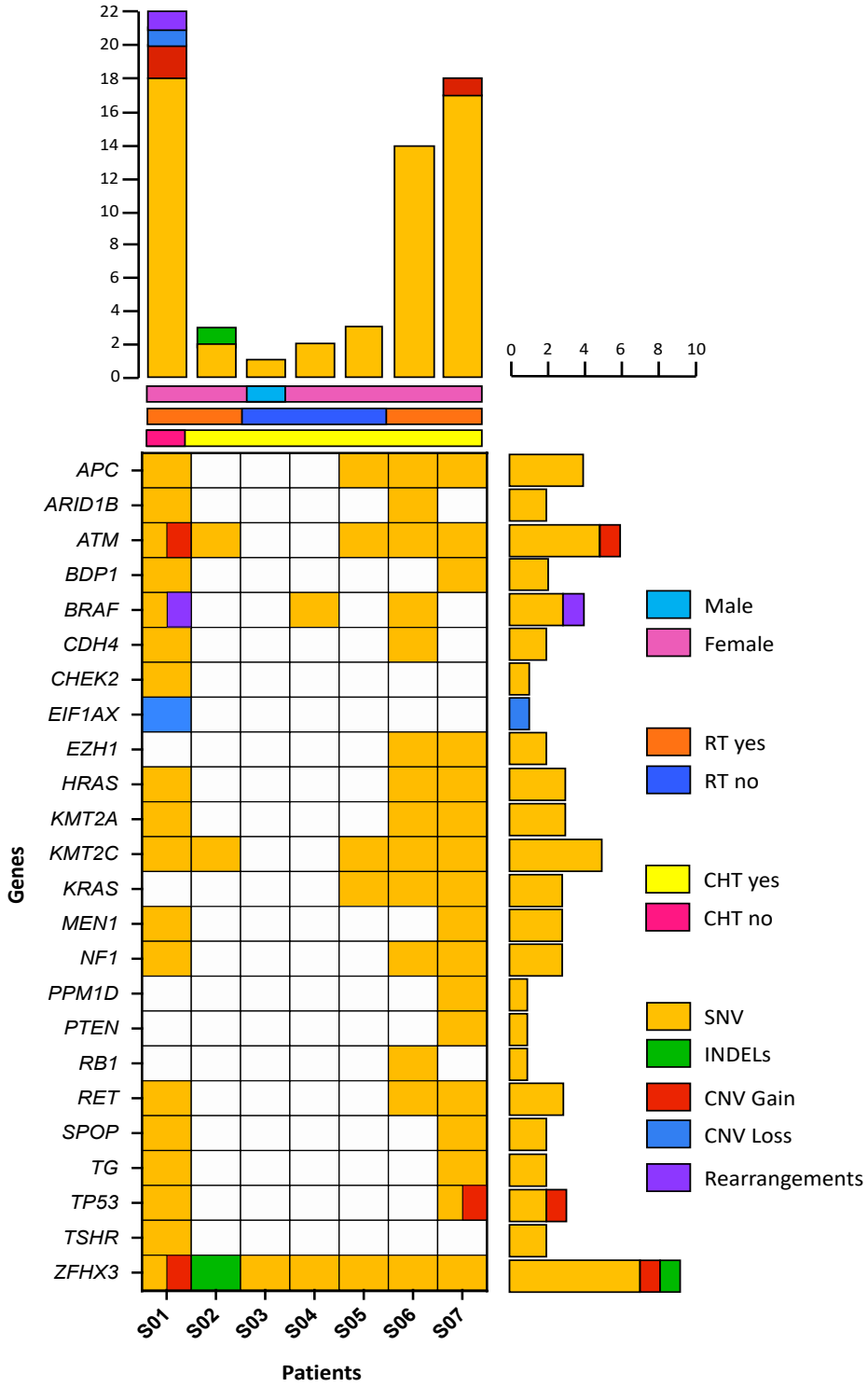



**Figure 5.2. Tile Plot of Molecular Alterations in Secondary Papillary Thyroid Cancers**

The tile plot shows the mutational landscape of the examined secondary papillary thyroid cancer samples. The color of each tile represents a different alteration type. Patients' gender and previous exposure to radiotherapy and/or chemotherapy are displayed in the horizontal bars. The numbered lines indicate the total alterations retrieved in each sample (vertical) and for each gene (horizontal). **Figure Legend:** RT radiotherapy; CHT chemotherapy; SNV single nucleotide variation; INDEL insertion/deletion; CNV copy number variation.

### 5.2.3. Comparison Between De Novo and Secondary Papillary Thyroid Cancers

**Figure 5.3** shows the retrieved molecular alterations in both de novo and secondary PTC. At a glance, it can be noticed that secondary PTC harbor a significantly higher burden of alterations (average 9.3), compared to de novo PTC (average 3.4). This difference is mainly driven by the presence of hypermutated patients (S01, S06 and S07) in the secondary PTC cohort. Single nucleotide variations were the prevalent alteration type in both groups, but secondary PTC displayed a higher CNV rate, including 3 gains and 1 loss, compared to one gain only among de novo PTC.

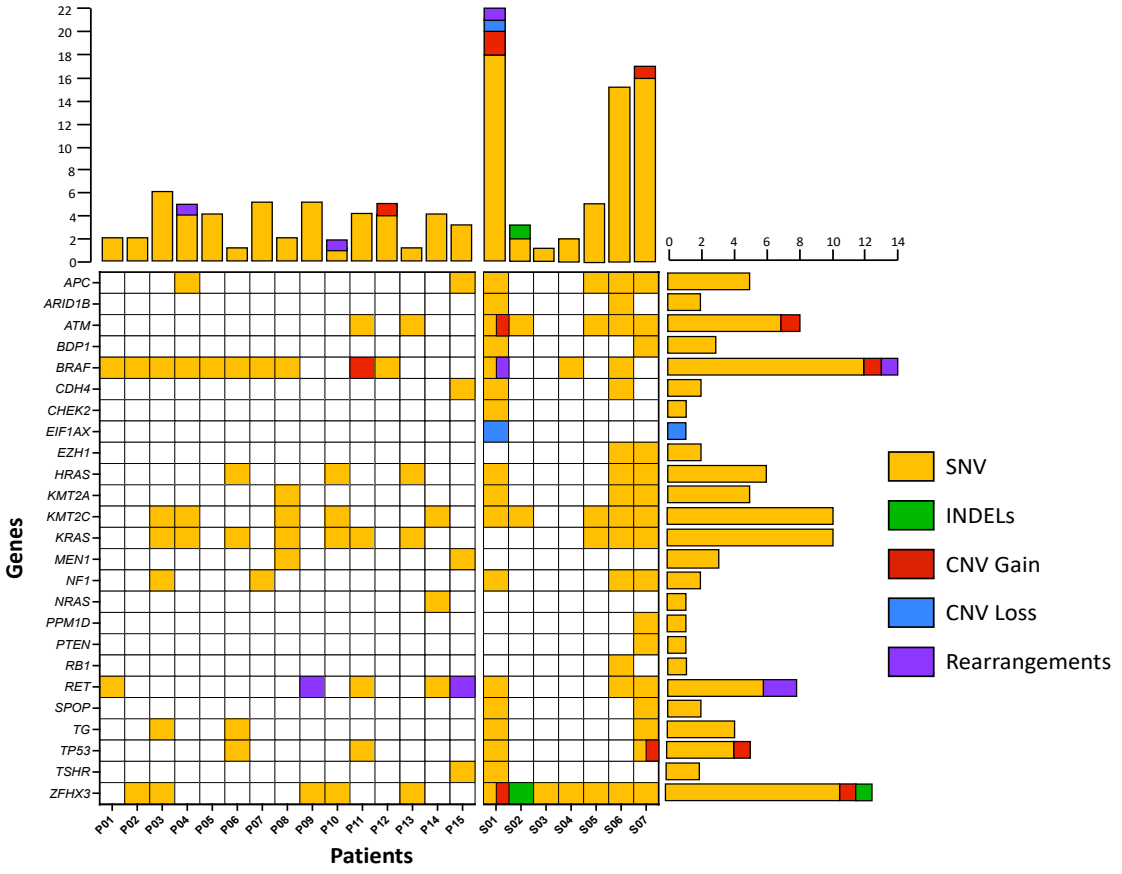



**Figure 5.3. Tile Plot of Molecular Alterations in De Novo and Secondary Papillary Thyroid Cancers**

The tile plot shows the mutational landscape of the examined papillary thyroid cancer samples, both the novo and secondary. The color of each tile represents a different alteration type. The numbered lines indicate the total alterations retrieved in each sample (vertical) and for each gene (horizontal). **Figure Legend:** SNV single nucleotide variation; INDEL insertion/deletion; CNV copy number variation.

### 5.2.4. Thyroid Adenomas

A total of 44 alterations in 14 genes were observed in the adenoma samples (**Table 5.6, Figure 5.4**). Of these, 84% were SNV (n=37), 11% were CNV (n=5) and 5% were INDEL (n=2). The mean number of alteration per subject was 2.8.

**Table 5.6. Molecular Alterations of Thyroid Adenomas**

| Gene Name | Alteration n° (%) | SNV n° (%) | AA substitution | INDEL n° (%) | INDEL | CNV n° (%) | CNV |
|---|---|---|---|---|---|---|---|
| ***APC*** | 4 (9%) | 3 (8%) | p.Ser2621Cys<br>p.Gly2502Ser | 1 (50%) | p.Ser1757PhefsTer12 | NR | - |
| ***ATM*** | 3 (7%) | 3 (8%) | p.Asp1853Asn | NR | - | NR | - |
| ***BDP1*** | 2 (5%) | 2 (5%) | p.Gln1676Glu | NR | - | NR | - |
| ***BRAF*** | 2 (5%) | 1 (3%) | p.Arg682Gln | NR | - | 1 (20%) | Gain |
| ***EIF1AX*** | 3 (7%) | NR | - | NR | - | 3 (60%) | Loss |
| ***HRAS*** | 1 (2%) | 1 (3%) | p.Gly13Arg | NR | - | NR | - |
| ***KMT2A*** | 1 (2%) | NR | - | 1 (50%) | p.Ser10LysfsTer32 | NR | - |
| ***KMT2C*** | 7 (16%) | 6 (16%) | p.Val4539Met<br>p.Gln3478Glu<br>p.Phe3226Leu<br>p.Asp348Asn<br>p.Leu291Phe<br>p.Arg284Gln | NR | - | 1 (20%) | Gain |
| ***KRAS*** | 8 (18%) | 8 (21%) | p.Gly12Asp<br>NCR | NR | - | NR | - |
| ***PTEN*** | 1 (2%) | 1 (3%) | p.Ile101Thr | NR | - | NR | - |
| ***RET*** | 1 (2%) | 1 (3%) | p.Arg982Cys | NR | - | NR | - |
| ***SPOP*** | 3 (7%) | 3 (8%) | p.Pro94Arg | NR | - | NR | - |
| ***TSHR*** | 1 (2%) | 1 (3%) | p.Arg531Trp | NR | - | NR | - |
| ***ZFHX3*** | 7 (16%) | 7 (19%) | p.Asp265Gly<br>p.Ala62Val | NR | - | NR | - |
| **Total (%)** | **44 (100%)** | **37 (84%)** | **-** | **2 (5%)** | | **5 (11%)** | **-** |

**Table Legend**: SNV single nucleotide variation; AA amino acid; NCR non-coding region; INDEL insertion/deletion; CNV copy number alteration; NR not retrieved

Single nucleotide variations occurred most commonly in *KRAS* sequence (n=8), more frequently in the non-coding region, followed *ZFHX3* (n=7) and *KMT2C* (n=6). *APC*, *ATM* and *SPOP* were mutated in 3 instances and BDP1 in 2. *BRAF*, *HRAS*, *PTEN*, *RET* and *TSHR* presented one SNV each in this population.

The retrieved INDEL were an insertion in *APC* and a deletion in *KTM2A*.

Of the 5 CNV observed, 2 were copy number gains, both occurring in patient A02 and involving BRAF and *KMT2C* on chromosome 7 (**Supplementary Figure 4**). The remaining 3 CNV were copy number losses of *EIF1AX* gene, on chromosome X (**Supplementary Figure 5-7**).

No significant differences emerged between female and male subjects, even though the presence of only two men in this cohort does not allow to drive any conclusion.

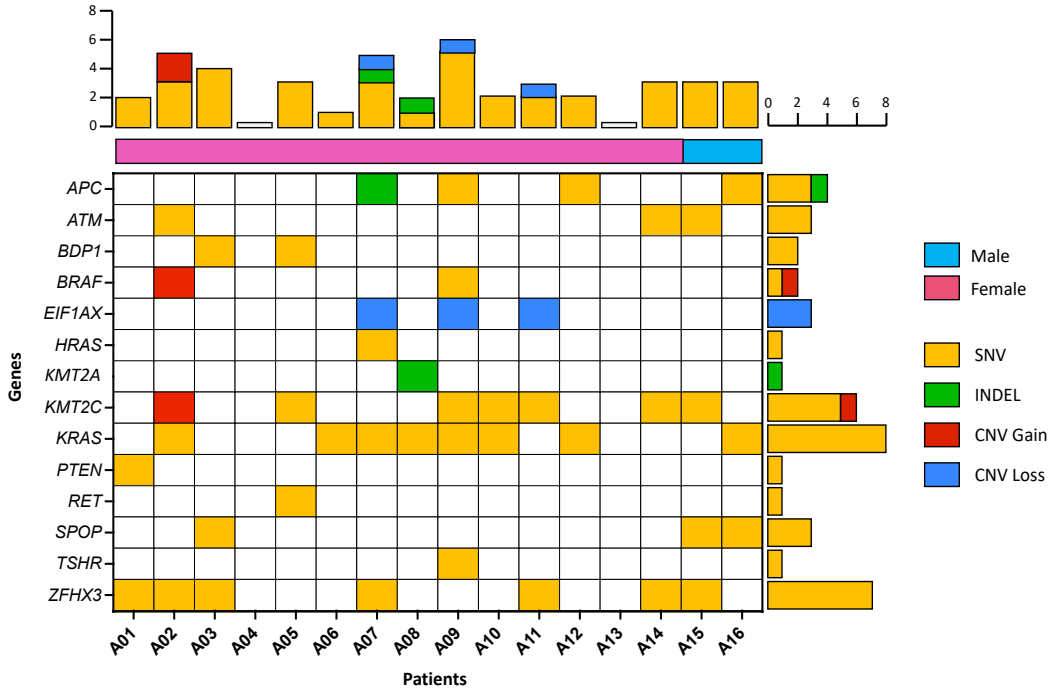



**Figure 5.4. Tile Plot of Molecular Alterations in Thyroid Adenomas**

The tile plot shows the mutational landscape of the examined adenoma samples. The color of each tile represents a different alteration type. Patients' gender is displayed in the horizontal bar. The numbered lines indicate the total alterations retrieved in each sample (vertical) and for each gene (horizontal). **Figure Legend:** SNV single nucleotide variation; INDEL insertion/deletion; CNV copy number variation.

### 5.2.5. Comparison of Molecular Alterations in the Populations

To highlight the differentially altered genes in the examined populations we plotted them in a Venn diagram (**Figure 5.5**). It can be noticed that de novo PTC, secondary PTC and adenomas shared most of the altered genes (10/25). Five genes were mutated in both de novo and secondary PTC, namely *CDH4*, *MEN1*, *NF1*, *TG* and *TP53* and 4 in both secondary PTC and adenomas (*BDP1*, *EIF1AX*, *PTEN*, *SPOP*). *NRAS* was mutated in a de novo PTC only, whereas *ARID1B*, *CHEK2*, *EZH1*, *PPM1D* and *RB1* were altered in secondary PTC but not in the other study groups. Of note, all the alterations retrieved in the adenoma samples were present in at least another study group.

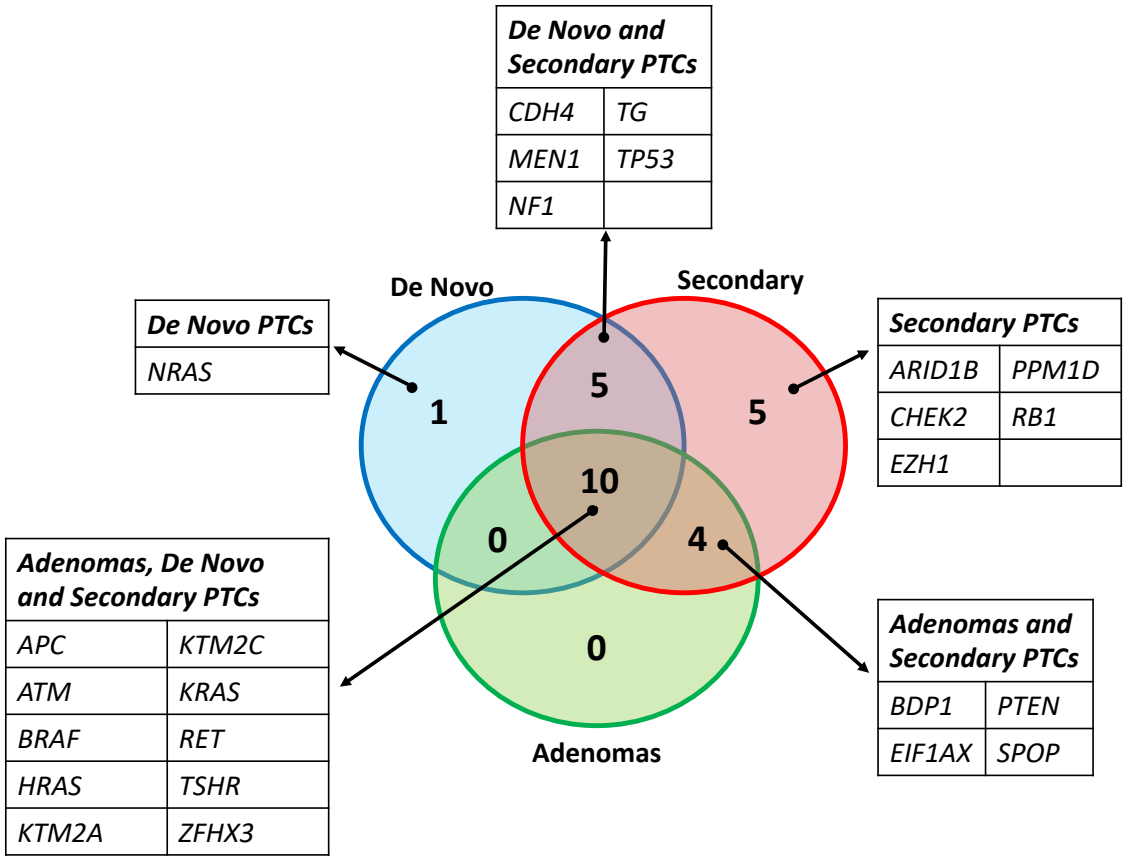



**Figure 5.5. Venn Diagram of Altered Genes in the Examined Populations**
The diagram illustrates the differentially altered gene in each population.

## 5.3. STRING Analysis

In order to explore how the molecular alterations found in our populations integrate and interact each other at a protein level, we performed STRING analysis with clusters generation. The minimum required interaction score to include predicted networks was set at 0.4 (medium confidence). Clustering was generated using Markov Cluster (MCL) algorithm (123). Protein-protein interaction networks and clustering were studied separately for de novo PTC, secondary PTC and adenomas.

### 5.3.1. De Novo Papillary Thyroid Cancers

According to STRING analysis, 16 proteins derived from altered genes in de novo PTC, each representing a network node, established 53 connections (**Figure 5.6**). The average node degree (i.e. the mean number of interactions of each protein in the network, at the score threshold) was 6.62. The average clustering coefficient (i.e. a measure of how connected the nodes in the network are, with higher values for higher connected networks) was 0.76. The number of connections was higher than the predicted one (=12), suggesting that most of the proteins are likely biologically related ($p < 1.0e-16$). In fact, 14 out of 16 proteins grouped in one main cluster (red nodes). The core of this cluster included 12 nodes, linked by multiple interaction types. CHD4 and ZFHX3, which stood at the periphery of the cluster, were connected by interactions extrapolated from databases and textmining only, respectively. The second minor cluster (green nodes) included TG and TSHR only, linked by a single textmining-derived interaction.

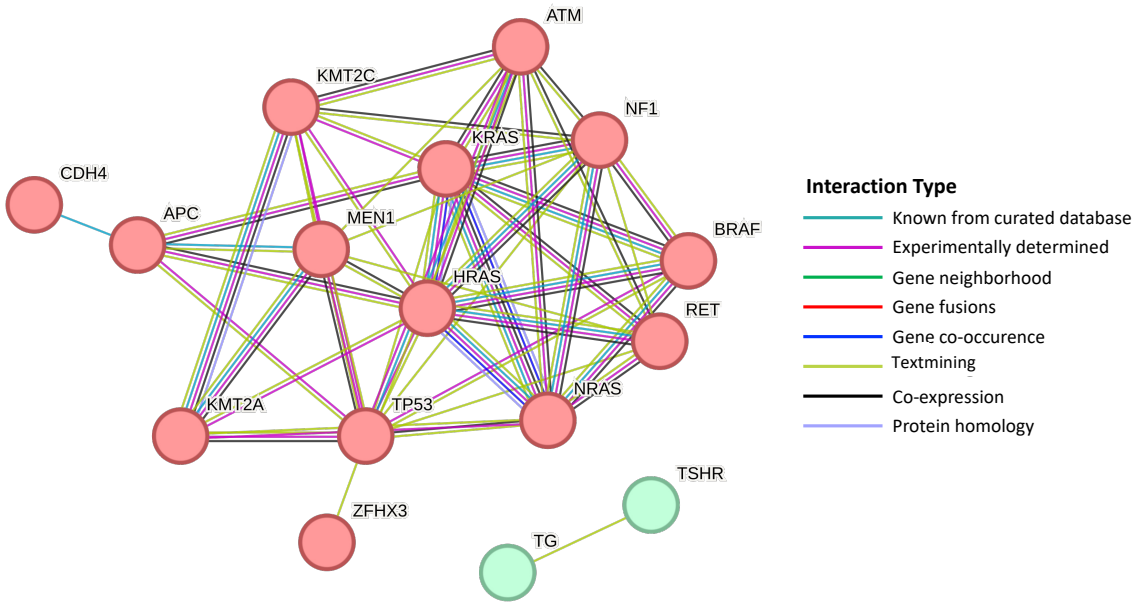



**Figure 5.6. STRING Analysis of De Novo Papillary Thyroid Cancers**
The figure shows the protein-protein interaction (PPI) network and clustering of gene alterations retrieved in the de novo papillary thyroid cancer samples. Network nodes represent the proteins produced by the genes. The color of the nodes indicates their cluster. Different colors of connecting lines indicates different types of interactions, as shown in figure legend.

### 5.3.2. Secondary Papillary Thyroid Cancers

STRING analysis of protein encoded by 24 altered genes in the secondary PTC population showed 93 connections in the PPI network (**Figure 5.7**). The average node degree was 7.8 and the clustering coefficient was 0.7. As for de novo PTC, the number of connections was by far higher than the expected (n=23), with a PPI enrichment p value< 1.0e-16. Even in this case, two clusters emerged. The main one included 19 of 24 proteins, establishing a plethora of different interaction types. Again, CDH4 and ZFHX3 were connected by a single textmining-derived interaction to the core of the cluster. The second network included TG and TSHR only (green nodes). Of note, EZH1 was the only non-clustered protein.

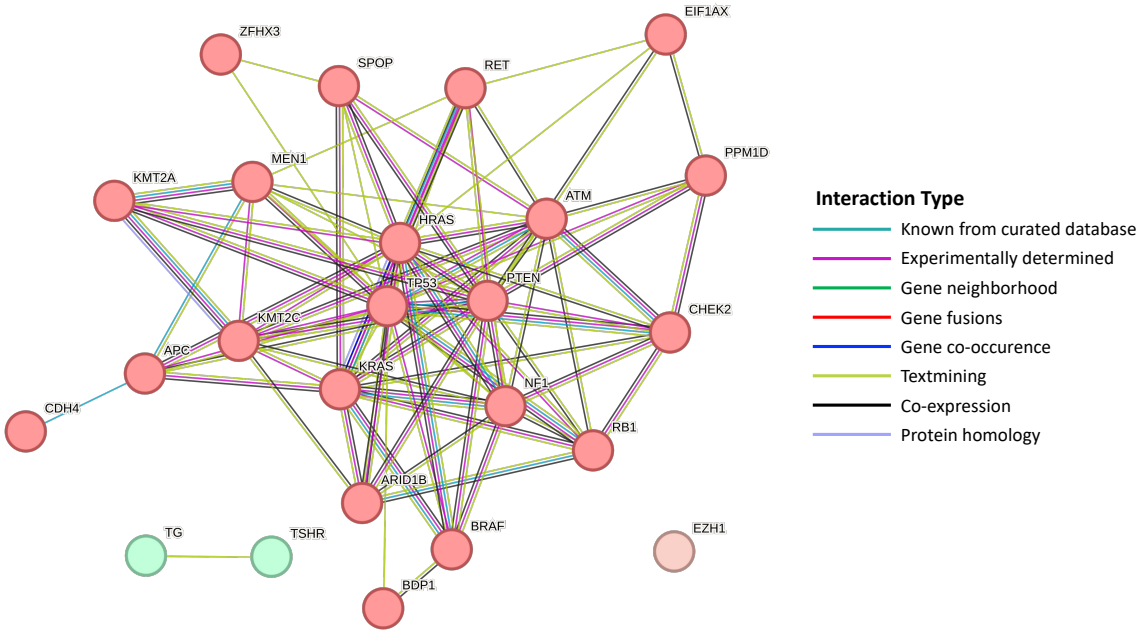



**Figure 5.7. STRING Analysis of Secondary Papillary Thyroid Cancers**
See Figure 5.6 caption.

### 5.3.3. Thyroid Adenomas

The 14 nodes representing the altered proteins in the adenoma samples established 34 connections in the STRING analysis output (**Figure 5.8**). The average node degree was 4.86, with a clustering coefficient of 0.75. The number of connections exceeded the one predicted by the algorithm (n=8), indicating again a strong biological relationship for the given entries. The PPI enrichment was highly statistically significant (p value 1.71e-12). Of note, all the 14 proteins clustered together (red nodes). The cluster included multiple interaction types for the majority of proteins, excluding ZFH3 and TSHR, that were linked only by textmining-derived connections.

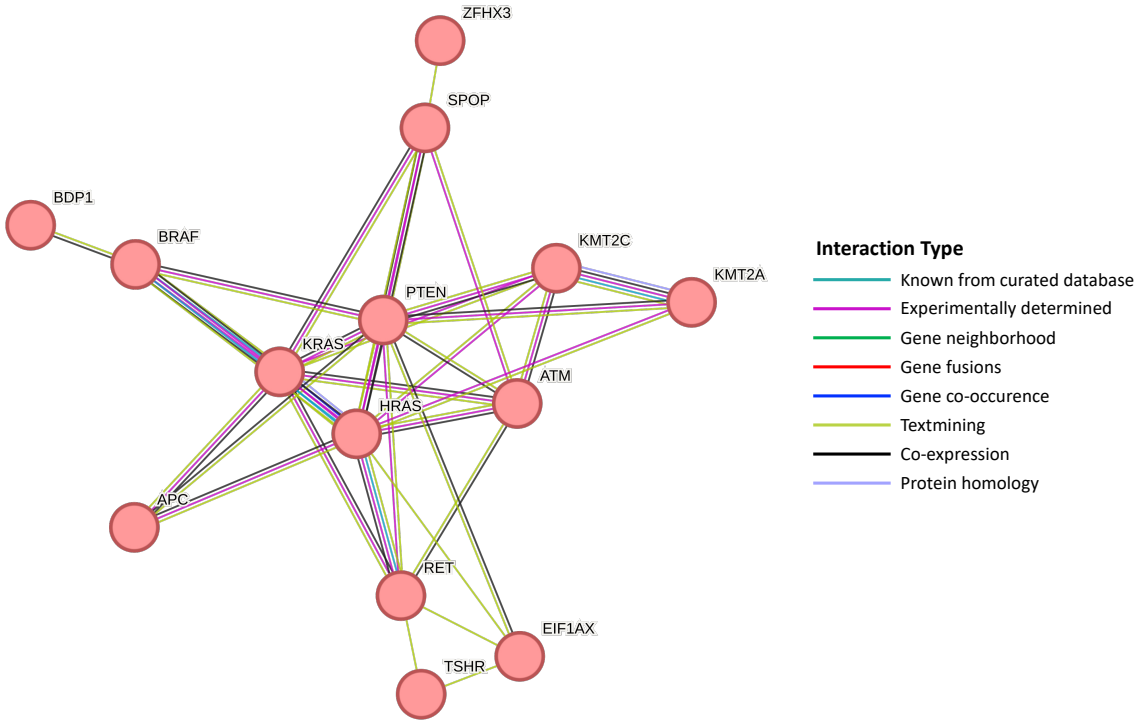



**Figure 5.8. STRING Analysis of Thyroid Adenomas**
See Figure 5.6 caption.

## 5.4. Biological Processes Analysis

To understand how altered genes in our populations influences biological processes, we performed a bio-informatic analysis according to Gene Ontology 2021, using the EnrichR online tool. Bi-dimensional visualization of altered biological processes was obtained through the Appyter online tool. We report here the top 10 statistically significant Gene Ontology enriched terms. The analysis was separately carried out for de novo PTC, secondary PTC and adenomas. Information about the top 100 significantly enriched terms for each population are provided in **Supplementary Tables.2-4**.

### 5.4.1. De Novo Papillary Thyroid Cancers

**Figure 5.9** shows the top ten enriched biological processes in the de novo PTC population, according to Gene Ontology 2021 library. The correlated Gene Ontology identification numbers, p values, q-values and overlapping genes are displayed in **Table 5.7**. The q-value is an adjusted p-value calculated using the Benjamini-Hochberg method, used to control the false discovery rate for multiple hypotheses testing (124).

Among the enriched terms, “Small GTPase Mediated Signal Transduction”, “RAS Protein Signal Transduction”, “Regulation Of Cell Cycle And Histone H3-K4 Methylation” clustered together, with p values ranging from 1.59e-08 (“Small GTPase Mediated Signal Transduction”) to 2.62e-06 (“Regulation Of Cell Cycle”).

Three positively regulated processes converged in a single cluster, with p values spanning from 1.23e-06 (“Positive Regulation Of Transcription, DNA-Templated”) to 5.57e-06 (“Positive Regulation Of Cellular Process”).

“Response To Gamma Radiation” (p value 2.87e-09) and “Cellular Response To Gamma Radiation” (p value 2.84e-07) were grouped in another cluster.

Lastly, “MAPK Cascade” did not cluster with any other term and represented the biological process with the highest statistical significance in this population (1.74e-09).

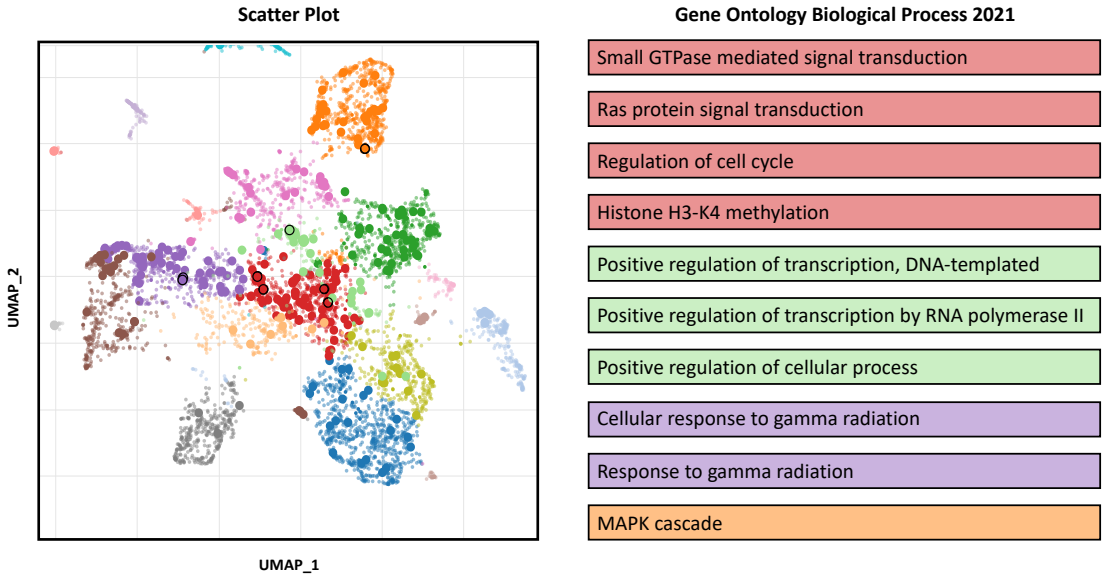



**Figure 5.9. Scatter Plot Visualization of Enriched Biological Process in De Novo Papillary Thyroid Cancers, According to Gene Ontology Analysis**

On the left: each point represents a term in the Gene Ontology library. Term frequency-inverse document frequency (TF-IDF) values were computed for the gene set corresponding to each term, and UMAP was applied to the resulting values. The terms are plotted based on the first two UMAP dimensions. Terms are colored by automatically identified clusters computed with the Leiden algorithm applied to the TF-IDF values. Top ten significantly enriched terms are represented as larger, black-outlined, points. On the right: the top ten significant Gene Ontology terms are reported in the boxes, with colors indicating the corresponding cluster. **Figure Legend:** UMAP Uniform Manifold Approximation and Projection.

**Table 5.7. Enriched Biological Processes in De Novo Papillary Thyroid Cancers, According to Gene Ontology Analysis**

| Biological Process | ID | p-value | q-value | Overlapping Genes |
|---|---|---|---|---|
| **MAPK Cascade** | GO:0000165 | 1.7395487929341005e-09 | 6.853287586558382e-07 | *RET, NRAS, NF1, BRAF, KRAS, HRAS, MEN1* |
| **Response To Gamma Radiation** | GO:0010332 | 2.873495843420705e-09 | 6.853287586558382e-07 | *ATM, HRAS, TP53, MEN1* |
| **Small GTPase Mediated Signal Transduction** | GO:0007264 | 1.5852753307949112e-08 | 2.520587775963909e-06 | *NRAS, NF1, KRAS, HRAS, TP53* |
| **Ras Protein Signal Transduction** | GO:0007265 | 6.894202313846511e-08 | 8.221336259261964e-06 | *NRAS, NF1, KRAS, HRAS, TP53* |
| **Cellular Response To Gamma Radiation** | GO:0071480 | 2.8366829343744953e-07 | 2.7061955193932686e-05 | *ATM, HRAS, TP53* |
| **Positive Regulation Of Transcription, DNA-Templated** | GO:0045893 | 1.2288524430338584e-06 | 9.769376922119175e-05 | *RET, ZFHX3, KMT2A, KMT2C, ATM, HRAS, TP53, MEN1* |
| **Histone H3-K4 Methylation** | GO:0051568 | 1.683001407154634e-06 | 0.00011468452445896577 | *KMT2A, KMT2C, MEN1* |
| **Regulation Of Cell Cycle** | GO:0051726 | 2.6231602249832543e-06 | 0.00015640592841462654 | *ZFHX3, KMT2A, KMT2C, ATM, HRAS, TP53, MEN1* |
| **Positive Regulation Of Transcription By RNA Polymerase II** | GO:0045944 | 3.0996875449344348e-06 | 0.00016428343988152504 | *ATM, TP53, MEN1* |
| **Positive Regulation Of Cellular Process** | GO:0048522 | 5.571151643809081e-06 | 0.00026574393340969316 | *ATM, HRAS, TP53* |

### 5.4.2. Secondary Papillary Thyroid Cancers

The top ten statistically significant enriched Gene Ontology terms in secondary PTC are shown in **Figure 5.10** and the detailed list is reported in **Table 5.8**.

The terms “Small GTPase Mediated Signal Transduction”, “Ras Protein Signal Transduction”, “Negative Regulation Of Serine/Threonine Kinase Activity” and “Regulation Of Cell Cycle” clustered together. The biological process with the highest p-value in this cluster was “Negative Regulation Of Protein Serine/Threonine Kinase Activity” (3.18e-08), while “Regulation Of Cell Cycle” was the one with the lowest statistical significance (p value 1.07e-06).

“Negative Regulation Of Cell Population Proliferation” (p value 2.18e-07) and “Positive Regulation Of Transcription, DNA Templated” (p value 4.61e-07) were grouped in another cluster.

“Response To Gamma Radiation” was the most significantly enriched process in this population (p value 1.67e-08) and, as for de novo PTC, it clustered with “Cellular Response To Gamma Radiation” (p value 1.02e-06).

Even in secondary PTC the “MAPK Cascade” term did not cluster but, differently from de novo PTC, it was the least statistically significant enriched process of the list (p value 1.23e-06).

“Replicative Senescence” term also did not cluster and exhibited a p value of 3.32e-07.

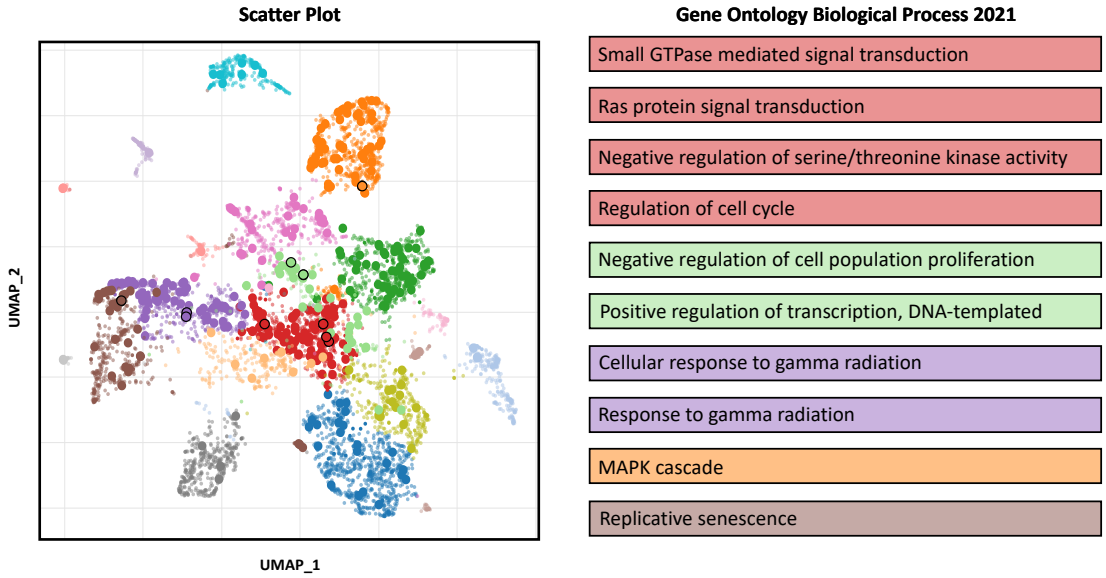



**Figure 5.10. Scatter Plot Visualization of Enriched Biological Process in Secondary Papillary Thyroid Cancers, According to Gene Ontology Analysis**

See Figure 5.9 caption.

**Table 5.8. Enriched Biological Process in Secondary Papillary Thyroid Cancers, According to Gene Ontology Analysis**

| Biological Process | ID | p-value | q-value | Overlapping Genes |
|---|---|---|---|---|
| **Response To Gamma Radiation** | GO:0010332 | 1.667021133489588e-08 | 1.0277037422197609e-05 | *ATM, HRAS, TP53, MEN1* |
| **Negative Regulation Of Protein Serine/Threonine Kinase Activity** | GO:0071901 | 3.17682764210127e-08 | 1.0277037422197609e-05 | *RB1, APC, PTEN, NF1, MEN1* |
| **Small GTPase Mediated Signal Transduction** | GO:0007264 | 1.4915175157189225e-07 | 3.216706108900476e-05 | *RB1, NF1, KRAS, HRAS, TP53* |
| **Negative Regulation Of Cell Population Proliferation** | GO:0008285 | 2.1762177729142788e-07 | 3.520032247688846e-05 | *APC, PTEN, NF1, PPM1D, HRAS, TP53, MEN1* |
| **Replicative Senescence** | GO:0090399 | 3.3161472134701576e-07 | 4.291094494230384e-05 | *CHEK2, ATM, TP53* |
| **Positive Regulation Of Transcription, DNA-Templated** | GO:0045893 | 4.614368874307938e-07 | 4.9758277694620595e-05 | *RET, ZFHX3, KMT2A, KMT2C, CHEK2, ATM, ARID1B, HRAS, TP53, MEN1* |
| **Ras Protein Signal Transduction** | GO:0007265 | 6.408939107990364e-07 | 5.923690861242522e-05 | *RB1, NF1, KRAS, HRAS, TP53* |
| **Cellular Response To Gamma Radiation** | GO:0071480 | 1.0209769829080983e-06 | 7.7231362910504e-05 | *ATM, HRAS, TP53* |
| **Regulation Of Cell Cycle** | GO:0051726 | 1.0743157128199938e-06 | 7.7231362910504e-05 | *RB1, APC, ATM, HRAS, TP53, MEN1* |
| **MAPK Cascade** | GO:0000165 | 1.2307995938249777e-06 | 7.963273372047606e-05 | *RB1, APC, ATM, HRAS, TP53, MEN1* |

### 5.4.3. Thyroid Adenomas

In the adenoma population, Gene Ontology analysis showed that the terms “Positive Regulation Of Cellular Processes”, “Regulation Of Cell Migration” and “Positive Regulation Of Cell Transcription, DNA Templated” clustered together (**Figure 5.11**), with p values ranging from 7.98e-08 (“Positive Regulation Of Cellular Processes”, the term with the highest significance in this population) to 8.41e-05 (“Positive Regulation Of Cell Transcription, DNA Templated”) (**Table 5.9**).

A second cluster was represented by the terms “Positive Regulation Of Cell Motility” (p value 1.33e-05) and “Positive Regulation Of Cell Migration” (p value 2.17e-05), while “Histone H3-K4 Monomethylation” (p value 1.27e-05) and “Positive Regulation Of Cell Adhesion” (p value 2.17e-05) converged in a third group.

The biological processes “Cellular Response To Gamma Radiation” and “Negative Regulation Of Cyclin-Dependent Protein Serine/Threonine Kinase Activity” clustered together with a p-value of 6.15e-05 and 8.58e-05, respectively.

Lastly, as for the previous populations, the “MAPK Cascade” process (4.59e-05) was not part cluster of cluster of enriched terms.

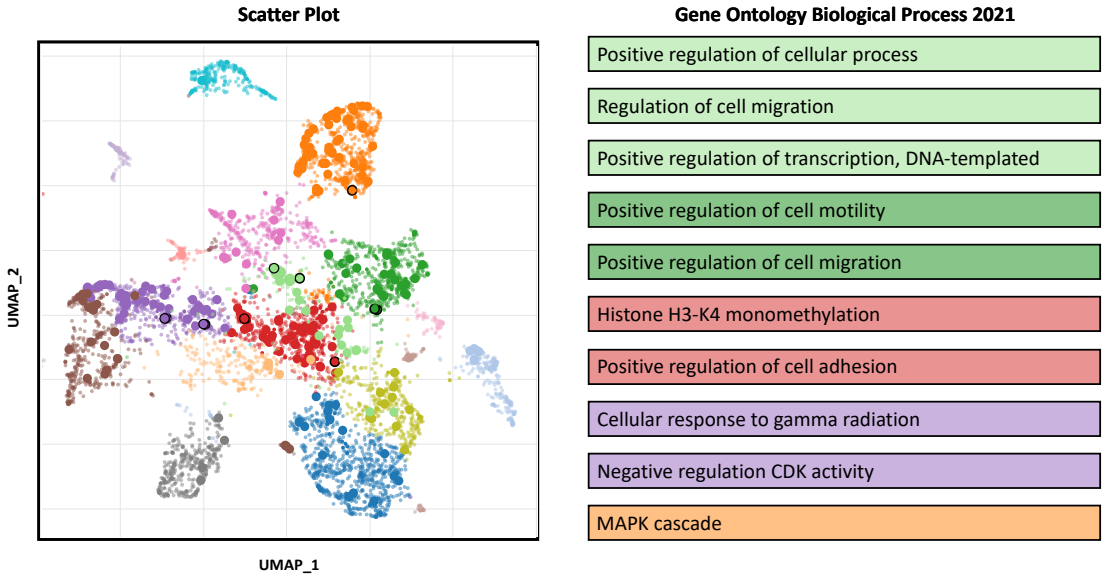



**Figure 5.11. Scatter Plot Visualization of altered Biological Process in Thyroid Adenomas According to Gene Ontology Analysis**

See Figure 5.9 caption.

## Table 5.9. Altered Biological Process in Thyroid Adenomas According to Gene Ontology Analysis

| Biological Process | ID | p-value | q-value | Overlapping Genes |
|---|---|---|---|---|
| **Positive Regulation Of Cellular Process** | GO:0048522 | 7.979924444029842e-08 | 3.0084315153992504e-05 | *ZFHX3, APC, PTEN, ATM, KRAS, HRAS, TSHR* |
| **Regulation Of Cell Migration** | GO:0030334 | 5.929673566571532e-06 | 0.0011177434672987338 | *RET, APC, PTEN, ATM, HRAS* |
| **Histone H3-K4 Monomethylation** | GO:0097692 | 1.270931618352025e-05 | 0.0013281066431612809 | *KMT2A, KMT2C* |
| **Positive Regulation Of Cell Motility** | GO:2000147 | 1.3314352312393792e-05 | 0.0013281066431612809 | *RET, APC, ATM, HRAS* |
| **Positive Regulation Of Cell Adhesion** | GO:0045785 | 2.1729931620316324e-05 | 0.0017340485433012425 | *RET, ZFHX3, ATM* |
| **Positive Regulation Of Cell Migration** | GO:0030335 | 2.8806883666937314e-05 | 0.0019156577638513314 | *RET, APC, ATM, HRAS* |
| **MAPK Cascade** | GO:0000165 | 4.585699206328194e-05 | 0.0026001924264631646 | *RET, BRAF, KRAS, HRAS* |
| **Cellular Response To Gamma Radiation** | GO:0071480 | 6.15100567126949e-05 | 0.0026001924264631646 | *ATM, HRAS* |
| **Positive Regulation Of Transcription, DNA-Templated** | GO:0045893 | 8.411821210449226e-05 | 0.0026001924264631646 | *RET, ZFHX3, KMT2A, KMT2C, ATM, HRAS* |
| **Negative Regulation Of Cyclin-Dependent Protein Serine/Threonine Kinase Activity** | GO:0045736 | 8.583035665856277e-05 | 0.0026001924264631646 | *APC, PTEN* |

### 5.4.4. Comparison of Enriched Biological Processes in the Populations

To highlight similarities and difference between the enriched biological pathways in our populations plotted the top ten significant Gene Ontology term in a Venn diagram (**Figure 5.12**).

Three biological processes were commonly enriched in de novo PTC, secondary PTC and adenomas, namely “Positive Regulation of Transcription, DNA Templated”, “Cellular Response to Gamma Radiation” and “MAPK Cascade”. Each of these terms belong to a different cluster, as shown in the scatter plots reported above. Four biological processes were enriched in both de novo and secondary PTC. Of note, 3 of them (“Small GTPase Mediated Signal Transduction”, “Ras Protein Signal Transduction” and “Regulation of Cell Cycle”) clustered together. “Positive Regulation of Cellular Process” was a shared term between de novo PTC and adenomas, while this latter group did not show any common biological process with secondary PTC.

The remaining 11 terms were enriched in one population only, of which 2 in de novo PTC, 3 in secondary PTC and 6 in adenomas. Among the terms enriched in adenomas only, four were related with cell adhesion, migration or motility.

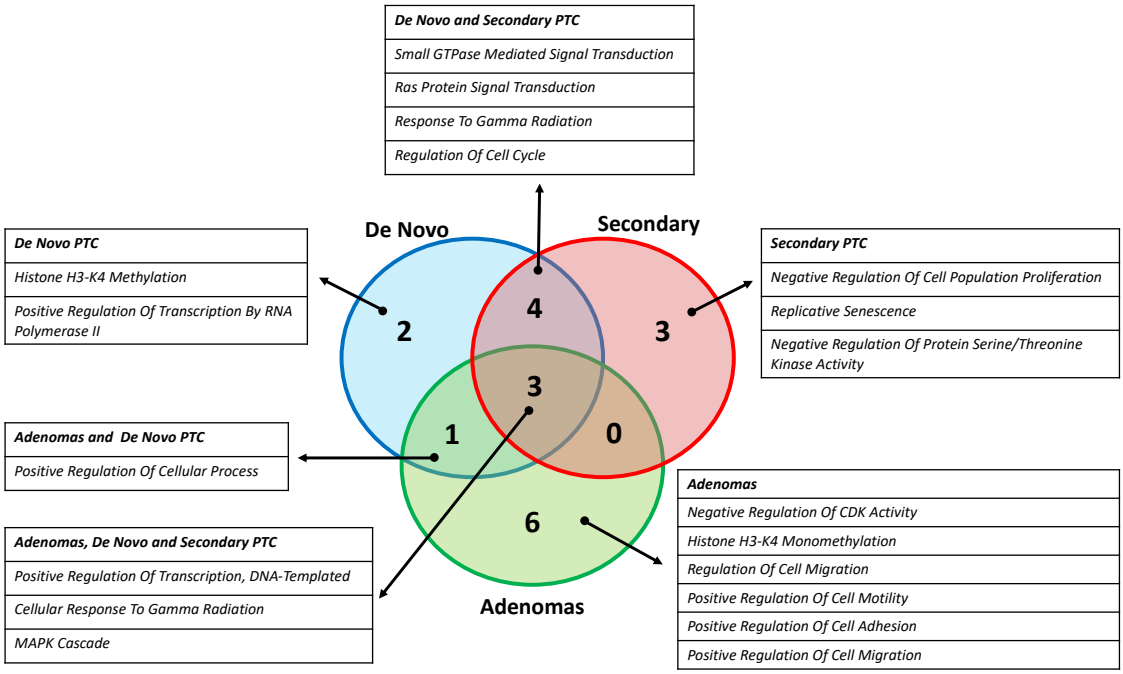



**Figure 5.12. Venn Diagram of Enriched Biological Processes in the Examined Populations**
The diagram illustrates the differentially enriched biological processes in each population.

## 5.5. Pathway Analysis

After assessing the enriched biological processes in our populations, we wanted to determine how altered genes influence cellular pathways. To this end, we carried out a bio-informatic analysis according to Reactome 2022, using the EnrichR online tool. Bi-dimensional visualization of pathways was obtained through the Appyter online tool. We report here the top 10 statistically significant Reactome enriched terms. The analysis was performed separately for de novo PTC, secondary PTC and adenomas. Information about the top 100 significantly enriched terms for each population are provided in **Supplementary Tables.5-7.**

### 5.5.1. De Novo Papillary Thyroid Cancers

**Figure 5.12** displays the top ten significantly enriched pathways in de novo PTC and their bidimensional spatial distribution, according to the Reactome 2022 library. Details about the pathway identification number, the p and q values and overlapping genes are provided in **Table 5.10.**

Of note, 9 out of 10 terms correlated to MAPK signaling and converged in the same cluster, with p values ranging from 3.76e-07 ("Oncogenic MAPK Signaling") to 3.72e-06 ("MAPK Family Signaling Cascades").

The only term not clustering with the others was "RAS Signaling Downstream Of NF1 Loss-Of-Function Variants" and was also the one with the highest statistical significance in this population (p value 8.39e-09).

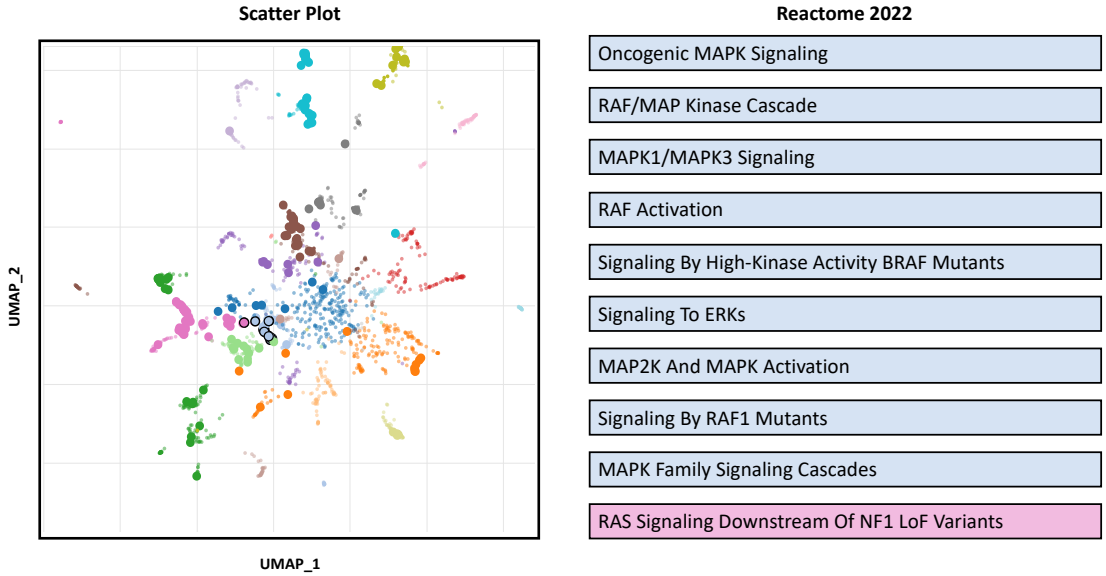



**Figure 5.13. Scatter Plot Visualization of Altered Pathways in De Novo Papillary Thyroid Cancers According to Reactome Analysis**

On the left: each point represents a term in the Reactome library. Term frequency-inverse document frequency (TF-IDF) values were computed for the gene set corresponding to each term, and UMAP was applied to the resulting values. The terms are plotted based on the first two UMAP dimensions. Terms are colored by automatically identified clusters computed with the Leiden algorithm applied to the TF-IDF values. Top ten significantly enriched terms are represented as larger, black-outlined, points. On the right: the top ten significant Gene Ontology terms are reported in the boxes, with colors indicating the corresponding cluster. **Figure Legend:** UMAP Uniform Manifold Approximation and Projection

**Table 5.10. Altered Pathways in De Novo Papillary Thyroid Cancers According to Reactome Analysis**

| Pathway | ID | p-value | q-value | Overlapping Genes |
|---|---|---|---|---|
| **RAS Signaling Downstream Of NF1 Loss-Of-Function Variants** | R-HSA-6802953 | 8.387593360272853e-09 | 2.5246656014421285e-06 | *NRAS, NF1, HRAS* |
| **Oncogenic MAPK Signaling** | R-HSA-6802957 | 3.758816366507574e-07 | 5.657018631593899e-05 | *NRAS, NF1, BRAF, HRAS* |
| **RAF/MAP Kinase Cascade** | R-HSA-5673001 | 1.7016183300408428e-06 | 9.449934319120904e-05 | *RET, NRAS, NF1, BRAF, HRAS* |
| **MAPK1/MAPK3 Signaling** | R-HSA-5684996 | 1.8947926377871997e-06 | 9.449934319120904e-05 | *RET, NRAS, NF1, BRAF, HRAS* |
| **RAF Activation** | R-HSA-5673000 | 2.0540807489581383e-06 | 9.449934319120904e-05 | *NRAS, BRAF, HRAS* |
| **Signaling By High-Kinase Activity BRAF Mutants** | R-HSA-6802948 | 2.258389565450525e-06 | 9.449934319120904e-05 | *NRAS, BRAF, HRAS* |
| **Signaling To ERKs** | R-HSA-187687 | 2.258389565450525e-06 | 9.449934319120904e-05 | *NRAS, BRAF, HRAS* |
| **MAP2K And MAPK Activation** | R-HSA-5674135 | 3.209962527762303e-06 | 9.449934319120904e-05 | *NRAS, BRAF, HRAS* |
| **Signaling By RAF1 Mutants** | R-HSA-9656223 | 3.4834058980573822e-06 | 9.449934319120904e-05 | *NRAS, BRAF, HRAS* |
| **MAPK Family Signaling Cascades** | R-HSA-5683057 | 3.7248458900597863e-06 | 9.449934319120904e-05 | *NRAS, BRAF, HRAS* |

### 5.5.2. Secondary Papillary Thyroid Cancers

In the secondary PTC population, pathway analysis showed more heterogeneous results compared to de novo tumors (**Figure 5.13**, **Table 5.11**).

The terms “Gene Expression (Transcription)”, “Generic Transcription Pathway”, “RNA Polymerase II Transcription”, “Regulation of TP53 Degradation” and “Regulation of TP53 Expression and Degradation” clustered together. The significance of these pathways enrichment went from 2.38e-08 [“Gene Expression (Transcription)”] to 1.15e-05 (“Regulation of TP53 Expression and Degradation”).

All the other terms did not cluster. Among them, “Recruitment And ATM-mediated Phosphorylation Of Repair And Signal Proteins At DNA Double Strand Breaks” showed the lowest p value of the list (4.97e-5)

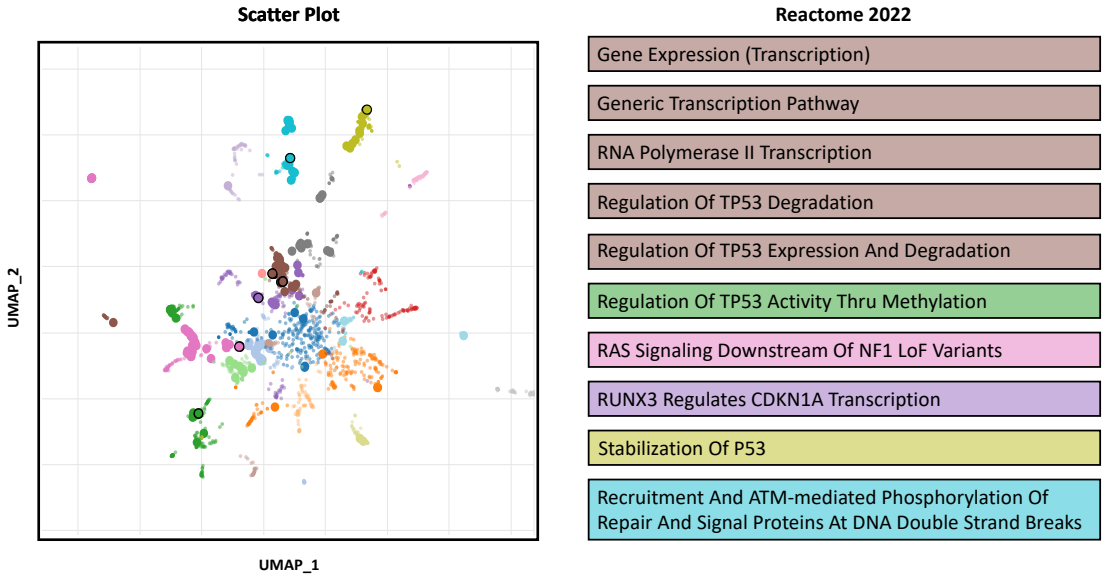



**Figure 5.14. Scatter Plot Visualization of Altered Pathways in Secondary Papillary Thyroid Cancers According to Reactome Analysis**

See Figure 5.12 caption.

**Table 5.11. Altered Pathways in Secondary Papillary Thyroid Cancers According to Reactome Analysis**

| Pathway | ID | p-value | q-value | Overlapping Genes |
|---|---|---|---|---|
| **Gene Expression (Transcription)** | R-HSA-74160 | 2.382084e-08 | 6.900994667936107e-06 | *RB1, ZFHX3, KMT2A, KMT2C, CHEK2, PTEN, ATM, PPM1D, BDP1, ARID1B, TP53, MEN1* |
| **Generic Transcription Pathway** | R-HSA-212436 | 3.844565e-08 | 6.900994667936107e-06 | *RB1, ZFHX3, KMT2A, KMT2C, CHEK2, PTEN, ATM, PPM1D, ARID1B, TP53, MEN1* |
| **RNA Polymerase II Transcription** | R-HSA-73857 | 1.046341e-07 | 1.2521212996541085e-05 | *RB1, ZFHX3, KMT2A, KMT2C, CHEK2, PTEN, ATM, PPM1D, ARID1B, TP53, MEN1* |
| **Regulation Of TP53 Activity Thru Methylation** | R-HSA-6804760 | 1.452608e-06 | 0.00013037159734931091 | *CHEK2, ATM, TP53* |
| **Regulation Of TP53 Degradation** | R-HSA-6804757 | 1.056131e-05 | 0.0006871347948762902 | *CHEK2, ATM, TP53* |
| **Regulation Of TP53 Expression And Degradation** | R-HSA-6806003 | 1.148415e-05 | 0.0006871347948762902 | *NRAS, BRAF, HRAS* |
| **RAS Signaling Downstream Of NF1 Loss-Of-Function Variants** | R-HSA-6802953 | 2.063923e-05 | 0.0010584976131054683 | *NF1, HRAS* |
| **RUNX3 Regulates CDKN1A Transcription** | R-HSA-8941855 | 2.887383e-05 | 0.0012957131774954857 | *ZFHX3, TP53* |
| **Stabilization Of P53** | R-HSA-69541 | 4.036253e-05 | 0.0015616062277531565 | *CHEK2, ATM, TP53* |
| **Recruitment And ATM-mediated Phosphorylation Of Repair And Signal Proteins At DNA Double Strand Breaks** | R-HSA-5693565 | 4.967047e-05 | 0.0015616062277531565 | *CHEK2, ATM, TP53* |

### 5.5.3. Thyroid Adenomas

Similarly to de novo PTC, most of the enriched pathways in the adenoma population (7/10) were closely related to the MAPK signaling and clustered together as shown in **Figure 5.11** and in **Table 5.9**. Among these terms, "RAF Activation" displayed the highest p-value (2.23e-04) and "Negative Regulation of MAPK Pathway" the lowest (p value 3.67e-04).

Other enriched terms were "Signaling By FGFR3" (p value 3.32e-04) and "Signaling by FGFR4" (p value 3.5e-04), that were grouped in a second cluster.

Lastly, "Gene Expression (Trasnscription)" did not cluster with other terms, with a p-value of 2.59e-04.

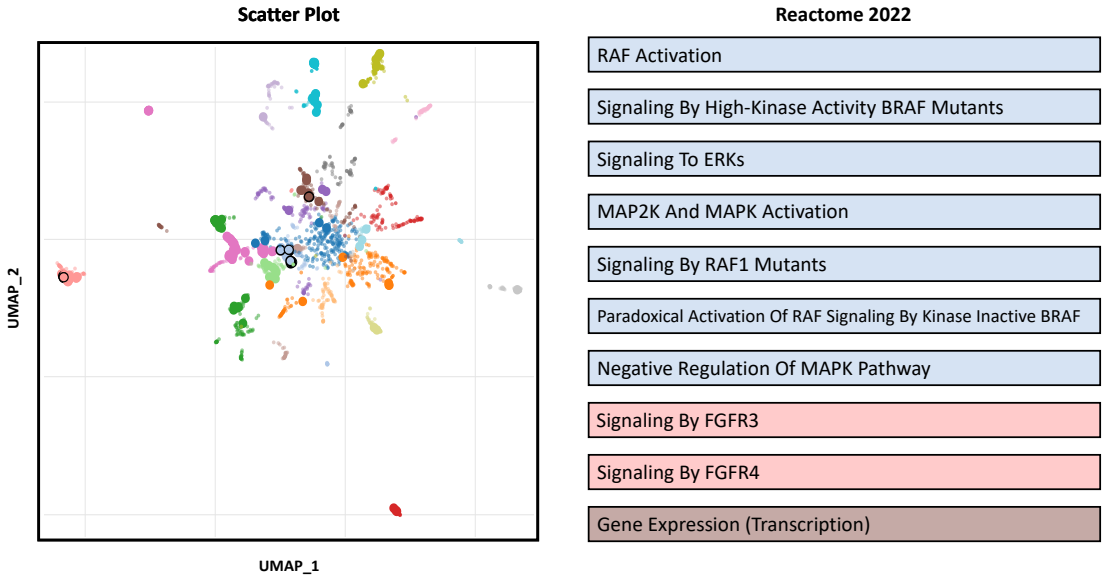



**Figure 5.15. Scatter Plot Visualization of Altered Pathways in Thyroid Adenomas According to Reactome Analysis**
See Figure 5.12 caption.

**Table 5.12. Altered Pathways in Thyroid Adenomas According to Reactome Analysis**

| Pathway | ID | p-value | q-value | Overlapping Genes |
|---|---|---|---|---|
| **RAF Activation** | R-HSA-5673000 | 0.00022299126039333032 | 0.00936695865197297 | *NRAS, BRAF, HRAS* |
| **Signaling By High-Kinase Activity BRAF Mutants** | R-HSA-6802948 | 0.00023728298476867357 | 0.00936695865197297 | *NRAS, BRAF, HRAS* |
| **Signaling To ERKs** | R-HSA-187687 | 0.00023728298476867357 | 0.00936695865197297 | *NRAS, BRAF, HRAS* |
| **Gene Expression (Transcription)** | R-HSA-74160 | 0.0002587167633104477 | 0.00936695865197297 | *ZFHX3, KMT2A, KMT2C, PTEN, ATM, BDP1* |
| **MAP2K And MAPK Activation** | R-HSA-5674135 | 0.0002988222171700243 | 0.00936695865197297 | *BRAF, HRAS* |
| **Signaling By RAF1 Mutants** | R-HSA-9656223 | 0.000315297453838473 | 0.00936695865197297 | *BRAF, HRAS* |
| **Signaling By FGFR3** | R-HSA-5654741 | 0.0003322078014627923 | 0.00936695865197297 | *BRAF, HRAS* |
| **Signaling By FGFR4** | R-HSA-5654743 | 0.00034955273047183893 | 0.00936695865197297 | *BRAF, HRAS* |
| **Paradoxical Activation Of RAF Signaling By Kinase Inactive BRAF** | R-HSA-6802955 | 0.0003673317118420773 | 0.00936695865197297 | *BRAF, HRAS* |
| **Negative Regulation Of MAPK Pathway** | R-HSA-5675221 | 0.0003673317118420773 | 0.00936695865197297 | *BRAF, HRAS* |

### 5.5.4. Comparison of Enriched Pathways in the Populations

As already done for Gene Ontology terms, we compared enriched pathways in the study populations by displaying them in a Venn diagram (**Figure 5.16**).

Of note, none of the enriched pathways were shared between de novo PTC, secondary PTC and adenomas. Only one term was common between de novo and secondary PTC ("RAS Signaling Downstream Of NF1 Loss-Of-Function Variants"). By contrast, de novo PTC and adenomas displayed the highest number of common terms (n=5), all related to MAPK signaling and clustering together. Lastly, secondary PTC and adenomas presented a shared term [Gene Expression (Transcription)].

Of the remaining pathways, 4 were enriched in de novo PTC only, 8 in secondary PTC and 4 in adenomas.

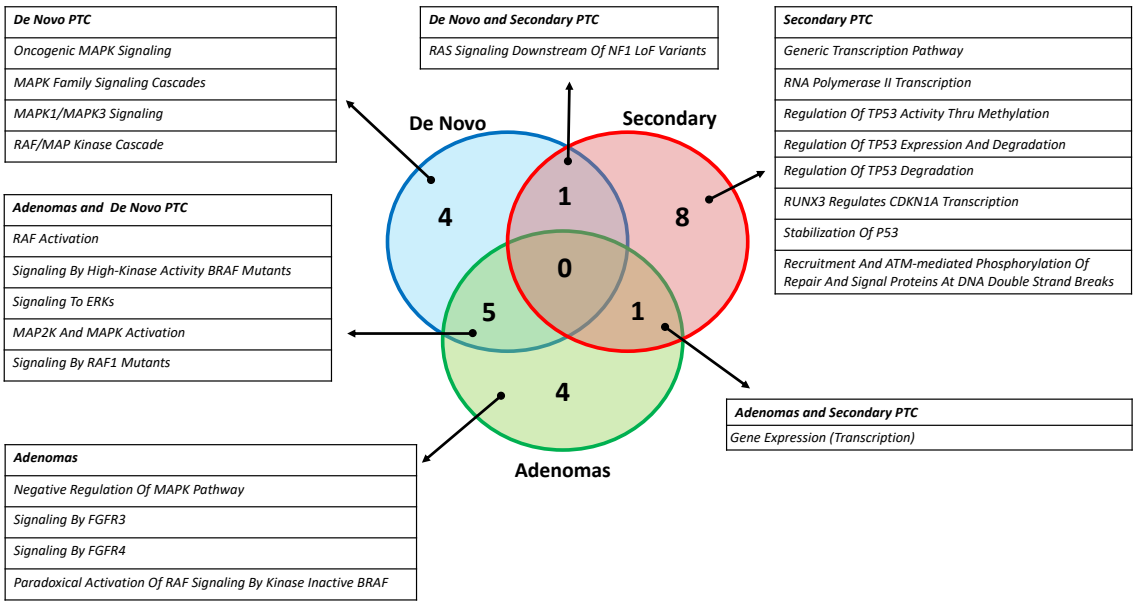



**Figure 5.16. Venn Diagram of Enriched Pathways in the Examined Populations**
The diagram illustrates the differentially enriched pathways in each population.

## 6. DISCUSSION

The molecular landscape of PTC has been largely studied and the genomic hallmarks of this disease are well established (78, 83). However, it is still unknown whether PTC developing in subjects exposed to a volcanic environment present peculiar features. On top of that, while the biological profile of radiation-induced PTC has been investigated (89), the molecular characteristics of PTC diagnosed in patients previously treated with chemotherapy and/or radiotherapy for a pediatric cancer (i.e. secondary PTC) remains controversial. Lastly, it is known that thyroid adenomas may present molecular alterations coherent with their follicular origin (116-118), but it still to determine if the exposure to volcanic pollution may interfere with their clonal evolution and malignancy potential.

To shed light on these controversies, we performed a NGS analysis on a cohort of thyroid neoplasms, including de novo PTC, secondary PTC and thyroid adenomas, diagnosed in patients from the volcanic area of Mount Etna Eastern Sicily (i.e. the province of Catania).

The NGS analysis was performed using a custom panel, specifically designed to address the most common genomic alterations of PTC, as reported by TCGA (78). Validation of this panel is ongoing and will be finalized once an adequate number of target sequencing will be performed. Samples sequenced for the present study will be included in the validation procedures.

Despite small (i.e. 15 de novo PTC, 7 secondary PTC and 16 adenomas), our study population recapitulate the main epidemiological characteristics of thyroid neoplasms, including the prevalence of female patients and the mean age of onset (125). In line with

what already reported, in the secondary PTC cohort previous malignancies were mainly hematological (Hodgkin lymphomas and leukemias). Also, the mean time from primary tumor to PTC diagnosis (i.e. 18.6 years) is coherent with previous literature data (48).

Looking at the results of molecular analysis in the de novo PTC population, it can be noticed that the mean number of molecular alterations is higher than expected (3.4 average alterations per patient). Indeed in the TCGA a low mutational density was observed among the analyzed PTC (78). In line with TCGA data, these tumors were BRAF-driven in 60% of cases, harboring the typical V600E substitution in 9/11 cases and one, less common, copy number gain. All the BRAF mutations were observed in classical variant PTC. Of note, half of the *BRAF*-altered samples also displayed a *HRAS* or *KRAS* SNV. Thus far, these alterations have been considered mutually exclusive (15). However, all the retrieved *KRAS* mutations occurred in the non-coding region of the gene, and despite pathogenic according to the applied prediction methods, their oncogenicity and clinical significance need further clarification.

In the same cohort, FV PTC were RAS-driven in 2/3 cases, while the third presented a *RET/NCOA4* rearrangement. This latter event has been rarely reported in FV PTC, being these tumors rather enriched in PPARG fusions (126-128).

The small cohort of secondary PTC showed a broadly heterogenous profile and a very high number of molecular alterations (average 9.2 per patient). However, these samples presented two significantly different mutational patterns, according to previous exposure to radiotherapy. Those occurring in patients who received radiation therapy, with or without chemotherapy, presented an extremely high mutational load. These hypermutated patients (S01, S06, S07) harbored redundant molecular alterations in genes involved in MAPK signaling (*BRAF, HRAS, KRAS, NRAS, NF1*), but also in cell

adhesion (*APC, CDH4*), epigenetic regulation (*KMT2A, KMT2C*), DNA repair (*ARID1B, ATM, CHEK2*) and cell cycle regulation (*TP53, RB1, ZFHX3*). The abovementioned complexity mirrors a vast clonal heterogeneity in radiation-induced tumors. This is in line with a recent report regarding a large cohort of survivors to Chernobyl disaster who later developed a PTC (89). According to this study, DNA double-strand breaks are the first oncogenic event and promote subsequent PTC development through the canonical aberrations of MAPK pathway (89).

On the other hand, patients from our secondary PTC cohort previously exposed to chemotherapy only showed a low mutational burden (3.6 average alteration per patient), resembling the one of de novo PTC. Among these subjects, SNV were heterogeneously scattered and involved BRAF V600E in one case only.

Even though the small number of patients does not allow to drive definitive conclusions, these findings seem to suggest that two distinct molecular profile co-exist in the population of secondary PTC. In the future, correlation with clinical outcomes could elucidate whether hypermutated tumors display an aggressive clinical behavior as expected.

Of note all of the tumors in this cohort, either radiotherapy or chemotherapy exposed, showed a mutation of *ZFHX3*. This gene, also known as AT-motif binding factor 1, codifies for a zinc-finger transcription factor with multiple homeodomains, and functionally exert an oncosuppressor role (129). Loss-of-function of *ZFHX3* have been retrieved in several tumor types, including prostate and breast cancer (130, 131). However, in TCGA the prevalence of this gene's mutations was only 1.7% among de novo PTC (78).

Taken together, our PTC populations, either de novo or secondary, showed a lower incidence of gene rearrangements compared to those observed in TCGA. Especially, we failed to identify any fusion in *PPARG*, *NTRK1*, *NTRK3* and *THADA*.

In the follicular adenomas cohort, *KRAS* SNV were the most frequent alterations, retrieved in half of the cases. As for the de novo PTC population, most of these mutations occurred in the non-coding region of the gene. *BRAF* was seldom altered with only one non canonical R682Q SNV and one copy number gain. This is coherent with the follicular origin of these lesions which represent a known precursors of FTC (82). A copy number loss of *EIF1AX* occurred in 3 of these samples. This gene codifies for an X-linked eukaryotic translation factor, and has been found mutated in different cancer types (132). *EIF1AX* was among the significantly mutated genes in PTC according to TCGA, and SNV have also been reported in adenomas (78). However, the biological significance of *EIF1AX* loss is unknow.

An integrated analysis of alterations found in the three study populations showed an overlap in genes related to the MAPK cascade (*BRAF, HRAS, KRAS*) and DNA methylation (KTM2A, KTM2C). Secondary PTC display the higher number of non-shared genes (n=5), and they were mainly related to DNA repair, gene expression and cell cycle regulation (*ARID1B, CHEK2, EZH1, PPM1D, RB1*). On the contrary, all the altered genes retrieved in adenomas were in common with secondary PTC or with both de novo and secondary PTC. These findings confirm that MAPK signaling plays a central role in thyroid neoplasm, that secondary PTC are also enriched with alterations related to other biological functions and that follicular adenomas share common features with PTC, despite being historically related to the development of FTC (82).

Protein-protein interaction networks explored by STRING analysis corroborate these assumptions. Overall, STRING analysis showed a simple clustering pattern in the three populations. In the de novo PTC cohort, altered genes corresponded to a main altered cluster of PPI, with MAPK components establishing most of the connections. In the secondary PTC population, protein clustering was similar, but additional proteins were present at the core of the cluster, including TP53, PTEN and ATM. Lastly, in the adenoma cohort a single PPI cluster emerged, again centered on MAPK signaling proteins, but with simpler interaction networks compared to de novo PTC.

A different scenario emerges from the analysis of enriched biological functions (according to Gene Ontology library) and pathways (according to Reactome library).

According to Gene Ontology analysis, de novo and secondary PTC share some enriched biological functions, mainly inherent to MAPK signaling and response to gamma radiation, while adenomas are rather enriched with terms not present in the other cohorts, including functions related to cell adhesion, motility and migrations.

The differences between the studied populations are even more straightforward looking at the results of the Reactome analysis. Indeed, de novo tumors are significantly enriched with pathways converging on MAPK signaling, whereas secondary tumors display an enrichment in several terms not clustering each other. These pathways are involved in TP53 activation and regulation and in DNA transcription and repair. It is possible that these biological differences between de novo and secondary PTC may underlie distinct clinical behaviors and potentially, a different response to targeted therapies.

Lastly, adenomas are also enriched in terms regarding MAPK pathway but display additional FGFR-related pathways.

The present study have several, worth mentioning, limitations. First of all, the study population included a limited number of cases. Of note, the recruitment has been severely slowed down by the break of Coronavirus pandemic since 2020. However, we plan to expand the analysis in the upcoming future. Hence, the results reported here should be considered as preliminary. For the secondary PTC cohort, we plan to sequence several cases from a non-volcanic area (Istituto dei Tumori, Milano, Italy), in order to score potential differences with secondary PTC diagnosed in patients exposed to a vlcanic environment. We also aim to include additional clinical information to correlate with the molecular findings.

Due to technical issues, the custom NGS panel does not include sequencing of *TERT* promoter, which can be altered in PTC and is known to confer aggressiveness and a worse prognosis (133, 134). Since we acknowledge this important limitation, we have decided to separately evaluate *TERT* promoter alterations in our cohort. Still, we do not deem that this additional test will change the overall results of our analysis.

In conclusion, this is the first attempt to molecularly characterize papillary thyroid tumors, either the novo or occurring after chemo-radiotherapy exposure, and thyroid adenomas in a cohort of patients from a volcanic area. Our study suggest that the molecular landscape of PTC developing in subjects exposed to a volcanic environment is partly concordant with pre-existing literature evidence. However, the simultaneous presence of *BRAF* V600E and *H or NRAS* mutations represent a novel and unexpected finding. Moreover, individuals previously exposed to radiotherapy, but not those exposed to chemotherapy only, seem to have a greater risk to develop a hypermutated, and probably more aggressive, PTC. Lastly, the profile of thyroid adenomas was enriched

with SNV and CNV converging on biological functions and pathways that was substantially different from those observed in PTC.

These findings may have significant clinical implications and thus deserve further validation on larger and prospective cohorts.

## 7. REFERENCES


1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7-34.

2. https://www.aiom.it/wp-content/uploads/2022/12/2022_AIOM_NDC-web.pdf. Accessed on Jan 31, 2023.

3. Force USPST, Bibbins-Domingo K, Grossman DC, Curry SJ, Barry MJ, Davidson KW, et al. Screening for Thyroid Cancer: US Preventive Services Task Force Recommendation Statement. JAMA. 2017;317(18):1882-7.

4. Pellegriti G, Frasca F, Regalbuto C, Squatrito S, Vigneri R. Worldwide increasing incidence of thyroid cancer: update on epidemiology and risk factors. J Cancer Epidemiol. 2013;2013:965212.

5. Miranda-Filho A, Lortet-Tieulent J, Bray F, Cao B, Franceschi S, Vaccarella S, et al. Thyroid cancer incidence trends by histology in 25 countries: a population-based study. Lancet Diabetes Endocrinol. 2021;9(4):225-34.

6. Vigneri R, Malandrino P, Giani F, Russo M, Vigneri P. Heavy metals in the volcanic environment and thyroid cancer. Mol Cell Endocrinol. 2017;457:73-80.

7. Vigneri R, Malandrino P, Vigneri P. The changing epidemiology of thyroid cancer: why is incidence increasing? Curr Opin Oncol. 2015;27(1):1-7.

8. Dal Maso L, Tavilla A, Pacini F, Serraino D, van Dijk BAC, Chirlaque MD, et al. Survival of 86,690 patients with thyroid cancer: A population-based study in 29 European countries from EUROCARE-5. Eur J Cancer. 2017;77:140-52.

9. Davies L, Welch HG. Current thyroid cancer trends in the United States. JAMA Otolaryngol Head Neck Surg. 2014;140(4):317-22.

10. Jonklaas J, Nogueras-Gonzalez G, Munsell M, Litofsky D, Ain KB, Bigos ST, et al. The impact of age and gender on papillary thyroid cancer survival. J Clin Endocrinol Metab. 2012;97(6):E878-87.

11. Khan YS, Farhana A. Histology, Thyroid Gland. StatPearls. Treasure Island (FL)2022.

12. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. Lancet. 2016;388(10061):2783-95.

13. Christofer Juhlin C, Mete O, Baloch ZW. The 2022 WHO classification of thyroid tumors: novel concepts in nomenclature and grading. Endocr Relat Cancer. 2023;30(2).

14. Baloch ZW, Asa SL, Barletta JA, Ghossein RA, Juhlin CC, Jung CK, et al. Overview of the 2022 WHO Classification of Thyroid Neoplasms. Endocr Pathol. 2022;33(1):27-63.

15. Giordano TJ. Genomic Hallmarks of Thyroid Neoplasia. Annu Rev Pathol. 2018;13:141-62.

16. de la Fouchardiere C, Decaussin-Petrucci M, Berthiller J, Descotes F, Lopez J, Lifante JC, et al. Predictive factors of outcome in poorly differentiated thyroid carcinomas. Eur J Cancer. 2018;92:40-7.

17. Saini S, Tulla K, Maker AV, Burman KD, Prabhakar BS. Therapeutic advances in anaplastic thyroid cancer: a current perspective. Mol Cancer. 2018;17(1):154.

18. Randle RW, Balentine CJ, Leverson GE, Havlena JA, Sippel RS, Schneider DF, et al. Trends in the presentation, treatment, and survival of patients with medullary thyroid cancer over the past 30 years. Surgery. 2017;161(1):137-46.

19. Thomas CM, Asa SL, Ezzat S, Sawka AM, Goldstein D. Diagnosis and pathologic characteristics of medullary thyroid carcinoma-review of current guidelines. Curr Oncol. 2019;26(5):338-44.

20. Sherman SI. Thyroid carcinoma. Lancet. 2003;361(9356):501-11.

21. Durante C, Haddy N, Baudin E, Leboulleux S, Hartl D, Travagli JP, et al. Long-term outcome of 444 patients with distant metastases from papillary and follicular thyroid carcinoma: benefits and limits of radioiodine therapy. J Clin Endocrinol Metab. 2006;91(8):2892-9.

22. Schlumberger M, Brose M, Elisei R, Leboulleux S, Luster M, Pitoia F, et al. Definition and management of radioactive iodine-refractory differentiated thyroid cancer. Lancet Diabetes Endocrinol. 2014;2(5):356-8.

23. Filetti S, Durante C, Hartl D, Leboulleux S, Locati LD, Newbold K, et al. Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-updagger. Ann Oncol. 2019;30(12):1856-83.

24. Pizzimenti C, Fiorentino V, Ieni A, Martini M, Tuccari G, Lentini M, et al. Aggressive variants of follicular cell-derived thyroid carcinoma: an overview. Endocrine. 2022;78(1):1-12.

25. Faugeras L, Pirson AS, Donckier J, Michel L, Lemaire J, Vandervorst S, et al. Refractory thyroid carcinoma: which systemic treatment to use? Ther Adv Med Oncol. 2018;10:1758834017752853.

26. Subbiah V, Kreitman RJ, Wainberg ZA, Cho JY, Schellens JHM, Soria JC, et al. Dabrafenib and Trametinib Treatment in Patients With Locally Advanced or Metastatic BRAF V600-Mutant Anaplastic Thyroid Cancer. J Clin Oncol. 2018;36(1):7-13.

27. Kiesewetter B, Riss P, Scheuba C, Raderer M. How I treat medullary thyroid cancer. ESMO Open. 2021;6(3):100183.

28. Kung TM, Ng WL, Gibson JB. Volcanoes and carcinoma of the thyroid: a possible association. Arch Environ Health. 1981;36(5):265-7.

29. Paksoy N, Montaville B, McCarthy SW. Cancer occurrence in Vanuatu in the South Pacific, 1980-86. Asia Pac J Public Health. 1989;3(3):231-6.

30. Curado M-P, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al. Cancer incidence in five continents, Volume IX: IARC Press, International Agency for Research on Cancer; 2007.

31. Truong T, Rougier Y, Dubourdieu D, Guihenneuc-Jouyaux C, Orsi L, Hemon D, et al. Time trends and geographic variations for thyroid cancer in New Caledonia, a very high incidence area (1985-1999). Eur J Cancer Prev. 2007;16(1):62-70.

32. Pellegriti G, De Vathaire F, Scollo C, Attard M, Giordano C, Arena S, et al. Papillary thyroid cancer incidence in the volcanic area of Sicily. J Natl Cancer Inst. 2009;101(22):1575-83.

33. Boffetta P, Memeo L, Giuffrida D, Ferrante M, Sciacca S. Exposure to emissions from Mount Etna (Sicily, Italy) and incidence of thyroid cancer: a geographic analysis. Sci Rep. 2020;10(1):21298.

34. Malandrino P, Russo M, Ronchi A, Minoia C, Cataldo D, Regalbuto C, et al. Increased thyroid cancer incidence in a basaltic volcanic area is associated with non-anthropogenic pollution and biocontamination. Endocrine. 2016;53(2):471-9.

35. Malandrino P, Scollo C, Marturano I, Russo M, Tavarelli M, Attard M, et al. Descriptive epidemiology of human thyroid cancer: experience from a regional registry and the "volcanic factor". Front Endocrinol (Lausanne). 2013;4:65.

36. Pellegriti G, Malandrino P, Vigneri R. Response: Re: Papillary Thyroid Cancer Incidence in the Volcanic Area of Sicily. JNCI: Journal of the National Cancer Institute. 2010;102(12):915-6.

37. Ferrante M, Fiore M, Ledda C, Cicciù F, Alonzo E, Fallico R, et al. [Monitoring of heavy metals and trace elements in the air, fruits and vegetables and soil in the province of Catania (Italy)]. Igiene e sanità pubblica. 2013;69(1):47-54.

38. Varrica D, Tamburo E, Dongarra G, Sposito F. Trace elements in scalp hair of children chronically exposed to volcanic activity (Mt. Etna, Italy). Sci Total Environ. 2014;470-471:117-26.

39. Marcello MA, Malandrino P, Almeida JF, Martins MB, Cunha LL, Bufalo NE, et al. The influence of the environment on the development of thyroid tumors: a new appraisal. Endocr Relat Cancer. 2014;21(5):T235-54.

40. Boelaert K. The association between serum TSH concentration and thyroid cancer. Endocr Relat Cancer. 2009;16(4):1065-72.

41. Giani F, Masto R, Trovato MA, Malandrino P, Russo M, Pellegriti G, et al. Heavy Metals in the Environment and Thyroid Cancer. Cancers (Basel). 2021;13(16).

42. Giani F, Pandini G, Scalisi NM, Vigneri P, Fazzari C, Malandrino P, et al. Effect of low-dose tungsten on human thyroid stem/precursor cells and their progeny. Endocr Relat Cancer. 2019;26(8):713-25.

43. Balali-Mood M, Naseri K, Tahergorabi Z, Khazdair MR, Sadeghi M. Toxic Mechanisms of Five Heavy Metals: Mercury, Lead, Chromium, Cadmium, and Arsenic. Front Pharmacol. 2021;12:643972.

44. Inskip PD, Curtis RE. New malignancies following childhood cancer in the United States, 1973-2002. Int J Cancer. 2007;121(10):2233-40.

45. Olsen JH, Moller T, Anderson H, Langmark F, Sankila R, Tryggvadottir L, et al. Lifelong cancer incidence in 47,697 patients treated for childhood cancer in the Nordic countries. J Natl Cancer Inst. 2009;101(11):806-13.

46. Taylor AJ, Croft AP, Palace AM, Winter DL, Reulen RC, Stiller CA, et al. Risk of thyroid cancer in survivors of childhood cancer: results from the British Childhood Cancer Survivor Study. Int J Cancer. 2009;125(10):2400-5.

47. Friedman DL, Whitton J, Leisenring W, Mertens AC, Hammond S, Stovall M, et al. Subsequent neoplasms in 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. J Natl Cancer Inst. 2010;102(14):1083-95.

48. Wijnen M, van den Heuvel-Eibrink MM, Medici M, Peeters RP, van der Lely AJ, Neggers SJ. Risk factors for subsequent endocrine-related cancer in childhood cancer survivors. Endocr Relat Cancer. 2016;23(6):R299-321.

49. Black P, Straaten A, Gutjahr P. Secondary thyroid carcinoma after treatment for childhood cancer. Med Pediatr Oncol. 1998;31(2):91-5.

50. Acharya S, Sarafoglou K, LaQuaglia M, Lindsley S, Gerald W, Wollner N, et al. Thyroid neoplasms after therapeutic radiation for malignancies during childhood or adolescence. Cancer. 2003;97(10):2397-403.

51. Cohen A, Rovelli A, Merlo DF, van Lint MT, Lanino E, Bresters D, et al. Risk for secondary thyroid carcinoma after hematopoietic stem-cell transplantation: an EBMT Late Effects Working Party Study. J Clin Oncol. 2007;25(17):2449-54.

52. Constine LS, Tarbell N, Hudson MM, Schwartz C, Fisher SG, Muhs AG, et al. Subsequent malignancies in children treated for Hodgkin's disease: associations with gender and radiation dose. Int J Radiat Oncol Biol Phys. 2008;72(1):24-33.

53. Diallo I, Haddy N, Adjadj E, Samand A, Quiniou E, Chavaudra J, et al. Frequency distribution of second solid cancer locations in relation to the irradiated volume among 115 patients treated for childhood cancer. Int J Radiat Oncol Biol Phys. 2009;74(3):876-83.

54. Veiga LH, Bhatti P, Ronckers CM, Sigurdson AJ, Stovall M, Smith SA, et al. Chemotherapy and thyroid cancer risk: a report from the childhood cancer survivor study. Cancer Epidemiol Biomarkers Prev. 2012;21(1):92-101.

55. Danner-Koptik KE, Majhail NS, Brazauskas R, Wang Z, Buchbinder D, Cahn JY, et al. Second malignancies after autologous hematopoietic cell transplantation in children. Bone Marrow Transplant. 2013;48(3):363-8.

56. Brignardello E, Felicetti F, Castiglione A, Gallo M, Maletta F, Isolato G, et al. Ultrasound surveillance for radiation-induced thyroid carcinoma in adult survivors of childhood cancer. Eur J Cancer. 2016;55:74-80.

57. Bhatia S. Genetic variation as a modifier of association between therapeutic exposure and subsequent malignant neoplasms in cancer survivors. Cancer. 2015;121(5):648-63.

58. Bhatti P, Veiga LH, Ronckers CM, Sigurdson AJ, Stovall M, Smith SA, et al. Risk of second primary thyroid cancer after radiotherapy for a childhood cancer in a large cohort study: an update from the childhood cancer survivor study. Radiat Res. 2010;174(6):741-52.

59. Ronckers CM, Sigurdson AJ, Stovall M, Smith SA, Mertens AC, Liu Y, et al. Thyroid cancer in childhood cancer survivors: a detailed evaluation of radiation dose response and its modifiers. Radiat Res. 2006;166(4):618-28.

60. Turcotte LM, Neglia JP, Reulen RC, Ronckers CM, van Leeuwen FE, Morton LM, et al. Risk, Risk Factors, and Surveillance of Subsequent Malignant Neoplasms in Survivors of Childhood Cancer: A Review. J Clin Oncol. 2018;36(21):2145-52.

61. Sherif RS, Elshemey WM, Attalla EM. The risk of secondary cancer in pediatric medulloblastoma patients due to three-dimensional conformal radiotherapy and intensity-modulated radiotherapy. Indian J Cancer. 2018;55(4):372-6.

62. Seravalli E, Bosman M, Lassen-Ramshad Y, Vestergaard A, Oldenburger F, Visser J, et al. Dosimetric comparison of five different techniques for craniospinal irradiation across 15 European centers: analysis on behalf of the SIOP-E-BTG (radiotherapy working group)(). Acta Oncol. 2018;57(9):1240-9.

63. Waguespack SG. Thyroid Sequelae of Pediatric Cancer Therapy. Horm Res Paediatr. 2019;91(2):104-17.

64. de Vathaire F, Haddy N, Allodji RS, Hawkins M, Guibout C, El-Fayech C, et al. Thyroid Radiation Dose and Other Risk Factors of Thyroid Carcinoma Following Childhood Cancer. J Clin Endocrinol Metab. 2015;100(11):4282-90.

65. Veiga LH, Holmberg E, Anderson H, Pottern L, Sadetzki S, Adams MJ, et al. Thyroid Cancer after Childhood Exposure to External Radiation: An Updated Pooled Analysis of 12 Studies. Radiat Res. 2016;185(5):473-84.

66. Gebauer J, Higham C, Langer T, Denzer C, Brabant G. Long-Term Endocrine and Metabolic Consequences of Cancer Treatment: A Systematic Review. Endocr Rev. 2019;40(3):711-67.

67. Maule M, Scelo G, Pastore G, Brennan P, Hemminki K, Pukkala E, et al. Risk of second malignant neoplasms after childhood central nervous system malignant tumours: an international study. Eur J Cancer. 2008;44(6):830-9.

68. Maule M, Scelo G, Pastore G, Brennan P, Hemminki K, Olsen JH, et al. Second malignancies after childhood noncentral nervous system solid cancer: Results from 13 cancer registries. Int J Cancer. 2011;129(8):1940-52.

69. Maule M, Scelo G, Pastore G, Brennan P, Hemminki K, Tracey E, et al. Risk of second malignant neoplasms after childhood leukemia and lymphoma: an international study. J Natl Cancer Inst. 2007;99(10):790-800.

70. Shiovitz S, Korde LA. Genetics of breast cancer: a topic in evolution. Ann Oncol. 2015;26(7):1291-9.

71. Strahm B, Malkin D. Hereditary cancer predisposition in children: genetic basis and clinical implications. Int J Cancer. 2006;119(9):2001-6.

72. Pilarski R, Nagy R. Genetic testing by cancer site: endocrine system. Cancer J. 2012;18(4):364-71.

73. Foulkes WD, Priest JR, Duchaine TF. DICER1: mutations, microRNAs and mechanisms. Nat Rev Cancer. 2014;14(10):662-72.

74. Akulevich NM, Saenko VA, Rogounovitch TI, Drozd VM, Lushnikov EF, Ivanov VK, et al. Polymorphisms of DNA damage response genes in radiation-related and sporadic papillary thyroid carcinoma. Endocr Relat Cancer. 2009;16(2):491-503.

75. Best T, Li D, Skol AD, Kirchhoff T, Jackson SA, Yasui Y, et al. Variants at 6q21 implicate PRDM1 in the etiology of therapy-induced second malignancies after Hodgkin's lymphoma. Nat Med. 2011;17(8):941-3.

76. Takahashi M, Saenko VA, Rogounovitch TI, Kawaguchi T, Drozd VM, Takigawa-Imamura H, et al. The FOXE1 locus is a major genetic determinant for radiation-related thyroid carcinoma in Chernobyl. Hum Mol Genet. 2010;19(12):2516-23.

77. Damiola F, Byrnes G, Moissonnier M, Pertesi M, Deltour I, Fillon A, et al. Contribution of ATM and FOXE1 (TTF2) to risk of papillary thyroid carcinoma in Belarusian children exposed to radiation. Int J Cancer. 2014;134(7):1659-68.

78. Cancer Genome Atlas Research N. Integrated genomic characterization of papillary thyroid carcinoma. Cell. 2014;159(3):676-90.

79. de Biase D, Cesari V, Visani M, Casadei GP, Cremonini N, Gandolfi G, et al. High-sensitivity BRAF mutation analysis: BRAF V600E is acquired early during tumor development but is heterogeneously distributed in a subset of papillary thyroid carcinomas. J Clin Endocrinol Metab. 2014;99(8):E1530-8.

80. Knauf JA, Sartor MA, Medvedovic M, Lundsmith E, Ryder M, Salzano M, et al. Progression of BRAF-induced thyroid cancer is associated with epithelial-mesenchymal transition requiring concomitant MAP kinase and TGFbeta signaling. Oncogene. 2011;30(28):3153-62.

81. Abdullah MI, Junit SM, Ng KL, Jayapalan JJ, Karikalan B, Hashim OH. Papillary Thyroid Cancer: Genetic Alterations and Molecular Biomarker Investigations. Int J Med Sci. 2019;16(3):450-60.

82. Nikiforov YE, Nikiforova MN. Molecular genetics and diagnosis of thyroid cancer. Nat Rev Endocrinol. 2011;7(10):569-80.

83. Valvo V, Nucera C. Coding Molecular Determinants of Thyroid Cancer Development and Progression. Endocrinol Metab Clin North Am. 2019;48(1):37-59.

84. Santoro M, Melillo RM, Fusco A. RET/PTC activation in papillary thyroid carcinoma: European Journal of Endocrinology Prize Lecture. Eur J Endocrinol. 2006;155(5):645-53.

85. Vuong HG, Le HT, Le TTB, Le T, Hassell L, Kakudo K. Clinicopathological significance of major fusion oncogenes in papillary thyroid carcinoma: An individual patient data meta-analysis. Pathol Res Pract. 2022;240:154180.

86. Liu Y, Cope L, Sun W, Wang Y, Prasad N, Sangenario L, et al. DNA copy number variations characterize benign and malignant thyroid tumors. J Clin Endocrinol Metab. 2013;98(3):E558-66.

87. Ory C, Ugolin N, Levalois C, Lacroix L, Caillou B, Bidart JM, et al. Gene expression signature discriminates sporadic from post-radiotherapy-induced thyroid tumors. Endocr Relat Cancer. 2011;18(1):193-206.

88. Handkiewicz-Junak D, Swierniak M, Rusinek D, Oczko-Wojciechowska M, Dom G, Maenhaut C, et al. Gene signature of the post-Chernobyl papillary thyroid cancer. Eur J Nucl Med Mol Imaging. 2016;43(7):1267-77.

89. Morton LM, Karyadi DM, Stewart C, Bogdanova TI, Dawson ET, Steinberg MK, et al. Radiation-related genomic profile of papillary thyroid carcinoma after the Chernobyl accident. Science. 2021;372(6543).

90. Selmansberger M, Braselmann H, Hess J, Bogdanova T, Abend M, Tronko M, et al. Genomic copy number analysis of Chernobyl papillary thyroid carcinoma in the Ukrainian-American Cohort. Carcinogenesis. 2015;36(11):1381-7.

91. Hess JR, Newbern DK, Beebe KL, Walsh AM, Schafernak KT. High Prevalence of Gene Fusions and Copy Number Alterations in Pediatric Radiation Therapy-Induced Papillary and Follicular Thyroid Carcinomas. Thyroid. 2022;32(4):411-20.

92. Prescott JD, Zeiger MA. The RET oncogene in papillary thyroid carcinoma. Cancer. 2015;121(13):2137-46.

93. Salvatore D, Santoro M, Schlumberger M. The importance of the RET gene in thyroid cancer and therapeutic implications. Nat Rev Endocrinol. 2021;17(5):296-306.

94. Ciampi R, Knauf JA, Kerler R, Gandhi M, Zhu Z, Nikiforova MN, et al. Oncogenic AKAP9-BRAF fusion is a novel mechanism of MAPK pathway activation in thyroid cancer. J Clin Invest. 2005;115(1):94-101.

95. Ciampi R, Knauf JA, Rabes HM, Fagin JA, Nikiforov YE. BRAF kinase activation via chromosomal rearrangement in radiation-induced and sporadic thyroid cancer. Cell Cycle. 2005;4(4):547-8.

96. Solomon JP, Benayed R, Hechtman JF, Ladanyi M. Identifying patients with NTRK fusion cancer. Ann Oncol. 2019;30 Suppl 8:viii16-viii22.

97. Greco A, Miranda C, Pierotti MA. Rearrangements of NTRK1 gene in papillary thyroid carcinoma. Mol Cell Endocrinol. 2010;321(1):44-9.

98. Brzezianska E, Karbownik M, Migdalska-Sek M, Pastuszak-Lewandoska D, Wloch J, Lewinski A. Molecular analysis of the RET and NTRK1 gene rearrangements in papillary thyroid carcinoma in the Polish population. Mutat Res. 2006;599(1-2):26-35.

99. Tirro E, Martorana F, Romano C, Vitale SR, Motta G, Di Gregorio S, et al. Molecular Alterations in Thyroid Cancer: From Bench to Clinical Practice. Genes (Basel). 2019;10(9).

100. McHenry CR, Phitayakorn R. Follicular adenoma and carcinoma of the thyroid gland. Oncologist. 2011;16(5):585-93.

101. Bisi H, Fernandes VS, de Camargo RY, Koch L, Abdo AH, de Brito T. The prevalence of unsuspected thyroid pathology in 300 sequential autopsies, with special reference to the incidental carcinoma. Cancer. 1989;64(9):1888-93.

102. Arem R, Padayatty SJ, Saliby AH, Sherman SI. Thyroid microcarcinoma: prevalence, prognosis, and management. Endocr Pract. 1999;5(3):148-56.

103. Zablotska LB, Nadyrov EA, Polyanskaya ON, McConnell RJ, O'Kane P, Lubin J, et al. Risk of thyroid follicular adenoma among children and adolescents in Belarus exposed to iodine-131 after the Chornobyl accident. Am J Epidemiol. 2015;182(9):781-90.

104. Mete O, Asa SL. Pitfalls in the diagnosis of follicular epithelial proliferations of the thyroid. Adv Anat Pathol. 2012;19(6):363-73.

105. Kamilaris CDC, Faucz FR, Voutetakis A, Stratakis CA. Carney Complex. Exp Clin Endocrinol Diabetes. 2019;127(2-03):156-64.

106. Spencer T, Pan KS, Collins MT, Boyce AM. The Clinical Spectrum of McCune-Albright Syndrome and Its Management. Horm Res Paediatr. 2019;92(6):347-56.

107. Jung CK, Bychkov A, Kakudo K. Update from the 2022 World Health Organization Classification of Thyroid Tumors: A Standardized Diagnostic Approach. Endocrinol Metab (Seoul). 2022;37(5):703-18.

108. Alexander EK, Doherty GM, Barletta JA. Management of thyroid nodules. Lancet Diabetes Endocrinol. 2022;10(7):540-8.

109. Kamran SC, Marqusee E, Kim MI, Frates MC, Ritner J, Peters H, et al. Thyroid nodule size and prediction of cancer. J Clin Endocrinol Metab. 2013;98(2):564-70.

110. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid. 2016;26(1):1-133.

111. Duan H, Liu X, Ren X, Zhang H, Wu H, Liang Z. Mutation profiles of follicular thyroid tumors by targeted sequencing. Diagn Pathol. 2019;14(1):39.

112. Madsen MB, Kiss K, Cilius Nielsen F, Bennedbaek FN, Rossing M. Amplicon-Based NGS Panels for Actionable Cancer Target Identification in Follicular Cell-Derived Thyroid Neoplasia. Front Endocrinol (Lausanne). 2020;11:146.

113. Barden CB, Shister KW, Zhu B, Guiter G, Greenblatt DY, Zeiger MA, et al. Classification of follicular thyroid tumors by molecular signature: results of gene profiling. Clin Cancer Res. 2003;9(5):1792-800.

114. Finley DJ, Zhu B, Barden CB, Fahey TJ, 3rd. Discrimination of benign and malignant thyroid nodules by molecular profiling. Ann Surg. 2004;240(3):425-36; discussion 36-7.

115. Fryknas M, Wickenberg-Bolin U, Goransson H, Gustafsson MG, Foukakis T, Lee JJ, et al. Molecular markers for discrimination of benign and malignant follicular thyroid tumors. Tumour Biol. 2006;27(4):211-20.

116. Dom G, Frank S, Floor S, Kehagias P, Libert F, Hoang C, et al. Thyroid follicular adenomas and carcinomas: molecular profiling provides evidence for a continuous evolution. Oncotarget. 2018;9(12):10343-59.

117. Giordano TJ, Beaudenon-Huibregtse S, Shinde R, Langfield L, Vinco M, Laosinchai-Wolf W, et al. Molecular testing for oncogenic gene mutations in thyroid lesions: a case-control validation study in 413 postsurgical specimens. Hum Pathol. 2014;45(7):1339-47.

118. Schulten HJ, Salama S, Al-Ahmadi A, Al-Mansouri Z, Mirza Z, Al-Ghamdi K, et al. Comprehensive survey of HRAS, KRAS, and NRAS mutations in proliferative thyroid lesions from an ethnically diverse population. Anticancer Res. 2013;33(11):4779-84.

119. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248-9.

120. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021;49(D1):D605-D12.

121. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1):W90-7.

122. Clarke DJB, Jeon M, Stein DJ, Moiseyev N, Kropiwnicki E, Dai C, et al. Appyters: Turning Jupyter Notebooks into data-driven web apps. Patterns (N Y). 2021;2(3):100213.

123. Dongen SV. Graph Clustering Via a Discrete Uncoupling Process. SIAM Journal on Matrix Analysis and Applications. 2008;30(1):121-41.

124. Haynes W. Benjamini–Hochberg Method. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H, editors. Encyclopedia of Systems Biology. New York, NY: Springer New York; 2013. p. 78-.

125. Carling T, Udelsman R. Thyroid cancer. Annu Rev Med. 2014;65:125-37.

126. Zhu Z, Gandhi M, Nikiforova MN, Fischer AH, Nikiforov YE. 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. 2003;120(1):71-7.

127. Rivera M, Ricarte-Filho J, Knauf J, Shaha A, Tuttle M, Fagin JA, et al. Molecular genotyping of papillary thyroid carcinoma follicular variant according to its histological

subtypes (encapsulated vs infiltrative) reveals distinct BRAF and RAS mutation patterns. Mod Pathol. 2010;23(9):1191-200.

128. Wreesmann VB, Ghossein RA, Hezel M, Banerjee D, Shaha AR, Tuttle RM, et al. Follicular variant of papillary thyroid carcinoma: genome-wide appraisal of a controversial entity. Genes Chromosomes Cancer. 2004;40(4):355-64.

129. Miura Y, Tam T, Ido A, Morinaga T, Miki T, Hashimoto T, et al. Cloning and characterization of an ATBF1 isoform that expresses in a neuronal differentiation-dependent manner. J Biol Chem. 1995;270(45):26840-8.

130. Dayoub A, Fokin AI, Lomakina ME, James J, Plays M, Jacquin T, et al. Inactivation of PTEN and ZFHX3 in Mammary Epithelial Cells Alters Patterns of Collective Cell Migration. Int J Mol Sci. 2022;24(1).

131. Zhang J, Zhou N, Lin A, Luo P, Chen X, Deng H, et al. ZFHX3 mutation as a protective biomarker for immune checkpoint blockade in non-small cell lung cancer. Cancer Immunol Immunother. 2021;70(1):137-51.

132. Martin M, Masshofer L, Temming P, Rahmann S, Metz C, Bornfeld N, et al. Exome sequencing identifies recurrent somatic mutations in EIF1AX and SF3B1 in uveal melanoma with disomy 3. Nat Genet. 2013;45(8):933-6.

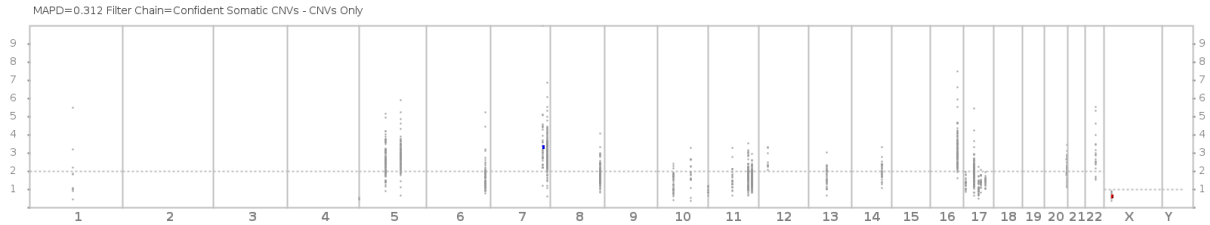
133. Liu R, Xing M. TERT promoter mutations in thyroid cancer. Endocr Relat Cancer. 2016;23(3):R143-55.

134. Liu X, Bishop J, Shan Y, Pai S, Liu D, Murugan AK, et al. Highly prevalent TERT promoter mutations in aggressive thyroid cancers. Endocr Relat Cancer. 2013;20(4):603-10.

# 8. SUPPLEMENTARY MATERIALS

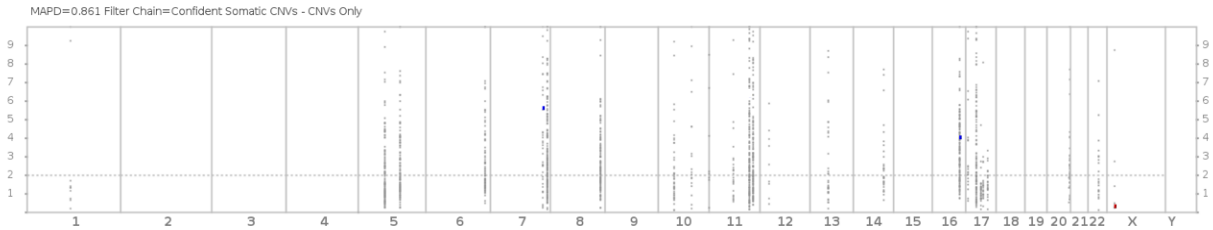
## Supplementary Table 1. Genes targeted by the Custom NGS DNA Panel

| # | Gene | Official Full Name | Ensemble Gene ID | Exon n° | Chromosome |
|---|---|---|---|---|---|
| **1** | *APC* | APC Regulator Of WNT Signaling Pathway | ENSG00000134982 | 20 | 5q22.2 |
| **2** | *ARID1B* | AT-rich interaction domain 1B | ENSG00000049618 | 26 | 6q25.3 |
| **3** | *ATM* | Ataxia Telangiectasia Mutated | ENSG00000149311 | 67 | 11q22.3 |
| **4** | *BDP1* | B Double Prime 1, Subunit Of RNA Polymerase III Transcription Initiation Factor IIIB | ENSG00000145734 | 41 | 5q13.2 |
| **5** | *BRAF* | B-Raf Proto-Oncogene, Serine/Threonine Kinase | ENSG00000157764 | 24 | 7q34 |
| **6** | *CDH4* | Cadherin 4 | ENSG00000179242 | 20 | 20q13.33 |
| **7** | *CHEK2* | Checkpoint kinase 2 | ENSG00000183765 | 21 | 22q12.1 |
| **8** | *EIF1AX* | Eukaryotic translation initiation factor 1A X-linked | ENSG00000173674 | 7 | Xp22.12 |
| **9** | *EZH1* | Enhancer of zeste 1 polycomb repressive complex 2 subunit | ENSG00000108799 | 21 | 17q21.2 |
| **10** | *HRAS* | HRas proto-oncogene, GTPase | ENSG00000174775 | 7 | 11p15.5 |
| **11** | *KRAS* | KRAS proto-oncogene, GTPase | ENSG00000133703 | 7 | 12p12.1 |
| **12** | *NRAS* | NRAS proto-oncogene, GTPase | ENSG00000213281 | 7 | 1p13.2 |
| **13** | *MEN1* | Menin 1 | ENSG00000133895 | 13 | 11q13.1 |
| **14** | *KMT2A* | Lysine methyltransferase 2A | ENSG00000118058 | 37 | 11q23.3 |
| **15** | *KMT2C* | Lysine methyltransferase 2C | ENSG00000055609 | 59 | 7q36.1 |
| **16** | *NF1* | Neurofibromin 1 | ENSG00000196712 | 58 | 17q11.2 |
| **17** | *PPMID* | Protein Phosphatase, Mg2+/Mn2+ Dependent 1D | ENSG00000170836 | 8 | 17q23.2 |
| **18** | *PTEN* | Phosphatase and tensin homolog | ENSG00000171862 | 10 | 10q23.31 |
| **19** | *RB1* | RB transcriptional corepressor 1 | ENSG00000139687 | 27 | 13q14.2 |
| **20** | *RET* | Ret proto-oncogene | ENSG00000165731 | 20 | 10q11.21 |
| **21** | *SPOP* | Speckle type BTB/POZ protein | ENSG00000121067 | 16 | 17q21.33 |
| **22** | *TG* | Thyroglobulin | ENSG00000042832 | 52 | 8q24.22 |
| **23** | *TP53* | Tumor protein p53 | ENSG00000141510 | 11 | 17p13.1 |
| **24** | *TSHR* | Thyroid stimulating hormone receptor | ENSG00000165409 | 11 | 14q31.1 |
| **25** | *ZFHX3* | Zinc finger homebox 3 | ENSG00000140836 | 23 | 16q22.2-q22.3 |



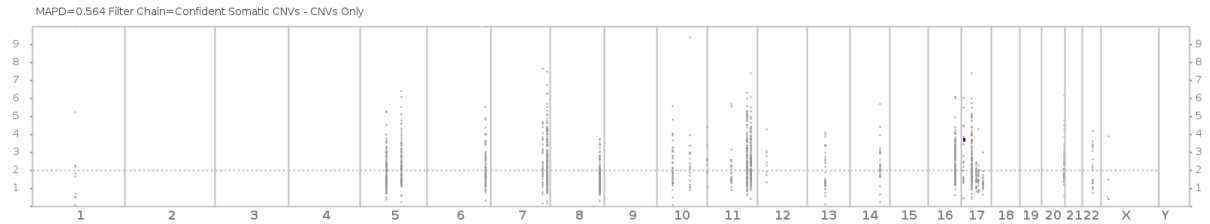

## Supplementary Figure 1. CNV identified in P11

The x axis represents the chromosome number; y axis represents the ploidy for each chromosome. The dashed horizontal line represents the normal ploidy condition. The blue spot indicates a copy number gain on chromosome 7, corresponding to *BRAF* gene.



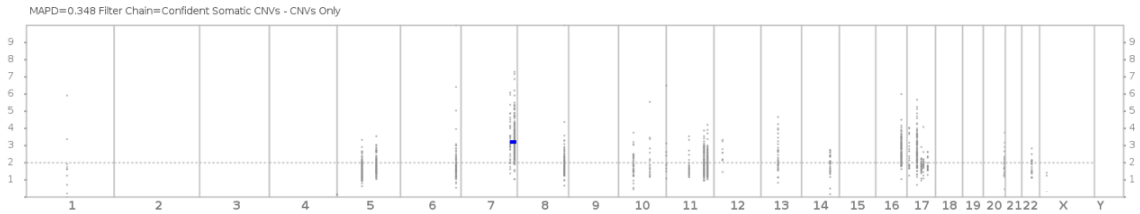

## Supplementary Figure 2. CNV identified in S01

The x axis represents the chromosome number; y axis represents the ploidy for each chromosome. The dashed horizontal line represents the normal ploidy condition. The blue spots indicate a copy number gain on chromosome 7, corresponding to *BRAF* gene, and on chromosome 16, corresponding to *ZFHX3* gene. The red spot indicates copy number loss on chromosome X, corresponding to *EIF1AX* gene.



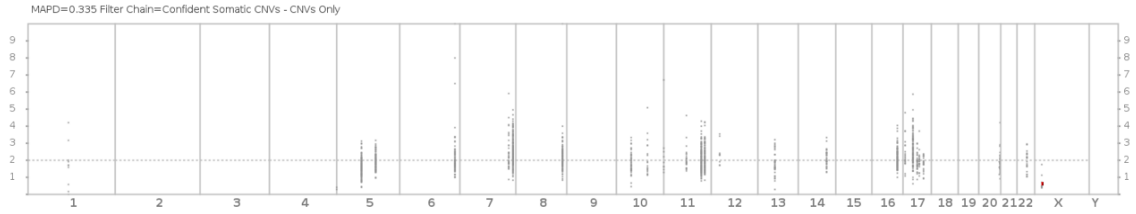

## Supplementary Figure 3. CNV identified in S07

The x axis represents the chromosome number; y axis represents the ploidy for each chromosome. The dashed horizontal line represents the normal ploidy condition. The blue spot indicates copy number gain on chromosome 17, corresponding to *TP53* gene



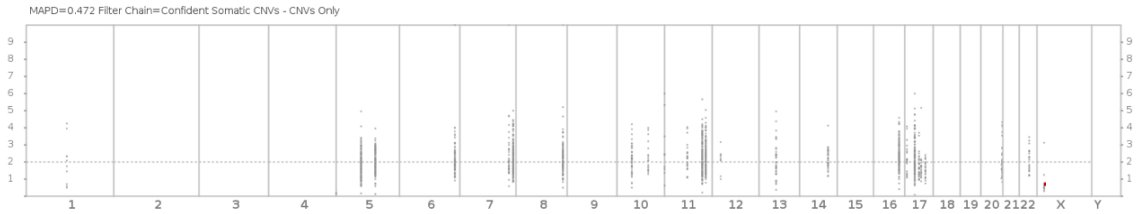

## Supplementary Figure 4. CNV identified in A02

The x axis represents the chromosome number; y axis represents the ploidy for each chromosome. The dashed horizontal line represents the normal ploidy condition. The blue spot indicates a copy number gain on chromosome 7, corresponding to *BRAF* and *KTM2C* genes.



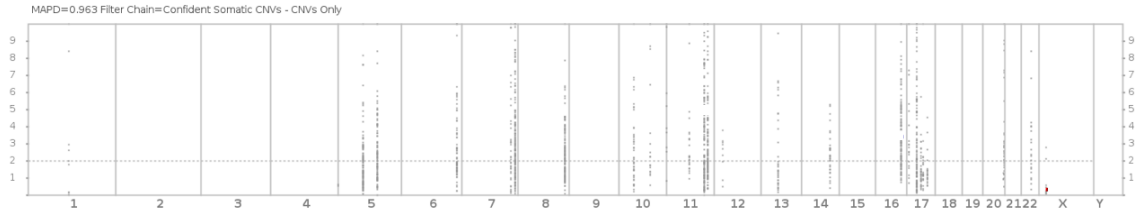

## Supplementary Figure 5. CNV identified in A07

The x axis represents the chromosome number; y axis represents the ploidy for each chromosome. The dashed horizontal line represents the normal ploidy condition. The red spot indicates a copy number loss on chromosome X, corresponding to *EIF1AX* gene.



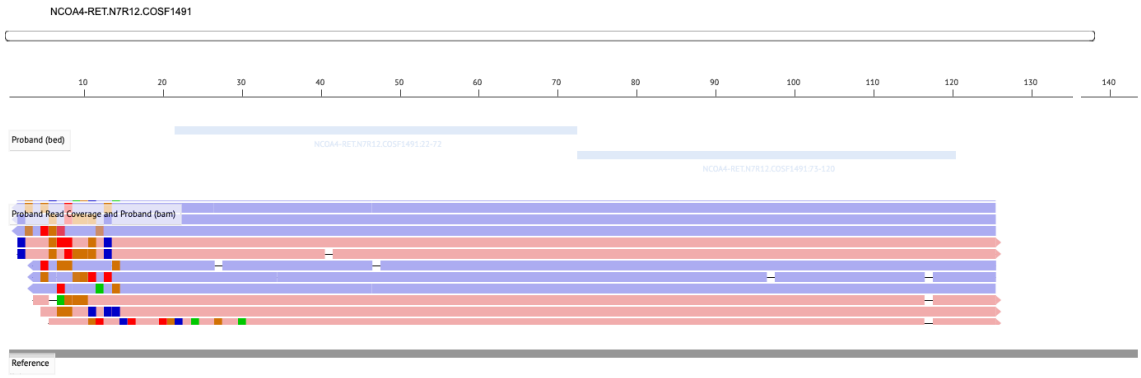

## Supplementary Figure 6. CNV identified in A09

The x axis represents the chromosome number; y axis represents the ploidy for each chromosome. The dashed horizontal line represents the normal ploidy condition. The red spot indicates a copy number loss on chromosome X, corresponding to *EIF1AX* gene.



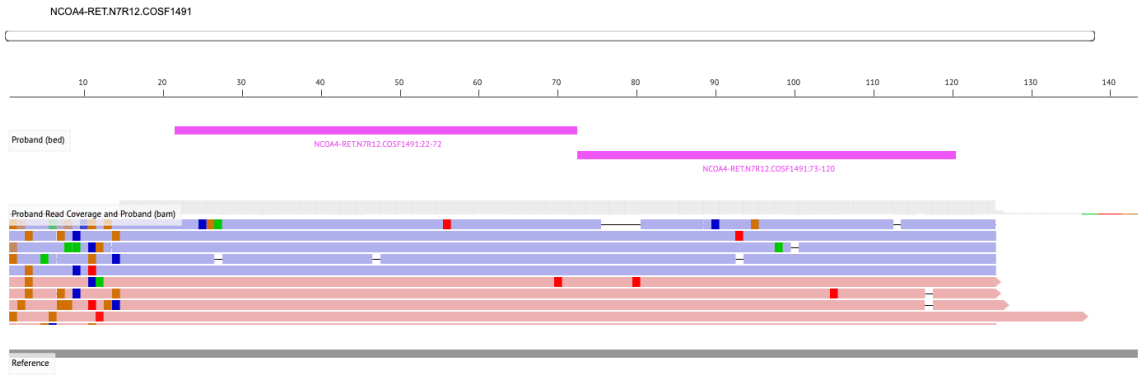

## Supplementary Figure 7. CNV identified in A11

The x axis represents the chromosome number; y axis represents the ploidy for each chromosome. The dashed horizontal line represents the normal ploidy condition. The red spot indicates a copy number loss on chromosome X, corresponding to *EIF1AX* gene.



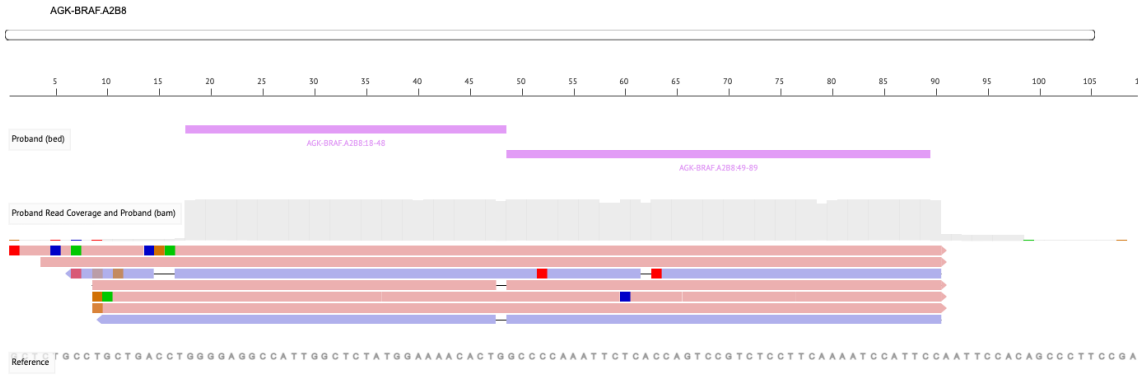

## Supplementary Figure 8. Fusion identified in P09

The IGV shows the fusion regions between *RET* and *NCOA4* (also known as PTC3). The blue and red lines indicate the reads in both forward and reverse direction, covering the fusion regions. The numbered scale represents the base pairs of the amplicon.




## Supplementary Figure 9. Fusion identified in P15

The IGV shows the fusion regions between *RET* and *NCOA4* (also known as PTC3). The blue and red lines indicate the reads in both forward and reverse direction, covering the fusion regions. The numbered scale represents the base pairs of the amplicon.




## Supplementary Figure 10. Fusion identified in S01

The IGV shows the fusion regions between *RET* and *NCOA4* (PTC2). The blue and red lines indicate the reads in both forward and reverse direction, covering the fusion regions. The numbered scale represents the base pairs of the amplicon.

## Supplementary Table 2. Top 100 Enriched Biological Processes in De Novo Papillary Thyroid Cancers According to Gene Ontology Analysis

| # | Term (Gene Ontology ID) | p-value | q-value |
|---|---|---|---|
| 1 | MAPK cascade (GO:0000165) | 1.7395487929341005e-09 | 6.853287586558382e-07 |
| 2 | response to gamma radiation (GO:0010332) | 2.873495843420705e-09 | 6.853287586558382e-07 |
| 3 | small GTPase mediated signal transduction (GO:0007264) | 1.5852753307949112e-08 | 2.520587775963909e-06 |
| 4 | Ras protein signal transduction (GO:0007265) | 6.894202313846511e-08 | 8.221336259261964e-06 |
| 5 | cellular response to gamma radiation (GO:0071480) | 2.8366829343744953e-07 | 2.7061955193932686e-05 |
| 6 | positive regulation of transcription, DNA-templated (GO:0045893) | 1.2288524430338584e-06 | 9.769376922119175e-05 |
| 7 | histone H3-K4 methylation (GO:0051568) | 1.683001407154634e-06 | 0.00011468452445896577 |
| 8 | regulation of cell cycle (GO:0051726) | 2.6231602249832543e-06 | 0.00015640592841462654 |
| 9 | positive regulation of transcription by RNA polymerase II (GO:0045944) | 3.0996875449344348e-06 | 0.00016428343988152504 |
| 10 | positive regulation of cellular process (GO:0048522) | 5.571151643809081e-06 | 0.00026574393340969316 |
| 11 | regulation of telomerase activity (GO:0051972) | 6.243776171906146e-06 | 0.00027075283945447563 |
| 12 | cellular response to ionizing radiation (GO:0071479) | 7.107193251517127e-06 | 0.00028252109317478058 |
| 13 | negative regulation of cell population proliferation (GO:0008285) | 8.752228882663223e-06 | 0.00032113947515618134 |
| 14 | regulation of cell migration (GO:0030334) | 1.250846752810919e-05 | 0.0004261813579220059 |
| 15 | histone H3-K4 monomethylation (GO:0097692) | 1.675288186561917e-05 | 0.0005236292507734916 |
| 16 | regulation of cell population proliferation (GO:0042127) | 1.7564083883387558e-05 | 0.0005236292507734916 |
| 17 | positive regulation of cell motility (GO:2000147) | 2.3790975884101065e-05 | 0.0006418835609547428 |
| 18 | response to ionizing radiation (GO:0010212) | 2.4222021168103504e-05 | 0.0006418835609547428 |
| 19 | negative regulation of telomerase activity (GO:0051974) | 2.6899281719267998e-05 | 0.0006753135463205702 |
| 20 | negative regulation of protein serine/threonine kinase activity (GO:0071901) | 3.080704022356698e-05 | 0.0007347479093320725 |
| 21 | positive regulation of cell adhesion (GO:0045785) | 3.323833204244041e-05 | 0.0007549849706782893 |
| 22 | replicative senescence (GO:0090399) | 3.941561644068303e-05 | 0.0008546022291911731 |
| 23 | protein-containing complex assembly (GO:0065003) | 4.9808109411792704e-05 | 0.0010191195652287746 |
| 24 | positive regulation of cell migration (GO:0030335) | 5.12764561121396e-05 | 0.0010191195652287746 |
| 25 | positive regulation of programmed cell death (GO:0043068) | 6.507117586083098e-05 | 0.001241558035424655 |
| 26 | positive regulation of apoptotic process (GO:0043065) | 8.892973141241418e-05 | 0.0015522561767788033 |
| 27 | protein-containing complex subunit organization (GO:0043933) | 9.077851676846927e-05 | 0.0015522561767788033 |
| 28 | negative regulation of fibroblast proliferation (GO:0048147) | 9.11177629974979e-05 | 0.0015522561767788033 |
| 29 | cellular response to lectin (GO:1990858) | 9.819625472168796e-05 | 0.0015613204500748387 |
| 30 | stimulatory C-type lectin receptor signaling pathway (GO:0002223) | 9.819625472168796e-05 | 0.0015613204500748387 |
| 31 | innate immune response activating cell surface receptor signaling pathway (GO:0002220) | 0.00010868787723630718 | 0.00165477889494936 |
| 32 | negative regulation of cyclin-dependent protein serine/threonine kinase activity (GO:0045736) | 0.00011304743026189121 | 0.00165477889494936 |
| 33 | regulation of gene expression (GO:0010468) | 0.00012194035140767226 | 0.00165477889494936 |
| 34 | spindle assembly checkpoint signaling (GO:0071173) | 0.000124888732037253 | 0.00165477889494936 |
| 35 | mitotic spindle assembly checkpoint signaling (GO:0007094) | 0.000124888732037253 | 0.00165477889494936 |
| 36 | mitotic spindle checkpoint signaling (GO:0071174) | 0.000124888732037253 | 0.00165477889494936 |
| 37 | negative regulation of cyclin-dependent protein kinase activity (GO:1904030) | 0.00013731388963879662 | 0.0017236506673080523 |

| | | | |
|---|---|---|---|
| **38** | **negative regulation of mitotic metaphase/anaphase transition (GO:0045841)** | 0.00013731388963879662 | 0.0017236506673080523 |
| **39** | **cellular response to DNA damage stimulus (GO:0006974)** | 0.00014205024123087765 | 0.001737383719669965 |
| **40** | **regulation of cell adhesion (GO:0030155)** | 0.00015111116263504502 | 0.001802000614422912 |
| **41** | **endocrine system development (GO:0035270)** | 0.00016391054960449903 | 0.00190695932100844 |
| **42** | **positive regulation of protein phosphorylation (GO:0001934)** | 0.00017769606897435948 | 0.0020181196404945112 |
| **43** | **positive regulation of macromolecule metabolic process (GO:0010604)** | 0.00020277492704402623 | 0.0022493869813953607 |
| **44** | **regulation of apoptotic process (GO:0042981)** | 0.00021533597833675084 | 0.0023344377651506854 |
| **45** | **negative regulation of transferase activity (GO:0051348)** | 0.00024055368393234716 | 0.0025087802338273034 |
| **46** | **regulation of cell differentiation (GO:0045595)** | 0.0002419368778952955 | 0.0025087802338273034 |
| **47** | **regulation of MAPK cascade (GO:0043408)** | 0.0002904298199196884 | 0.0029163363241536447 |
| **48** | **negative regulation of DNA biosynthetic process (GO:2000279)** | 0.00029346780620414034 | 0.0029163363241536447 |
| **49** | **histone lysine methylation (GO:0034968)** | 0.00033161703785482386 | 0.003228190348096959 |
| **50** | **positive regulation of gene expression (GO:0010628)** | 0.00048155431370325503 | 0.0045940281527290525 |
| **51** | **cellular response to retinoic acid (GO:0071300)** | 0.0005315449008354333 | 0.0049715081901667 |
| **52** | **positive regulation of RNA metabolic process (GO:0051254)** | 0.0005822138818500083 | 0.005340692723893346 |
| **53** | **positive regulation of nucleic acid-templated transcription (GO:1903508)** | 0.0006002599226353449 | 0.00537415154663229 |
| **54** | **regulation of fibroblast proliferation (GO:0048145)** | 0.0006083945147130895 | 0.00537415154663229 |
| **55** | **intrinsic apoptotic signaling pathway in response to DNA damage (GO:0008630)** | 0.0007477269425466376 | 0.0064848318471772025 |
| **56** | **signal transduction in response to DNA damage (GO:0042770)** | 0.0007772736651313045 | 0.006620706040493433 |
| **57** | **regulation of megakaryocyte differentiation (GO:0045652)** | 0.0008380409440356887 | 0.0070130794790355 |
| **58** | **negative regulation of cellular process (GO:0048523)** | 0.0008806880558634352 | 0.007242900045635493 |
| **58** | **DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest (GO:0006977)** | 0.0009010346774456689 | 0.0072521773583253 |
| **60** | **regulation of transcription by RNA polymerase II (GO:0006357)** | 0.0009122235670849434 | 0.0072521773583253 |
| **61** | **positive regulation of phosphorylation (GO:0042327)** | 0.0009916339043174551 | 0.007691132102921517 |
| **62** | **mitotic DNA damage checkpoint signaling (GO:0044773)** | 0.0009996859337130693 | 0.007691132102921517 |
| **63** | **cellular component assembly (GO:0022607)** | 0.0010848587453735211 | 0.00821393050068523 |
| **64** | **response to retinoic acid (GO:0032526)** | 0.001103310897929822 | 0.008223114036133204 |
| **65** | **regulation of protein phosphorylation (GO:0001932)** | 0.0011458567755711943 | 0.008408825876114764 |
| **66** | **mitotic G1 DNA damage checkpoint signaling (GO:0031571)** | 0.0012118881637081286 | 0.00875846274072385 |
| **67** | **positive regulation of MAPK cascade (GO:0043410)** | 0.001247911716894404 | 0.008884386402367623 |
| **68** | **regulation of myeloid cell differentiation (GO:0045637)** | 0.0013253963875063156 | 0.009297265835889891 |
| **69** | **gland development (GO:0048732)** | 0.001443814289271468 | 0.009981150956267974 |
| **70** | **DNA damage response, signal transduction by p53 class mediator (GO:0030330)** | 0.001567120651291163 | 0.010669258588383294 |
| **71** | **DNA repair (GO:0006281)** | 0.0015880867081241382 | 0.010669258588383294 |
| **72** | **negative regulation of cell cycle (GO:0045786)** | 0.0018283141997896284 | 0.012112581573606288 |
| **73** | **regulation of cyclin-dependent protein serine/threonine kinase activity (GO:0000079)** | 0.001919676073191219 | 0.012543636807016596 |
| **74** | **positive regulation of plasma membrane bounded cell projection assembly (GO:0120034)** | 0.001966159263820519 | 0.01267375633570794 |
| **75** | **cellular response to oxygen-containing compound (GO:1901701)** | 0.0019992555261276893 | 0.012715265146172104 |
| **76** | **regulation of endothelial cell proliferation (GO:0001936)** | 0.0024084361889752624 | 0.015116106080805266 |

| | | | |
|---|---|---|---|
| **77** | **hemopoiesis (GO:0030097)** | 0.002512535109003552 | 0.015564665545385641 |
| **78** | **positive regulation of cold-induced thermogenesis (GO:0120162)** | 0.002672621458078075 | 0.01613722070257268 |
| **79** | **regulation of MAP kinase activity (GO:0043405)** | 0.002672621458078075 | 0.01613722070257268 |
| **80** | **response to UV (GO:0009411)** | 0.0027819620672849227 | 0.016587448826186353 |
| **81** | **positive regulation of protein localization to membrane (GO:1905477)** | 0.0028933886457841386 | 0.017038844247395483 |
| **82** | **intrinsic apoptotic signaling pathway (GO:0097193)** | 0.0029498822665064593 | 0.017159680989311966 |
| **83** | **regulation of JNK cascade (GO:0046328)** | 0.0031224753229706615 | 0.01794482806092778 |
| **84** | **positive regulation of metabolic process (GO:0009893)** | 0.0036054127233747734 | 0.020473593679163894 |
| **85** | **phosphorylation (GO:0016310)** | 0.003663338969162722 | 0.020485075043104727 |
| **86** | **transmembrane receptor protein tyrosine kinase signaling pathway (GO:0007169)** | 0.003767249707723238 | 0.020485075043104727 |
| **87** | **positive regulation of protein-containing complex assembly (GO:0031334)** | 0.003923876780538224 | 0.020485075043104727 |
| **88** | **glial cell-derived neurotrophic factor receptor signaling pathway (GO:0035860)** | 0.003993945448655638 | 0.020485075043104727 |
| **89** | **beta-catenin destruction complex assembly (GO:1904885)** | 0.003993945448655638 | 0.020485075043104727 |
| **90** | **bone marrow development (GO:0048539)** | 0.003993945448655638 | 0.020485075043104727 |
| **91** | **cellular response to antibiotic (GO:0071236)** | 0.003993945448655638 | 0.020485075043104727 |
| **92** | **Schwann cell differentiation (GO:0014037)** | 0.003993945448655638 | 0.020485075043104727 |
| **93** | **regulation of miRNA metabolic process (GO:2000628)** | 0.003993945448655638 | 0.020485075043104727 |
| **94** | **positive regulation of epithelial cell proliferation (GO:0050679)** | 0.00425501903988099 | 0.021591958319396084 |
| **95** | **regulation of thymocyte apoptotic process (GO:0070243)** | 0.004790942275014269 | 0.023559582115276353 |
| **96** | **astrocyte development (GO:0014002)** | 0.004790942275014269 | 0.023559582115276353 |
| **97** | **cellular response to X-ray (GO:0071481)** | 0.004790942275014269 | 0.023559582115276353 |
| **98** | **pexophagy (GO:0000425)** | 0.005587341707821481 | 0.025143037685196663 |
| **99** | **positive regulation of RNA polymerase II transcription preinitiation complex assembly (GO:0045899)** | 0.005587341707821481 | 0.025143037685196663 |
| **100** | **positive regulation of protein localization to centrosome (GO:1904781)** | 0.005587341707821481 | 0.025143037685196663 |

## Supplementary Table 3. Top 100 Enriched Biological Processes in Secondary Papillary Thyroid Cancers According to Gene Ontology Analysis

| # | Term (Gene Ontology ID) | p-value | q-value |
|---|---|---|---|
| **1** | **response to gamma radiation (GO:0010332)** | 1.667021133489588e-08 | 1.0277037422197609e-05 |
| **2** | **negative regulation of protein serine/threonine kinase activity (GO:0071901)** | 3.17682764210127e-08 | 1.0277037422197609e-05 |
| **3** | **small GTPase mediated signal transduction (GO:0007264)** | 1.4915175157189225e-07 | 3.216706108900476e-05 |
| **4** | **negative regulation of cell population proliferation (GO:0008285)** | 2.1762177729142788e-07 | 3.520032247688846e-05 |
| **5** | **replicative senescence (GO:0090399)** | 3.3161472134701576e-07 | 4.291094494230384e-05 |
| **6** | **positive regulation of transcription, DNA-templated (GO:0045893)** | 4.614368874307938e-07 | 4.9758277694620595e-05 |
| **7** | **Ras protein signal transduction (GO:0007265)** | 6.408939107990364e-07 | 5.923690861242522e-05 |
| **8** | **cellular response to gamma radiation (GO:0071480)** | 1.0209769829080983e-06 | 7.7231362910504e-05 |
| **9** | **regulation of cell cycle (GO:0051726)** | 1.0743157128199938e-06 | 7.7231362910504e-05 |
| **10** | **MAPK cascade (GO:0000165)** | 1.2307995938249777e-06 | 7.963273372047606e-05 |

| | | | |
|---|---|---|---|
| **11** | **negative regulation of cyclin-dependent protein serine/threonine kinase activity (GO:0045736)** | 1.7076087361357007e-06 | 0.0001004384411163453 |
| **12** | **regulation of cell population proliferation (GO:0042127)** | 1.8655911907378508e-06 | 0.00010058645836728245 |
| **13** | **negative regulation of cyclin-dependent protein kinase activity (GO:1904030)** | 2.303146781996938e-06 | 0.00010268397198568423 |
| **14** | **negative regulation of cell cycle (GO:0045786)** | 2.372566258739054e-06 | 0.00010268397198568423 |
| **15** | **regulation of transcription by RNA polymerase II (GO:0006357)** | 2.38061758853982e-06 | 0.00010268397198568423 |
| **16** | **negative regulation of cellular process (GO:0048523)** | 3.1887847471248214e-06 | 0.00012894648321185997 |
| **17** | **histone H3-K4 methylation (GO:0051568)** | 6.033858309199734e-06 | 0.00022019233479891921 |
| **18** | **positive regulation of cellular process (GO:0048522)** | 6.125907305070396e-06 | 0.00022019233479891921 |
| **19** | **regulation of cell migration (GO:0030334)** | 6.8478867883389276e-06 | 0.00023318856589764665 |
| **20** | **negative regulation of cell cycle G1/S phase transition (GO:1902807)** | 9.688815339849341e-06 | 0.0003134331762441262 |
| **21** | **negative regulation of G1/S transition of mitotic cell cycle (GO:2000134)** | 1.148414698957588e-05 | 0.00035382110010740927 |
| **22** | **regulation of telomerase activity (GO:0051972)** | 2.2277900155129167e-05 | 0.0006551727909258441 |
| **23** | **regulation of gene expression (GO:0010468)** | 2.3677212473519384e-05 | 0.0006555886948067028 |
| **24** | **positive regulation of nucleic acid-templated transcription (GO:1903508)** | 2.4570589079219964e-05 | 0.0006555886948067028 |
| **25** | **cellular response to ionizing radiation (GO:0071479)** | 2.5343390574460125e-05 | 0.0006555886948067028 |
| **26** | **negative regulation of protein phosphorylation (GO:0001933)** | 3.0267327982233596e-05 | 0.0007294900665945269 |
| **27** | **intrinsic apoptotic signaling pathway in response to DNA damage (GO:0008630)** | 3.044239845139448e-05 | 0.0007294900665945269 |
| **28** | **signal transduction in response to DNA damage (GO:0042770)** | 3.2280808026285865e-05 | 0.0007459172426073912 |
| **29** | **DNA damage induced protein phosphorylation (GO:0006975)** | 3.847032196562326e-05 | 0.0008296766103919416 |
| **30** | **histone H3-K4 monomethylation (GO:0097692)** | 3.847032196562326e-05 | 0.0008296766103919416 |
| **31** | **DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest (GO:0006977)** | 4.036253222957771e-05 | 0.0008424051081463476 |
| **32** | **mitotic DNA damage checkpoint signaling (GO:0044773)** | 4.722408723590282e-05 | 0.0009548120138009102 |
| **33** | **cellular response to DNA damage stimulus (GO:0006974)** | 5.155369460895416e-05 | 0.0009933074514399073 |
| **34** | **regulation of DNA-templated transcription, initiation (GO:2000142)** | 5.219853686082974e-05 | 0.0009933074514399073 |
| **35** | **regulation of transcription initiation from RNA polymerase II promoter (GO:0060260)** | 6.028541151003668e-05 | 0.0011043194367948595 |
| **36** | **negative regulation of telomerase activity (GO:0051974)** | 6.173695322587491e-05 | 0.0011043194367948595 |
| **37** | **mitotic G1 DNA damage checkpoint signaling (GO:0031571)** | 6.315273440712488e-05 | 0.0011043194367948595 |
| **38** | **positive regulation of transcription by RNA polymerase II (GO:0045944)** | 6.81505787688593e-05 | 0.0011603532753539992 |
| **39** | **regulation of G1/S transition of mitotic cell cycle (GO:2000045)** | 8.2245729511214e-05 | 0.0013644355639424479 |
| **40** | **response to ionizing radiation (GO:0010212)** | 8.575412779016238e-05 | 0.0013870730170058767 |
| **41** | **DNA damage response, signal transduction by p53 class mediator (GO:0030330)** | 9.306144097509942e-05 | 0.001468554934411935 |
| **42** | **positive regulation of cell adhesion (GO:0045785)** | 0.00011739350481991087 | 0.001808418990916246 |
| **43** | **regulation of cyclin-dependent protein serine/threonine kinase activity (GO:0000079)** | 0.00012633911432916592 | 0.001900962952813264 |
| **44** | **positive regulation of cell motility (GO:2000147)** | 0.00012959176509031649 | 0.0019055880003053356 |
| **45** | **cellular protein modification process (GO:0006464)** | 0.00014591475126229418 | 0.002097929868148985 |
| **46** | **negative regulation of mitotic cell cycle phase transition (GO:1901991)** | 0.0001777476099679932 | 0.0025000587749846002 |
| **47** | **negative regulation of fibroblast proliferation (GO:0048147)** | 0.00020868016842799032 | 0.0028726823185725473 |

| | | | |
|---|---|---|---|
| **48** | **intrinsic apoptotic signaling pathway (GO:0097193)** | 0.00024112207109460744 | 0.0032501245832960626 |
| **49** | **chromatin remodeling (GO:0006338)** | 0.000248160437794501 | 0.003276730678633513 |
| **50** | **regulation of protein phosphorylation (GO:0001932)** | 0.00026356075792133857 | 0.0033619921153435765 |
| **51** | **protein-containing complex assembly (GO:0065003)** | 0.00026735482249979416 | 0.0033619921153435765 |
| **52** | **positive regulation of cell migration (GO:0030335)** | 0.0002750608696568323 | 0.0033619921153435765 |
| **53** | **spindle assembly checkpoint signaling (GO:0071173)** | 0.0002857953111961309 | 0.0033619921153435765 |
| **54** | **mitotic spindle assembly checkpoint signaling (GO:0007094)** | 0.0002857953111961309 | 0.0033619921153435765 |
| **55** | **mitotic spindle checkpoint signaling (GO:0071174)** | 0.0002857953111961309 | 0.0033619921153435765 |
| **56** | **negative regulation of mitotic metaphase/anaphase transition (GO:0045841)** | 0.000314144742212557 | 0.003604643014965175 |
| **57** | **protein-containing complex subunit organization (GO:0043933)** | 0.00031756514969554097 | 0.003604643014965175 |
| **58** | **positive regulation of programmed cell death (GO:0043068)** | 0.00034717147767384006 | 0.003872757690603009 |
| **58** | **endocrine system development (GO:0035270)** | 0.0003747925243021608 | 0.004110012935991492 |
| **60** | **DNA repair (GO:0006281)** | 0.0004056348658293148 | 0.004374095969859444 |
| **61** | **positive regulation of apoptotic process (GO:0043065)** | 0.00047084533110052786 | 0.004994048020033467 |
| **62** | **regulation of cell adhesion (GO:0030155)** | 0.0005253190192256105 | 0.0054819581522414515 |
| **63** | **negative regulation of transferase activity (GO:0051348)** | 0.0005493106201692561 | 0.005641332876976329 |
| **64** | **negative regulation of DNA biosynthetic process (GO:2000279)** | 0.0006696067158070427 | 0.006665162232725486 |
| **65** | **DNA damage checkpoint signaling (GO:0000077)** | 0.0006696067158070427 | 0.006665162232725486 |
| **66** | **brain development (GO:0007420)** | 0.000745521638353832 | 0.007195492254134845 |
| **67** | **histone lysine methylation (GO:0034968)** | 0.000756249572304744 | 0.007195492254134845 |
| **68** | **phosphatidylinositol 3-kinase signaling (GO:0014065)** | 0.000756249572304744 | 0.007195492254134845 |
| **69** | **negative regulation of neurogenesis (GO:0050768)** | 0.0008014959660444926 | 0.007515476667112851 |
| **70** | **regulation of signal transduction by p53 class mediator (GO:1901796)** | 0.0008353133639175971 | 0.007611940090911061 |
| **71** | **regulation of cell differentiation (GO:0045595)** | 0.0008353133639175971 | 0.007611940090911061 |
| **72** | **positive regulation of protein phosphorylation (GO:0001934)** | 0.0009227137591005459 | 0.008291608363028516 |
| **73** | **DNA integrity checkpoint signaling (GO:0031570)** | 0.0009952474592520714 | 0.008741116293464466 |
| **74** | **regulation of MAPK cascade (GO:0043408)** | 0.0009997567321736792 | 0.008741116293464466 |
| **75** | **positive regulation of macromolecule metabolic process (GO:0010604)** | 0.0010485896500580098 | 0.009045833381167098 |
| **76** | **negative regulation of transcription by RNA polymerase II (GO:0000122)** | 0.0011410261158100504 | 0.009713735485909246 |
| **77** | **cellular response to retinoic acid (GO:0071300)** | 0.0012092868136118387 | 0.01011643893849145 |
| **78** | **phosphorylation (GO:0016310)** | 0.001219601603094796 | 0.01011643893849145 |
| **79** | **positive regulation of RNA metabolic process (GO:0051254)** | 0.0013238573520127186 | 0.010842224136104162 |
| **80** | **regulation of fibroblast proliferation (GO:0048145)** | 0.0013830203898899714 | 0.011185177403235144 |
| **81** | **negative regulation of cell differentiation (GO:0045596)** | 0.0014968368414389997 | 0.011875510185287637 |
| **82** | **regulation of protein metabolic process (GO:0051246)** | 0.0015050878441941053 | 0.011875510185287637 |
| **83** | **regulation of nucleic acid-templated transcription (GO:1903506)** | 0.0015916712196261307 | 0.012407364808410922 |
| **84** | **regulation of apoptotic process (GO:0042981)** | 0.0016380304979186093 | 0.012616734906587384 |
| **85** | **regulation of intracellular signal transduction (GO:1902531)** | 0.001688741052173828 | 0.012854299538311372 |
| **86** | **regulation of megakaryocyte differentiation (GO:0045652)** | 0.001901018240228249 | 0.014301846528228803 |
| **87** | **regulation of cellular macromolecule biosynthetic process (GO:2000112)** | 0.002168762461272382 | 0.016128612786703807 |
| **88** | **positive regulation of cell population proliferation (GO:0008284)** | 0.002271544162353265 | 0.01670101219366548 |

| | | | |
|---|---|---|---|
| **89** | **positive regulation of gene expression (GO:0010628)** | 0.0024137862334011554 | 0.017547412281017388 |
| **90** | **negative regulation of kinase activity (GO:0033673)** | 0.0024974544715199065 | 0.017756626846960215 |
| **91** | **response to retinoic acid (GO:0032526)** | 0.0024974544715199065 | 0.017756626846960215 |
| **92** | **protein phosphorylation (GO:0006468)** | 0.0026773798758388226 | 0.018828964996388243 |
| **93** | **regulation of ERK1 and ERK2 cascade (GO:0070372)** | 0.0027998197707377885 | 0.01947831603943386 |
| **94** | **canonical Wnt signaling pathway (GO:0060070)** | 0.0029954011693546405 | 0.020400258490236342 |
| **95** | **regulation of myeloid cell differentiation (GO:0045637)** | 0.0029954011693546405 | 0.020400258490236342 |
| **96** | **gland development (GO:0048732)** | 0.003260434667143527 | 0.0217474353571326 |
| **97** | **phosphatidylinositol-mediated signaling (GO:0048015)** | 0.003260434667143527 | 0.0217474353571326 |
| **98** | **positive regulation of phosphorylation (GO:0042327)** | 0.003326390002095202 | 0.021960962564853016 |
| **99** | **chromatin assembly (GO:0031497)** | 0.0034430218247567387 | 0.022501364854723333 |
| **100** | **regulation of transcription, DNA-templated (GO:0006355)** | 0.003486783729186308 | 0.02255949072783541 |

## Supplementary Table 4. Top 100 Enriched Biological Processes in Thyroid Adenomas According to Gene Ontology Analysis

| # | Term (Gene Ontology ID) | p-value | q-value |
|---|---|---|---|
| **1** | **positive regulation of cellular process (GO:0048522)** | 7.979924444029842e-08 | 3.1839898531679065e-05 |
| **2** | **regulation of cell migration (GO:0030334)** | 5.929673566571532e-06 | 0.0011829698765310206 |
| **3** | **histone H3-K4 monomethylation (GO:0097692)** | 1.270931618352025e-05 | 0.0013281066431612809 |
| **4** | **positive regulation of cell motility (GO:2000147)** | 1.3314352312393792e-05 | 0.0013281066431612809 |
| **5** | **positive regulation of cell adhesion (GO:0045785)** | 2.1729931620316324e-05 | 0.0017340485433012425 |
| **6** | **positive regulation of cell migration (GO:0030335)** | 2.8806883666937314e-05 | 0.0019156577638513314 |
| **7** | **MAPK cascade (GO:0000165)** | 4.585699206328194e-05 | 0.0026001924264631646 |
| **8** | **cellular response to gamma radiation (GO:0071480)** | 6.15100567126949e-05 | 0.0026001924264631646 |
| **9** | **positive regulation of transcription, DNA-templated (GO:0045893)** | 8.411821210449226e-05 | 0.0026001924264631646 |
| **10** | **negative regulation of cyclin-dependent protein serine/threonine kinase activity (GO:0045736)** | 8.583035665856277e-05 | 0.0026001924264631646 |
| **11** | **mitotic spindle checkpoint signaling (GO:0071174)** | 9.482727286527288e-05 | 0.0026001924264631646 |
| **12** | **mitotic spindle assembly checkpoint signaling (GO:0007094)** | 9.482727286527288e-05 | 0.0026001924264631646 |
| **13** | **spindle assembly checkpoint signaling (GO:0071173)** | 9.482727286527288e-05 | 0.0026001924264631646 |
| **14** | **regulation of cell adhesion (GO:0030155)** | 9.918353818155463e-05 | 0.0026001924264631646 |
| **15** | **negative regulation of cyclin-dependent protein kinase activity (GO:1904030)** | 0.00010426836797847276 | 0.0026001924264631646 |
| **16** | **negative regulation of mitotic metaphase/anaphase transition (GO:0045841)** | 0.00010426836797847276 | 0.0026001924264631646 |
| **17** | **regulation of cell population proliferation (GO:0042127)** | 0.00012053331852146332 | 0.0027593276780207205 |
| **18** | **response to gamma radiation (GO:0010332)** | 0.000124480947880634 | 0.0027593276780207205 |
| **19** | **histone H3-K4 methylation (GO:0051568)** | 0.00019572323532665868 | 0.004110187941859833 |
| **20** | **positive regulation of cell population proliferation (GO:0008284)** | 0.000258153226805406 | 0.004914003371001311 |
| **21** | **negative regulation of cell cycle G1/S phase transition (GO:1902807)** | 0.00026717919670534254 | 0.004914003371001311 |
| **22** | **positive regulation of transcription by RNA polymerase II (GO:0045944)** | 0.00027094755429079914 | 0.004914003371001311 |

| | | | |
|---|---|---|---|
| **23** | **negative regulation of G1/S transition of mitotic cell cycle (GO:2000134)** | 0.0002988222171700243 | 0.005183915854384335 |
| **24** | **cellular response to retinoic acid (GO:0071300)** | 0.0004041897176269234 | 0.006719654055476015 |
| **25** | **cellular response to ionizing radiation (GO:0071479)** | 0.0005038936977770158 | 0.008042143416521173 |
| **26** | **regulation of ERK1 and ERK2 cascade (GO:0070372)** | 0.0005497022368441449 | 0.008435815096185148 |
| **27** | **regulation of gene expression (GO:0010468)** | 0.0006019695838135628 | 0.008895772738578205 |
| **28** | **regulation of megakaryocyte differentiation (GO:0045652)** | 0.0006377173030824915 | 0.009035756310354895 |
| **29** | **positive regulation of phosphorylation (GO:0042327)** | 0.0006567341679205313 | 0.009035756310354895 |
| **30** | **regulation of protein phosphorylation (GO:0001932)** | 0.0007596086041793947 | 0.010102794435585951 |
| **31** | **positive regulation of MAPK cascade (GO:0043410)** | 0.0008277564804024493 | 0.010474050099352877 |
| **32** | **response to retinoic acid (GO:0032526)** | 0.0008400240681185265 | 0.010474050099352877 |
| **33** | **canonical Wnt signaling pathway (GO:0060070)** | 0.0010095147214072258 | 0.011798621729524695 |
| **34** | **regulation of myeloid cell differentiation (GO:0045637)** | 0.0010095147214072258 | 0.011798621729524695 |
| **35** | **regulation of cell cycle (GO:0051726)** | 0.001034966818379359 | 0.011798621729524695 |
| **36** | **regulation of G1/S transition of mitotic cell cycle (GO:2000045)** | 0.0010999291367171904 | 0.012190881265282194 |
| **37** | **negative regulation of protein serine/threonine kinase activity (GO:0071901)** | 0.0013254974295978832 | 0.013975350290673764 |
| **38** | **cellular response to oxygen-containing compound (GO:1901701)** | 0.00133098574196893 | 0.013975350290673764 |
| **39** | **negative regulation of cell cycle (GO:0045786)** | 0.0013936816342916061 | 0.014258435181598738 |
| **40** | **regulation of cyclin-dependent protein serine/threonine kinase activity (GO:0000079)** | 0.0014635188196625445 | 0.014588374263379927 |
| **41** | **positive regulation of plasma membrane bounded cell projection assembly (GO:0120034)** | 0.001499056002001446 | 0.014588374263379927 |
| **42** | **negative regulation of mitotic cell cycle phase transition (GO:1901991)** | 0.0018373565774843534 | 0.017048959870145514 |
| **43** | **Wnt signaling pathway (GO:0016055)** | 0.0018373565774843534 | 0.017048959870145514 |
| **44** | **positive regulation of protein phosphorylation (GO:0001934)** | 0.001979595558535306 | 0.017860029322279383 |
| **45** | **positive regulation of cold-induced thermogenesis (GO:0120162)** | 0.002039574841630082 | 0.017860029322279383 |
| **46** | **regulation of peptidyl-serine phosphorylation (GO:0033135)** | 0.0020812343433310263 | 0.017860029322279383 |
| **47** | **negative regulation of cell population proliferation (GO:0008285)** | 0.0021038129778123585 | 0.017860029322279383 |
| **48** | **positive regulation of macromolecule metabolic process (GO:0010604)** | 0.002183877397655296 | 0.01815348086800965 |
| **49** | **small GTPase mediated signal transduction (GO:0007264)** | 0.0024290159012020714 | 0.01918994069359022 |
| **50** | **regulation of transcription by RNA polymerase II (GO:0006357)** | 0.0024490031528466003 | 0.01918994069359022 |
| **51** | **phosphorylation (GO:0016310)** | 0.0024528495623385993 | 0.01918994069359022 |
| **52** | **transmembrane receptor protein tyrosine kinase signaling pathway (GO:0007169)** | 0.0025231748983624985 | 0.01936051508551225 |
| **53** | **protein-containing complex subunit organization (GO:0043933)** | 0.002706669841645333 | 0.020312408046847478 |
| **54** | **positive regulation of metabolic process (GO:0009893)** | 0.0027543382287582405 | 0.020312408046847478 |
| **55** | **cellular response to lectin (GO:1990858)** | 0.002850864287276839 | 0.020312408046847478 |
| **56** | **stimulatory C-type lectin receptor signaling pathway (GO:0002223)** | 0.002850864287276839 | 0.020312408046847478 |
| **57** | **innate immune response activating cell surface receptor signaling pathway (GO:0002220)** | 0.003048659628690299 | 0.021340617400832092 |
| **58** | **glial cell-derived neurotrophic factor receptor signaling pathway (GO:0035860)** | 0.0034953988931204698 | 0.022494583199275282 |
| **58** | **beta-catenin destruction complex assembly (GO:1904885)** | 0.0034953988931204698 | 0.022494583199275282 |
| **60** | **negative regulation of organ growth (GO:0046621)** | 0.0034953988931204698 | 0.022494583199275282 |
| **61** | **transcription initiation from RNA polymerase III promoter (GO:0006384)** | 0.0034953988931204698 | 0.022494583199275282 |

| | | | |
|---|---|---|---|
| **62** | **regulation of miRNA metabolic process (GO:2000628)** | 0.0034953988931204698 | 0.022494583199275282 |
| **63** | **positive regulation of gene expression (GO:0010628)** | 0.004153563986844202 | 0.024096760750734022 |
| **64** | **negative regulation of cell size (GO:0045792)** | 0.004193119798694676 | 0.024096760750734022 |
| **65** | **negative regulation of excitatory postsynaptic potential (GO:0090394)** | 0.004193119798694676 | 0.024096760750734022 |
| **66** | **positive regulation of extrinsic apoptotic signaling pathway via death domain receptors (GO:1902043)** | 0.004193119798694676 | 0.024096760750734022 |
| **67** | **cellular response to X-ray (GO:0071481)** | 0.004193119798694676 | 0.024096760750734022 |
| **68** | **postsynaptic density assembly (GO:0097107)** | 0.004193119798694676 | 0.024096760750734022 |
| **69** | **prepulse inhibition (GO:0060134)** | 0.004193119798694676 | 0.024096760750734022 |
| **70** | **Ras protein signal transduction (GO:0007265)** | 0.00430723167559448 | 0.024096760750734022 |
| **71** | **protein phosphorylation (GO:0006468)** | 0.004500687457340992 | 0.024096760750734022 |
| **72** | **cellular protein modification process (GO:0006464)** | 0.0045408809234693745 | 0.024096760750734022 |
| **73** | **brain development (GO:0007420)** | 0.004792708495455789 | 0.024096760750734022 |
| **74** | **pexophagy (GO:0000425)** | 0.004890387533564395 | 0.024096760750734022 |
| **75** | **positive regulation of protein localization to centrosome (GO:1904781)** | 0.004890387533564395 | 0.024096760750734022 |
| **76** | **negative regulation of telomere capping (GO:1904354)** | 0.004890387533564395 | 0.024096760750734022 |
| **77** | **negative regulation of dendritic spine development (GO:0061000)** | 0.004890387533564395 | 0.024096760750734022 |
| **78** | **histone H3-K4 dimethylation (GO:0044648)** | 0.004890387533564395 | 0.024096760750734022 |
| **79** | **positive regulation of miRNA metabolic process (GO:2000630)** | 0.004890387533564395 | 0.024096760750734022 |
| **80** | **negative regulation of cell cycle phase transition (GO:1901988)** | 0.004890387533564395 | 0.024096760750734022 |
| **81** | **positive regulation of nucleic acid-templated transcription (GO:1903508)** | 0.004891823611051268 | 0.024096760750734022 |
| **82** | **regulation of cell differentiation (GO:0045595)** | 0.005172725227483539 | 0.025169723972755274 |
| **83** | **excitatory synapse assembly (GO:1904861)** | 0.0055872022837974 | 0.02562406564638118 |
| **84** | **regulation of wound healing, spreading of epidermal cells (GO:1903689)** | 0.0055872022837974 | 0.02562406564638118 |
| **85** | **postsynaptic density organization (GO:0097106)** | 0.0055872022837974 | 0.02562406564638118 |
| **86** | **DNA damage induced protein phosphorylation (GO:0006975)** | 0.0055872022837974 | 0.02562406564638118 |
| **87** | **peptidyl-serine autophosphorylation (GO:0036289)** | 0.0055872022837974 | 0.02562406564638118 |
| **88** | **regulation of neuron projection development (GO:0010975)** | 0.005768069610737263 | 0.02615295198504736 |
| **89** | **regulation of MAPK cascade (GO:0043408)** | 0.005836081726626109 | 0.026164006841840646 |
| **90** | **positive regulation of ERK1 and ERK2 cascade (GO:0070374)** | 0.006251924576393368 | 0.0277168656220106 |
| **91** | **negative regulation of cellular process (GO:0048523)** | 0.006501163478400495 | 0.02841642855515429 |
| **92** | **regulation of attachment of spindle microtubules to kinetochore (GO:0051988)** | 0.006979473680213335 | 0.02841642855515429 |
| **93** | **negative regulation of phosphatidylinositol 3-kinase signaling (GO:0014067)** | 0.006979473680213335 | 0.02841642855515429 |
| **94** | **regulation of DNA methylation (GO:0044030)** | 0.006979473680213335 | 0.02841642855515429 |
| **95** | **presynaptic membrane assembly (GO:0097105)** | 0.006979473680213335 | 0.02841642855515429 |
| **96** | **regulation of organ growth (GO:0046620)** | 0.006979473680213335 | 0.02841642855515429 |
| **97** | **myelin maintenance (GO:0043217)** | 0.006979473680213335 | 0.02841642855515429 |
| **98** | **regulation of protein localization to centrosome (GO:1904779)** | 0.006979473680213335 | 0.02841642855515429 |
| **99** | **phosphorylated carbohydrate dephosphorylation (GO:0046838)** | 0.007674930787541144 | 0.029731042565329286 |
| **100** | **presynaptic membrane organization (GO:0097090)** | 0.007674930787541144 | 0.029731042565329286 |

## Supplementary Table 5. Top 100 Enriched Pathways in De Novo Papillary Thyroid Cancers According to Reactome Analysis

| # | Term (Reactome ID) | p-value | q-value |
|---|---|---|---|
| 1 | RAS Signaling Downstream Of NF1 Loss-Of-Function Variants R-HSA-6802953 | 8.387593360272853e-09 | 2.5246656014421285e-06 |
| 2 | Oncogenic MAPK Signaling R-HSA-6802957 | 3.758816366507574e-07 | 5.657018631593899e-05 |
| 3 | RAF/MAP Kinase Cascade R-HSA-5673001 | 1.7016183300408428e-06 | 9.449934319120904e-05 |
| 4 | MAPK1/MAPK3 Signaling R-HSA-5684996 | 1.8947926377871997e-06 | 9.449934319120904e-05 |
| 5 | RAF Activation R-HSA-5673000 | 2.0540807489581383e-06 | 9.449934319120904e-05 |
| 6 | Signaling By High-Kinase Activity BRAF Mutants R-HSA-6802948 | 2.258389565450525e-06 | 9.449934319120904e-05 |
| 7 | Signaling To ERKs R-HSA-187687 | 2.258389565450525e-06 | 9.449934319120904e-05 |
| 8 | MAP2K And MAPK Activation R-HSA-5674135 | 3.209962527762303e-06 | 9.449934319120904e-05 |
| 9 | Signaling By RAF1 Mutants R-HSA-9656223 | 3.4834058980573822e-06 | 9.449934319120904e-05 |
| 10 | MAPK Family Signaling Cascades R-HSA-5683057 | 3.7248458900597863e-06 | 9.449934319120904e-05 |
| 11 | Signaling By FGFR3 R-HSA-5654741 | 3.77185285167105e-06 | 9.449934319120904e-05 |
| 12 | Signaling By FGFR4 R-HSA-5654743 | 4.075693778113988e-06 | 9.449934319120904e-05 |
| 13 | Paradoxical Activation Of RAF Signaling By Kinase Inactive BRAF R-HSA-6802955 | 4.395318287963211e-06 | 9.449934319120904e-05 |
| 14 | Negative Regulation Of MAPK Pathway R-HSA-5675221 | 4.395318287963211e-06 | 9.449934319120904e-05 |
| 15 | Estrogen-stimulated Signaling Thru PRKCZ R-HSA-9634635 | 5.9914892032562716e-06 | 0.00012022921667867585 |
| 16 | Signaling By FGFR1 R-HSA-5654736 | 7.567018654927798e-06 | 0.0001423545384458292 |
| 17 | SOS-mediated Signaling R-HSA-112412 | 8.983080444348457e-06 | 0.00015905336551464035 |
| 18 | RUNX3 Regulates CDKN1A Transcription R-HSA-8941855 | 1.2570489476168942e-05 | 0.0001891858666163426 |
| 19 | Activated NTRK3 Signals Thru RAS R-HSA-9034864 | 1.2570489476168942e-05 | 0.0001891858666163426 |
| 20 | Activated NTRK2 Signals Thru RAS R-HSA-9026519 | 1.2570489476168942e-05 | 0.0001891858666163426 |
| 21 | Signaling By BRAF And RAF1 Fusions R-HSA-6802952 | 1.5435164714613006e-05 | 0.00021924423658919 |
| 22 | SHC-related Events Triggered By IGF1R R-HSA-2428933 | 1.675288186561917e-05 | 0.00021924423658919 |
| 23 | EGFR Transactivation By Gastrin R-HSA-2179392 | 1.675288186561917e-05 | 0.00021924423658919 |
| 24 | Regulation Of RAS By GAPs R-HSA-5658442 | 1.7800710274205427e-05 | 0.00022199712133303602 |
| 25 | Generic Transcription Pathway R-HSA-212436 | 1.843829911403954e-05 | 0.00022199712133303602 |
| 26 | Activated NTRK2 Signals Thru FRS2 And FRS3 R-HSA-9028731 | 2.152942370892626e-05 | 0.00024001320505136312 |
| 27 | MET Activates RAS Signaling R-HSA-8851805 | 2.152942370892626e-05 | 0.00024001320505136312 |
| 28 | Signaling By FGFR2 R-HSA-5654738 | 2.4222021168103504e-05 | 0.0002603867275571127 |
| 29 | Signaling By FGFR4 In Disease R-HSA-5655291 | 2.6899281719267998e-05 | 0.0002698894599166556 |
| 30 | Constitutive Signaling By Overexpressed ERBB2 R-HSA-9634285 | 2.6899281719267998e-05 | 0.0002698894599166556 |
| 31 | Signaling By PDGFRA Extracellular Domain Mutants R-HSA-9673770 | 3.286162327694617e-05 | 0.00031907576149550954 |
| 32 | RNA Polymerase II Transcription R-HSA-73857 | 3.477346985884713e-05 | 0.0003270879508597808 |
| 33 | TP53 Regulates Transcription Of Caspase Activators And Caspases R-HSA-6803207 | 3.941561644068303e-05 | 0.0003389743013898741 |
| 34 | GRB2 Events In EGFR Signaling R-HSA-179812 | 3.941561644068303e-05 | 0.0003389743013898741 |
| 35 | p38MAPK Events R-HSA-171007 | 3.941561644068303e-05 | 0.0003389743013898741 |
| 36 | Signaling By FGFR R-HSA-190236 | 4.128131894504711e-05 | 0.000345157694512755 |
| 37 | Erythropoietin Activates RAS R-HSA-9027284 | 4.6560429993752554e-05 | 0.00035036723570298797 |
| 38 | SHC1 Events In ERBB4 Signaling R-HSA-1250347 | 4.6560429993752554e-05 | 0.00035036723570298797 |

| | | | |
|---|---|---|---|
| **39** | **SHC1 Events In EGFR Signaling R-HSA-180336** | 4.6560429993752554e-05 | 0.0003503672357029879 7 |
| **40** | **PTK6 Regulates RHO GTPases, RAS GTPase And MAP Kinases R-HSA-8849471** | 4.6560429993752554e-05 | 0.0003503672357029879 7 |
| **41** | **Constitutive Signaling By EGFRvIII R-HSA-5637810** | 5.429523345740245e-05 | 0.0003983843868179691 |
| **42** | **Transcriptional Regulation By RUNX3 R-HSA-8878159** | 5.558851909087941e-05 | 0.0003983843868179691 |
| **43** | **GRB2 Events In ERBB2 Signaling R-HSA-1963640** | 6.261919708622724e-05 | 0.0004188528516212089 |
| **44** | **Signaling By ERBB2 ECD Mutants R-HSA-9665348** | 6.261919708622724e-05 | 0.0004188528516212089 |
| **45** | **Signaling By FLT3 ITD And TKD Mutants R-HSA-9703648** | 6.261919708622724e-05 | 0.0004188528516212089 |
| **46** | **Gene Expression (Transcription) R-HSA-74160** | 6.593810008933556e-05 | 0.0004314645244976088 |
| **47** | **Signaling By NTRK3 (TRKC) R-HSA-9034015** | 7.15314919064766e-05 | 0.0004581059375287118 3 |
| **48** | **SHC-mediated cascade:FGFR3 R-HSA-5654704** | 8.103128969688274e-05 | 0.0004878083639752340 7 |
| **49** | **Tie2 Signaling R-HSA-210993** | 8.103128969688274e-05 | 0.0004878083639752340 7 |
| **50** | **Gastrin-CREB Signaling Pathway Via PKC And MAPK R-HSA-881907** | 8.103128969688274e-05 | 0.0004878083639752340 7 |
| **51** | **VEGFR2 Mediated Cell Proliferation R-HSA-5218921** | 9.11177629974979e-05 | 0.0004941744457275818 |
| **52** | **Constitutive Signaling By Ligand-Responsive EGFR Cancer Variants R-HSA-1236382** | 9.11177629974979e-05 | 0.0004941744457275818 |
| **53** | **Signaling By FLT3 Fusion Proteins R-HSA-9703465** | 9.11177629974979e-05 | 0.0004941744457275818 |
| **54** | **Signaling By NTRK1 (TRKA) R-HSA-187037** | 9.568128026635027e-05 | 0.0004941744457275818 |
| **55** | **Signaling By PDGFR In Disease R-HSA-9671555** | 0.00010179008516647866 | 0.0004941744457275818 |
| **56** | **Signaling By KIT In Disease R-HSA-9669938** | 0.00010179008516647866 | 0.0004941744457275818 |
| **57** | **SHC-mediated cascade:FGFR4 R-HSA-5654719** | 0.00010179008516647866 | 0.0004941744457275818 |
| **58** | **Regulation Of TP53 Activity Thru Methylation R-HSA-6804760** | 0.00010179008516647866 | 0.0004941744457275818 |
| **58** | **Ras Activation Upon Ca2+ Influx Thru NMDA Receptor R-HSA-442982** | 0.00010179008516647866 | 0.0004941744457275818 |
| **60** | **RAS Processing R-HSA-9648002** | 0.00010179008516647866 | 0.0004941744457275818 |
| **61** | **FRS-mediated FGFR3 Signaling R-HSA-5654706** | 0.00010179008516647866 | 0.0004941744457275818 |
| **62** | **Signaling To RAS R-HSA-167044** | 0.00010179008516647866 | 0.0004941744457275818 |
| **63** | **TP53 Regulates Transcription Of Genes Involved In Cytochrome C Release R-HSA-6803204** | 0.00011304743026189121 | 0.0005234965616742963 |
| **64** | **CD209 (DC-SIGN) Signaling R-HSA-5621575** | 0.00011304743026189121 | 0.0005234965616742963 |
| **65** | **SHC-mediated cascade:FGFR1 R-HSA-5654688** | 0.00011304743026189121 | 0.0005234965616742963 |
| **66** | **Developmental Biology R-HSA-1266738** | 0.00011823408398537705 | 0.0005370225847760187 |
| **67** | **Signaling By FGFR3 In Disease R-HSA-5655332** | 0.0001248889732037253 | 0.0005370225847760187 |
| **68** | **Signaling By ERBB2 TMD/JMD Mutants R-HSA-9665686** | 0.0001248889732037253 | 0.0005370225847760187 |
| **69** | **SHC1 Events In ERBB2 Signaling R-HSA-1250196** | 0.0001248889732037253 | 0.0005370225847760187 |
| **70** | **FRS-mediated FGFR4 Signaling R-HSA-5654712** | 0.0001248889732037253 | 0.0005370225847760187 |
| **71** | **FRS-mediated FGFR1 Signaling R-HSA-5654693** | 0.00013731388963879662 | 0.0005740483441844137 |
| **72** | **SHC-mediated cascade:FGFR2 R-HSA-5654699** | 0.00013731388963879662 | 0.0005740483441844137 |
| **73** | **Signaling By NTRKs R-HSA-166520** | 0.00014777472952853766 | 0.0006093177203847922 |
| **74** | **Signaling By NTRK2 (TRKB) R-HSA-9006115** | 0.00015032135601996194 | 0.0006114422724595749 |
| **75** | **Signaling By Erythropoietin R-HSA-9006335** | 0.00016391054960449903 | 0.0006245199421639773 |
| **76** | **Signaling By ERBB2 KD Mutants R-HSA-9664565** | 0.00016391054960449903 | 0.0006245199421639773 |
| **77** | **FRS-mediated FGFR2 Signaling R-HSA-5654700** | 0.00016391054960449903 | 0.0006245199421639773 |
| **78** | **Downstream Signaling Of Activated FGFR3 R-HSA-5654708** | 0.00016391054960449903 | 0.0006245199421639773 |
| **79** | **Signaling By EGFR In Cancer R-HSA-1643713** | 0.00016391054960449903 | 0.0006245199421639773 |
| **80** | **Signaling By ERBB2 In Cancer R-HSA-1227990** | 0.00017808064838669893 | 0.0006700284395549547 |

| | | | |
|---|---|---|---|
| 81 | Downstream Signaling Of Activated FGFR4 R-HSA-5654716 | 0.0001928308312214584 | 0.0007165688913291232 |
| 82 | Disease R-HSA-1643685 | 0.00020765786041422078 | 0.0007459076617888676 |
| 83 | FLT3 Signaling In Disease R-HSA-9682385 | 0.00020816027770852118 | 0.0007459076617888676 |
| 84 | CREB1 Phosphorylation Thru NMDA Receptor-Mediated Activation Of RAS Signaling R-HSA-442742 | 0.00020816027770852118 | 0.0007459076617888676 |
| 85 | DAP12 Signaling R-HSA-2424491 | 0.00022406816820398208 | 0.0007842385887139373 |
| 86 | Downstream Signal Transduction R-HSA-186763 | 0.00022406816820398208 | 0.0007842385887139373 |
| 87 | Downstream Signaling Of Activated FGFR2 R-HSA-5654696 | 0.00024055368393234716 | 0.0008228029416322329 |
| 88 | FCERI Mediated MAPK Activation R-HSA-2871796 | 0.00024055368393234716 | 0.0008228029416322329 |
| 89 | Downstream Signaling Of Activated FGFR1 R-HSA-5654687 | 0.00025761600687002403 | 0.0008712631243581712 |
| 90 | Signal Transduction R-HSA-162582 | 0.00027033957947392775 | 0.0009041357046850251 |
| 91 | Diseases Of Signal Transduction By Growth Factor Receptors And Second Messengers R-HSA-5663202 | 0.00029606400432625703 | 0.0009792886296945424 |
| 92 | Regulation Of TP53 Degradation R-HSA-6804757 | 0.00037205718620846284 | 0.00120418508654567 |
| 93 | Signaling By FGFR1 In Disease R-HSA-5655302 | 0.00037205718620846284 | 0.00120418508654567 |
| 94 | Regulation Of TP53 Expression And Degradation R-HSA-6806003 | 0.00039313432145597 | 0.001245615060613126 |
| 95 | FLT3 Signaling R-HSA-9607240 | 0.00039313432145597 | 0.001245615060613126 |
| 96 | Signaling By FGFR2 In Disease R-HSA-5655253 | 0.0005070581642096471 | 0.0015734485301763276 |
| 97 | Signaling By SCF-KIT R-HSA-1433557 | 0.0005070581642096471 | 0.0015734485301763276 |
| 98 | DAP12 Interactions R-HSA-2172127 | 0.0005315449008354333 | 0.0016326021954231166 |
| 99 | TP53 Regulates Transcription Of Cell Death Genes R-HSA-5633008 | 0.000556597075446985 | 0.0016922799970660854 |
| 100 | PKMTs Methylate Histone Lysines R-HSA-3214841 | 0.0006351381693978915 | 0.001892837514740251 |

## Supplementary Table 6. Top 100 Enriched Pathways in Secondary Papillary Thyroid Cancers According to Reactome Analysis

| # | Term (Reactome ID) | p-value | q-value |
|---|---|---|---|
| 1 | Gene Expression (Transcription) R-HSA-74160 | 2.3820843431355153e-08 | 6.900994667936107e-06 |
| 2 | Generic Transcription Pathway R-HSA-212436 | 3.8445652746162157e-08 | 6.900994667936107e-06 |
| 3 | RNA Polymerase II Transcription R-HSA-73857 | 1.0463409189310099e-07 | 1.2521212996541085e-05 |
| 4 | Regulation Of TP53 Activity Thru Methylation R-HSA-6804760 | 1.4526083270118207e-06 | 0.0001303715973493109 1 |
| 5 | Regulation Of TP53 Degradation R-HSA-6804757 | 1.0561305190162971e-05 | 0.0006871347948762902 |
| 6 | Regulation Of TP53 Expression And Degradation R-HSA-6806003 | 1.148414698957588e-05 | 0.0006871347948762902 |
| 7 | RAS Signaling Downstream Of NF1 Loss-Of-Function Variants R-HSA-6802953 | 2.063922922489771e-05 | 0.0010584976131054683 |
| 8 | RUNX3 Regulates CDKN1A Transcription R-HSA-8941855 | 2.88738312533813e-05 | 0.0012957131774954857 |
| 9 | Stabilization Of P53 R-HSA-69541 | 4.036253222957771e-05 | 0.0015616062277531565 |
| 10 | Recruitment And ATM-mediated Phosphorylation Of Repair And Signal Proteins At DNA Double Strand Breaks R-HSA-5693565 | 4.96704652216321e-05 | 0.0015616062277531565 |
| 11 | DNA Double Strand Break Response R-HSA-5693606 | 5.219853686082974e-05 | 0.0015616062277531565 |
| 12 | DNA Damage/Telomere Stress Induced Senescence R-HSA-2559586 | 5.219853686082974e-05 | 0.0015616062277531565 |
| 13 | p53-Dependent G1 DNA Damage Response R-HSA-69563 | 6.315273440712488e-05 | 0.0017439870501659873 |
| 14 | G1/S DNA Damage Checkpoints R-HSA-69615 | 6.915234345700029e-05 | 0.00177326366436165 |
| 15 | TP53 Regulates Transcription Of Caspase Activators And Caspases R-HSA-6803207 | 9.041513427391057e-05 | 0.0021639355469555933 |

| | | | |
|---|---|---|---|
| **16** | **G2/M DNA Damage Checkpoint R-HSA-69473** | 0.00010476666355344596 | 0.0022991127121672907 |
| **17** | **Oncogenic MAPK Signaling R-HSA-6802957** | 0.00010887163260959316 | 0.0022991127121672907 |
| **18** | **Regulation Of TP53 Activity Thru Phosphorylation R-HSA-6804756** | 0.0001665446137518253 | 0.0033216397964947377 |
| **19** | **Formation Of Senescence-Associated Heterochromatin Foci (SAHF) R-HSA-2559584** | 0.00018562933966781003 | 0.003507417523197042 |
| **20** | **TP53 Regulates Transcription Of Genes Involved In Cytochrome C Release R-HSA-6803204** | 0.00025876610708752327 | 0.004644851622221042 |
| **21** | **RAF/MAP Kinase Cascade R-HSA-5673001** | 0.0002829257501974813 | 0.004836683062899799 |
| **22** | **MAPK1/MAPK3 Signaling R-HSA-5684996** | 0.00030749278017807745 | 0.005017723094724082 |
| **23** | **MAPK Family Signaling Cascades R-HSA-5683057** | 0.000518282211290909 | 0.008089709297975492 |
| **24** | **Disease R-HSA-1643685** | 0.0006567739102443203 | 0.00874985733198134 |
| **25** | **RAF Activation R-HSA-5673000** | 0.0006696067158070427 | 0.00874985733198134 |
| **26** | **Signaling By High-Kinase Activity BRAF Mutants R-HSA-6802948** | 0.0007122855268112139 | 0.00874985733198134 |
| **27** | **Signaling To ERKs R-HSA-187687** | 0.0007122855268112139 | 0.00874985733198134 |
| **28** | **Oncogene Induced Senescence R-HSA-2559585** | 0.0007122855268112139 | 0.00874985733198134 |
| **29** | **G2/M Checkpoints R-HSA-69481** | 0.0007170275110421054 | 0.00874985733198134 |
| **30** | **DNA Double-Strand Break Repair R-HSA-5693532** | 0.0007311858494691927 | 0.00874985733198134 |
| **31** | **Transcriptional Regulation By TP53 R-HSA-3700989** | 0.0007747619076066085 | 0.008972242736476529 |
| **32** | **Regulation Of TP53 Activity R-HSA-5633007** | 0.0008509164987106108 | 0.009546219469909665 |
| **33** | **MAP2K And MAPK Activation R-HSA-5674135** | 0.0008958242762467574 | 0.009656590212743071 |
| **34** | **Signaling By RAF1 Mutants R-HSA-9656223** | 0.0009449004432908564 | 0.009656590212743071 |
| **35** | **Cellular Senescence R-HSA-2559583** | 0.0009824607924292772 | 0.009656590212743071 |
| **36** | **Ovarian Tumor Domain Proteases R-HSA-5689896** | 0.0009952474592520714 | 0.009656590212743071 |
| **37** | **Signaling By FGFR3 R-HSA-5654741** | 0.0009952474592520714 | 0.009656590212743071 |
| **38** | **Signaling By FGFR4 R-HSA-5654743** | 0.0010468624608583276 | 0.009870189738613762 |
| **39** | **Paradoxical Activation Of RAF Signaling By Kinase Inactive BRAF R-HSA-6802955** | 0.0010997425892605864 | 0.009870189738613762 |
| **40** | **Negative Regulation Of MAPK Pathway R-HSA-5675221** | 0.0010997425892605864 | 0.009870189738613762 |
| **41** | **TP53 Regulates Transcription Of Cell Death Genes R-HSA-5633008** | 0.0012659452144418342 | 0.01108473980450289 |
| **42** | **PKMTs Methylate Histone Lysines R-HSA-3214841** | 0.0014434314962566937 | 0.012333596447752553 |
| **43** | **Signal Transduction R-HSA-162582** | 0.001486230786990175 | 0.012333596447752553 |
| **44** | **Diseases Of Signal Transduction By Growth Factor Receptors And Second Messengers R-HSA-5663202** | 0.0015116385618415386 | 0.012333596447752553 |
| **45** | **Signaling By FGFR1 R-HSA-5654736** | 0.0015679866104777358 | 0.01250904873692238 |
| **46** | **Autodegradation Of E3 Ubiquitin Ligase COP1 R-HSA-349425** | 0.0016975001307372168 | 0.013247881455101322 |
| **47** | **Transcriptional Regulation By RUNX1 R-HSA-8878171** | 0.0018070387636645207 | 0.013802700343735382 |
| **48** | **Regulation Of PTEN Gene Transcription R-HSA-8943724** | 0.0024186590963730416 | 0.01793172310551293 |
| **49** | **TP53 Regulates Transcription Of DNA Repair Genes R-HSA-6796648** | 0.0024974544715199065 | 0.01793172310551293 |
| **50** | **Signaling By BRAF And RAF1 Fusions R-HSA-6802952** | 0.0024974544715199065 | 0.01793172310551293 |
| **51** | **Regulation Of RAS By GAPs R-HSA-5658442** | 0.0027410497923989 | 0.01913295326532209 |
| **52** | **Chromatin Modifying Enzymes R-HSA-3247509** | 0.0027998197707377885 | 0.01913295326532209 |
| **53** | **RUNX1 Regulates Genes Involved In Megakaryocyte Differentiation And Platelet Function R-HSA-8936459** | 0.0028246421255210884 | 0.01913295326532209 |
| **54** | **Signaling By FGFR2 R-HSA-5654738** | 0.003351140242328576 | 0.022278876796221456 |
| **55** | **Cell Cycle Checkpoints R-HSA-69620** | 0.004033927749400419 | 0.026330546582450008 |
| **56** | **TP53 Regulates Metabolic Genes R-HSA-5628897** | 0.0042200861822271766 | 0.027053766775349222 |

| | | | |
|---|---|---|---|
| **57** | **Signaling By FGFR R-HSA-190236** | 0.004743270894726786 | 0.029874285108893263 |
| **58** | **Transcriptional Regulation By RUNX3 R-HSA-8878159** | 0.005756502851391889 | 0.03306190610072261 |
| **58** | **DNA Repair R-HSA-73894** | 0.005864436331777521 | 0.03306190610072261 |
| **60** | **Signaling To P38 Via RIT And RIN R-HSA-187706** | 0.0059861389875960165 | 0.03306190610072261 |
| **61** | **Synthesis Of PI R-HSA-1483226** | 0.0059861389875960165 | 0.03306190610072261 |
| **62** | **Activation Of NOXA And Translocation To Mitochondria R-HSA-111448** | 0.0059861389875960165 | 0.03306190610072261 |
| **63** | **SARS-CoV-1 Genome Replication And Transcription R-HSA-9679514** | 0.0059861389875960165 | 0.03306190610072261 |
| **64** | **SARS-CoV-2 Genome Replication And Transcription R-HSA-9694682** | 0.0059861389875960165 | 0.03306190610072261 |
| **65** | **Estrogen-stimulated Signaling Thru PRKCZ R-HSA-9634635** | 0.0059861389875960165 | 0.03306190610072261 |
| **66** | **Cell Cycle R-HSA-1640170** | 0.007149459485429744 | 0.037352882751123154 |
| **67** | **SOS-mediated Signaling R-HSA-112412** | 0.007179244874171303 | 0.037352882751123154 |
| **68** | **Sensing Of DNA Double Strand Breaks R-HSA-5693548** | 0.007179244874171303 | 0.037352882751123154 |
| **69** | **Negative Feedback Regulation Of MAPK Pathway R-HSA-5674499** | 0.007179244874171303 | 0.037352882751123154 |
| **70** | **Developmental Biology R-HSA-1266738** | 0.007969526144727176 | 0.0406105602432474 |
| **71** | **Signaling By NTRK1 (TRKA) R-HSA-187037** | 0.00818976833184501 | 0.0406105602432474 |
| **72** | **ARMS-mediated Activation R-HSA-170984** | 0.00837097899164431 | 0.0406105602432474 |
| **73** | **Activated NTRK2 Signals Thru RAS R-HSA-9026519** | 0.00837097899164431 | 0.0406105602432474 |
| **74** | **Activated NTRK3 Signals Thru RAS R-HSA-9034864** | 0.00837097899164431 | 0.0406105602432474 |
| **75** | **Transcriptional Regulation By RUNX2 R-HSA-8878166** | 0.008894763870200068 | 0.04257626972535766 |
| **76** | **EGFR Transactivation By Gastrin R-HSA-2179392** | 0.009561342717827156 | 0.04400669276538396 |
| **77** | **SHC-related Events Triggered By IGF1R R-HSA-2428933** | 0.009561342717827156 | 0.04400669276538396 |
| **78** | **Gain-of-function MRAS Complexes Activate RAF Signaling R-HSA-9726842** | 0.009561342717827156 | 0.04400669276538396 |
| **79** | **Activated NTRK2 Signals Thru FRS2 And FRS3 R-HSA-9028731** | 0.010750337464808102 | 0.046371643701802225 |
| **80** | **PI5P Regulates TP53 Acetylation R-HSA-6811555** | 0.010750337464808102 | 0.046371643701802225 |
| **81** | **Activation Of PUMA And Translocation To Mitochondria R-HSA-139915** | 0.010750337464808102 | 0.046371643701802225 |
| **82** | **Regulation Of PTEN Localization R-HSA-8948747** | 0.010750337464808102 | 0.046371643701802225 |
| **83** | **MET Activates RAS Signaling R-HSA-8851805** | 0.010750337464808102 | 0.046371643701802225 |
| **84** | **Signaling By NTRKs R-HSA-166520** | 0.010850189612677958 | 0.046371643701802225 |
| **85** | **Constitutive Signaling By Overexpressed ERBB2 R-HSA-9634285** | 0.011937964668319789 | 0.04941372189597947 |
| **86** | **Signaling By FGFR4 In Disease R-HSA-5655291** | 0.011937964668319789 | 0.04941372189597947 |
| **87** | **PTEN Regulation R-HSA-6807070** | 0.011974913105710904 | 0.04941372189597947 |
| **88** | **Signaling By PDGFRA Extracellular Domain Mutants R-HSA-9673770** | 0.013124225781325726 | 0.05177579181863666 |
| **89** | **Regulation Of PTEN mRNA Translation R-HSA-8943723** | 0.013124225781325726 | 0.05177579181863666 |
| **90** | **RHO GTPases Activate IQGAPs R-HSA-5626467** | 0.013124225781325726 | 0.05177579181863666 |
| **91** | **Pexophagy R-HSA-9664873** | 0.013124225781325726 | 0.05177579181863666 |
| **92** | **Frs2-mediated Activation R-HSA-170968** | 0.014309122256811026 | 0.05351015510619957 |
| **93** | **GRB2 Events In EGFR Signaling R-HSA-179812** | 0.014309122256811026 | 0.05351015510619957 |
| **94** | **TP53 Regulates Transcription Of Death Receptors And Ligands R-HSA-6803211** | 0.014309122256811026 | 0.05351015510619957 |
| **95** | **Hormone Ligand-Binding Receptors R-HSA-375281** | 0.014309122256811026 | 0.05351015510619957 |
| **96** | **p38MAPK Events R-HSA-171007** | 0.014309122256811026 | 0.05351015510619957 |
| **97** | **Erythropoietin Activates RAS R-HSA-9027284** | 0.01549265556616628 | 0.05452807204170289 |
| **98** | **SHC1 Events In ERBB4 Signaling R-HSA-1250347** | 0.01549265556616628 | 0.05452807204170289 |

| | | | |
|---|---|---|---|
| 99 | SHC1 Events In EGFR Signaling R-HSA-180336 | 0.01549265556616628 | 0.05452807204170289 |
| 100 | Inhibition Of Replication Initiation Of Damaged DNA By RB1/E2F1 R-HSA-113501 | 0.01549265556616628 | 0.05452807204170289 |

## Supplementary Table 6. Top 100 Enriched Pathways in Secondary Papillary Thyroid Cancers According to Reactome Analysis

| # | Term (Reactome ID) | p-value | q-value |
|---|---|---|---|
| 1 | RAF Activation R-HSA-5673000 | 0.00022299126039333032 | 0.00936695865197297 |
| 2 | Signaling By High-Kinase Activity BRAF Mutants R-HSA-6802948 | 0.00023728298476867357 | 0.00936695865197297 |
| 3 | Signaling To ERKs R-HSA-187687 | 0.00023728298476867357 | 0.00936695865197297 |
| 4 | Gene Expression (Transcription) R-HSA-74160 | 0.0002587167633104477 | 0.00936695865197297 |
| 5 | MAP2K And MAPK Activation R-HSA-5674135 | 0.0002988222171700243 | 0.00936695865197297 |
| 6 | Signaling By RAF1 Mutants R-HSA-9656223 | 0.000315297453838473 | 0.00936695865197297 |
| 7 | Signaling By FGFR3 R-HSA-5654741 | 0.0003322078014627923 | 0.00936695865197297 |
| 8 | Signaling By FGFR4 R-HSA-5654743 | 0.00034955273047183893 | 0.00936695865197297 |
| 9 | Paradoxical Activation Of RAF Signaling By Kinase Inactive BRAF R-HSA-6802955 | 0.0003673317118420773 | 0.00936695865197297 |
| 10 | Negative Regulation Of MAPK Pathway R-HSA-5675221 | 0.0003673317118420773 | 0.00936695865197297 |
| 11 | PKMTs Methylate Histone Lysines R-HSA-3214841 | 0.0004830911270410691 | 0.01115892971018672 2 |
| 12 | Signaling By FGFR1 R-HSA-5654736 | 0.0005251261040087869 | 0.01115892971018672 2 |
| 13 | RAF/MAP Kinase Cascade R-HSA-5673001 | 0.0008017612505235663 | 0.014270019879919035 |
| 14 | Signaling By BRAF And RAF1 Fusions R-HSA-6802952 | 0.0008400240681185265 | 0.014270019879919035 |
| 15 | MAPK1/MAPK3 Signaling R-HSA-5684996 | 0.0008542849371464446 | 0.014270019879919035 |
| 16 | Generic Transcription Pathway R-HSA-212436 | 0.0009415293855927125 | 0.014270019879919035 |
| 17 | RUNX1 Regulates Genes Involved In Megakaryocyte Differentiation And Platelet Function R-HSA-8936459 | 0.000951334658661269 | 0.014270019879919035 |
| 18 | Signaling By FGFR2 R-HSA-5654738 | 0.0011309036960398487 | 0.016021135693897858 |
| 19 | MAPK Family Signaling Cascades R-HSA-5683057 | 0.0012725729948198342 | 0.01690009222737301 |
| 20 | Oncogenic MAPK Signaling R-HSA-6802957 | 0.0013254974295978832 | 0.01690009222737301 |
| 21 | RNA Polymerase II Transcription R-HSA-73857 | 0.0014636283887934716 | 0.017772630435349296 |
| 22 | Signaling By FGFR R-HSA-190236 | 0.0016081357833763062 | 0.01863975567095264 |
| 23 | Signaling By NTRK1 (TRKA) R-HSA-187037 | 0.0028024032097844604 | 0.03074551174579278 |
| 24 | Diseases Of Signal Transduction By Growth Factor Receptors And Second Messengers R-HSA-5663202 | 0.0028936952231334383 | 0.03074551174579278 |
| 25 | Signaling To P38 Via RIT And RIN R-HSA-187706 | 0.0034953988931204698 | 0.03428179683637384 |
| 26 | Estrogen-stimulated Signaling Thru PRKCZ R-HSA-9634635 | 0.0034953988931204698 | 0.03428179683637384 |
| 27 | Signaling By NTRKs R-HSA-166520 | 0.0037348253857727 51 | 0.034491791892488466 |
| 28 | Sensing Of DNA Double Strand Breaks R-HSA-5693548 | 0.004193119798694676 | 0.034491791892488466 |
| 29 | Negative Feedback Regulation Of MAPK Pathway R-HSA-5674499 | 0.004193119798694676 | 0.034491791892488466 |
| 30 | RAS Signaling Downstream Of NF1 Loss-Of-Function Variants R-HSA-6802953 | 0.004193119798694676 | 0.034491791892488466 |
| 31 | SOS-mediated Signaling R-HSA-112412 | 0.004193119798694676 | 0.034491791892488466 |
| 32 | Signal Transduction R-HSA-162582 | 0.004330348739526121 | 0.03450746651809878 |
| 33 | ARMS-mediated Activation R-HSA-170984 | 0.004890387533564395 | 0.034640245029414464 |

| | | | |
|---|---|---|---|
| **34** | **Activated NTRK2 Signals Thru RAS R-HSA-9026519** | 0.004890387533564395 | 0.034640245029414464 |
| **35** | **Activated NTRK3 Signals Thru RAS R-HSA-9034864** | 0.004890387533564395 | 0.034640245029414464 |
| **36** | **RUNX3 Regulates CDKN1A Transcription R-HSA-8941855** | 0.004890387533564395 | 0.034640245029414464 |
| **37** | **Disease R-HSA-1643685** | 0.005033211416489899 | 0.03468834895148444 |
| **38** | **Developmental Biology R-HSA-1266738** | 0.005348183766701789 | 0.03474967274069115 |
| **39** | **Gain-of-function MRAS Complexes Activate RAF Signaling R-HSA-9726842** | 0.0055872022837974 | 0.03474967274069115 |
| **40** | **SHC-related Events Triggered By IGF1R R-HSA-2428933** | 0.0055872022837974 | 0.03474967274069115 |
| **41** | **EGFR Transactivation By Gastrin R-HSA-2179392** | 0.0055872022837974 | 0.03474967274069115 |
| **42** | **Activated NTRK2 Signals Thru FRS2 And FRS3 R-HSA-9028731** | 0.006283564258093123 | 0.03641611104122151 |
| **43** | **Regulation Of PTEN Localization R-HSA-8948747** | 0.006283564258093123 | 0.03641611104122151 |
| **44** | **MET Activates RAS Signaling R-HSA-8851805** | 0.006283564258093123 | 0.03641611104122151 |
| **45** | **Signaling By FGFR4 In Disease R-HSA-5655291** | 0.006979473680213335 | 0.03869056061857393 |
| **46** | **Constitutive Signaling By Overexpressed ERBB2 R-HSA-9634285** | 0.006979473680213335 | 0.03869056061857393 |
| **47** | **Pexophagy R-HSA-9664873** | 0.007674930787541144 | 0.03917702884937202 |
| **48** | **Regulation Of PTEN mRNA Translation R-HSA-8943723** | 0.007674930787541144 | 0.03917702884937202 |
| **49** | **Signaling By PDGFRA Extracellular Domain Mutants R-HSA-9673770** | 0.007674930787541144 | 0.03917702884937202 |
| **50** | **p38MAPK Events R-HSA-171007** | 0.008369935820942226 | 0.03917702884937202 |
| **51** | **Frs2-mediated Activation R-HSA-170968** | 0.008369935820942226 | 0.03917702884937202 |
| **52** | **TP53 Regulates Transcription Of Caspase Activators And Caspases R-HSA-6803207** | 0.008369935820942226 | 0.03917702884937202 |
| **53** | **GRB2 Events In EGFR Signaling R-HSA-179812** | 0.008369935820942226 | 0.03917702884937202 |
| **54** | **Hormone Ligand-Binding Receptors R-HSA-375281** | 0.008369935820942226 | 0.03917702884937202 |
| **55** | **Transcriptional Regulation By RUNX1 R-HSA-8878171** | 0.008690990362893284 | 0.03917702884937202 |
| **56** | **PTK6 Regulates RHO GTPases, RAS GTPase And MAP Kinases R-HSA-8849471** | 0.009064489027893918 | 0.03917702884937202 |
| **57** | **SHC1 Events In EGFR Signaling R-HSA-180336** | 0.009064489027893918 | 0.03917702884937202 |
| **58** | **SHC1 Events In ERBB4 Signaling R-HSA-1250347** | 0.009064489027893918 | 0.03917702884937202 |
| **58** | **Erythropoietin Activates RAS R-HSA-9027284** | 0.009064489027893918 | 0.03917702884937202 |
| **60** | **Constitutive Signaling By EGFRvIII R-HSA-5637810** | 0.009758590659951067 | 0.04079410849651676 |
| **61** | **Prolonged ERK Activation Events R-HSA-169893** | 0.009758590659951067 | 0.04079410849651676 |
| **62** | **Signaling By ERBB2 ECD Mutants R-HSA-9665348** | 0.0104522409709381 | 0.04162586683536024 |
| **63** | **Signaling By FLT3 ITD And TKD Mutants R-HSA-9703648** | 0.0104522409709381 | 0.04162586683536024 |
| **64** | **GRB2 Events In ERBB2 Signaling R-HSA-1963640** | 0.0104522409709381 | 0.04162586683536024 |
| **65** | **Spry Regulation Of FGF Signaling R-HSA-1295596** | 0.011145440218903038 | 0.04162586683536024 |
| **66** | **Signaling By NTRK3 (TRKC) R-HSA-9034015** | 0.011145440218903038 | 0.04162586683536024 |
| **67** | **Chromatin Modifying Enzymes R-HSA-3247509** | 0.011678264928131234 | 0.04162586683536024 |
| **68** | **Gastrin-CREB Signaling Pathway Via PKC And MAPK R-HSA-881907** | 0.01183818865932929 | 0.04162586683536024 |
| **69** | **Tie2 Signaling R-HSA-210993** | 0.01183818865932929 | 0.04162586683536024 |
| **70** | **SHC-mediated cascade:FGFR3 R-HSA-5654704** | 0.01183818865932929 | 0.04162586683536024 |
| **71** | **VEGFR2 Mediated Cell Proliferation R-HSA-5218921** | 0.01253048655713669 | 0.04162586683536024 |
| **72** | **Constitutive Signaling By Ligand-Responsive EGFR Cancer Variants R-HSA-1236382** | 0.01253048655713669 | 0.04162586683536024 |
| **73** | **Signaling By FLT3 Fusion Proteins R-HSA-9703465** | 0.01253048655713669 | 0.04162586683536024 |
| **74** | **Signaling By PDGFR In Disease R-HSA-9671555** | 0.013222334171232074 | 0.04162586683536024 |

| | | | |
|---|---|---|---|
| **75** | **Signaling By KIT In Disease R-HSA-9669938** | 0.013222334171232074 | 0.04162586683536024 |
| **76** | **Regulation Of TP53 Activity Thru Methylation R-HSA-6804760** | 0.013222334171232074 | 0.04162586683536024 |
| **77** | **SHC-mediated cascade:FGFR4 R-HSA-5654719** | 0.013222334171232074 | 0.04162586683536024 |
| **78** | **RAS Processing R-HSA-9648002** | 0.013222334171232074 | 0.04162586683536024 |
| **79** | **FRS-mediated FGFR3 Signaling R-HSA-5654706** | 0.013222334171232074 | 0.04162586683536024 |
| **80** | **Signaling To RAS R-HSA-167044** | 0.013222334171232074 | 0.04162586683536024 |
| **81** | **Ras Activation Upon Ca2+ Influx Thru NMDA Receptor R-HSA-442982** | 0.013222334171232074 | 0.04162586683536024 |
| **82** | **TP53 Regulates Transcription Of Genes Involved In Cytochrome C Release R-HSA-6803204** | 0.013913731760820029 | 0.042238114273917944 |
| **83** | **CD209 (DC-SIGN) Signaling R-HSA-5621575** | 0.013913731760820029 | 0.042238114273917944 |
| **84** | **SHC-mediated cascade:FGFR1 R-HSA-5654688** | 0.013913731760820029 | 0.042238114273917944 |
| **85** | **Signaling By FGFR3 In Disease R-HSA-5655332** | 0.014604679593987663 | 0.04232037836894153 |
| **86** | **Signaling By ERBB2 TMD/JMD Mutants R-HSA-9665686** | 0.014604679593987663 | 0.04232037836894153 |
| **87** | **SHC1 Events In ERBB2 Signaling R-HSA-1250196** | 0.014604679593987663 | 0.04232037836894153 |
| **88** | **FRS-mediated FGFR4 Signaling R-HSA-5654712** | 0.014604679593987663 | 0.04232037836894153 |
| **89** | **FRS-mediated FGFR1 Signaling R-HSA-5654693** | 0.01529517792933681 | 0.0433363374664543 |
| **90** | **SHC-mediated cascade:FGFR2 R-HSA-5654699** | 0.01529517792933681 | 0.0433363374664543 |
| **91** | **Signaling By NTRK2 (TRKB) R-HSA-9006115** | 0.015985227036696047 | 0.04383588588026018 |
| **92** | **Signaling By ERBB2 KD Mutants R-HSA-9664565** | 0.016674827177981324 | 0.04383588588026018 |
| **93** | **Signaling By Erythropoietin R-HSA-9006335** | 0.016674827177981324 | 0.04383588588026018 |
| **94** | **Signaling By EGFR In Cancer R-HSA-1643713** | 0.016674827177981324 | 0.04383588588026018 |
| **95** | **Downstream Signaling Of Activated FGFR3 R-HSA-5654708** | 0.016674827177981324 | 0.04383588588026018 |
| **96** | **Impaired BRCA2 Binding To PALB2 R-HSA-9709603** | 0.016674827177981324 | 0.04383588588026018 |
| **97** | **FRS-mediated FGFR2 Signaling R-HSA-5654700** | 0.016674827177981324 | 0.04383588588026018 |
| **98** | **Defective HDR Thru Homologous Recombination (HRR) Due To BRCA1 Loss-Of-Function R-HSA-9701192** | 0.017363978615388416 | 0.04466297938838022 |
| **99** | **Signaling By ERBB2 In Cancer R-HSA-1227990** | 0.017363978615388416 | 0.04466297938838022 |
| **100** | **Synthesis Of IP3 And IP4 In Cytosol R-HSA-1855204** | 0.01805268161758667 | 0.04466297938838022 |