ESCOLA
SUPERIOR
DE SAÚDE
POLITÉCNICO
DO PORTO

M

MESTRADO
TÉCNICAS LABORATORIAIS EM BIOPATOLOGIA- PATOLOGIA MOLECULAR

# DGCR8 microprocessor defect and deregulation of its expression in thyroid lesions

Lia Cristiana Gomes Rodrigues

10/**2021**

P.PORTO

ESCOLA SUPERIOR DE SAÚDE


# DGCR8 microprocessor defect and deregulation of its expression in thyroid lesions

**Autor**

Lia Cristiana Gomes Rodrigues

**Orientadores**

PhD/João Pedro Rico de Oliveira Vinagre/Instituto de Investigação e Inovação em Saúde da Universidade do Porto (I3S)

PhD/Ana Paula Soares Dias Ferreira/I3S

PhD/Regina Augusta Alves Pereira da Silva/Escola Superior de Saúde do Instituto Politécnico do Porto

Projeto apresentado para cumprimento dos requisitos necessários à obtenção do grau de Mestre em **Técnicas Laboratoriais em Biopatologia** – Ramo de Especialização em **Patologia Molecular** pela Escola Superior de Saúde do Instituto Politécnico do Porto.

## Agradecimentos

Ao meu orientador, Doutor João Vinagre, por todo o apoio, disponibilidade e conhecimentos transmitidos ao longo do último ano, que tão importantes foram neste percurso e que permitiram o meu enriquecimento a nível científico e profissional.

Às minha co-orientadoras, Professora Doutora Paula Soares, por me aceitar no grupo "Sinalização e metabolismo do cancro", bem como por toda a afabilidade e disponibilidade que sempre demonstrou; e Professora Doutora Regina Silva, por toda a disponibilidade, apoio e transmissão de conhecimentos durante todo o meu percurso académico.

A todo o grupo "Sinalização e metabolismo do cancro" por me receberem tão bem, por toda a disponibilidade e conhecimentos transmitidos.

Ao meu João, por todo o apoio e companheirismo ao longo dos últimos anos, obrigada por sempre acreditares em mim.

À minha mãe e irmãs, por toda a cumplicidade, apoio, união e amor que sempre partilhamos. Obrigada por tornarem isto possível.

Aos meus sobrinhos, Álvaro e Luana, que despertam os melhores sentimentos em mim e me fazem acreditar num futuro melhor.

Às minhas melhores amigas e a todos os meus grandes amigos, por me acompanharem, desencaminharem, mas acima de tudo me apoiarem e se orgulharem de mim.

A todos que se cruzaram no meu caminho e que, direta ou indiretamente, tornaram isto possível. Aos que foram o caminho, aos que ficaram pelo caminho e aos que estarão sempre à chegada: o meu muito obrigada!

**Resumo**

A alteração no processamento de microRNAs (miRNAs) pode ser um evento primário em tumores, incluindo o cancro da tireóide (CT). O gene DiGeorge Criticial Region 8 (*DGCR8*) tem um papel crítico na biogénese dos miRNAs e no desenvolvimento da tireóide como componente do complexo microprocessador. Estudos prévios identificaram a variante *DGCR8:* c.1552G>A p.(E518K) em casos de CT e numa síndrome familiar caracterizada por bócio multinodular e schwannomatose. O objetivo deste estudo foi caracterizar esta variante numa série de lesões da tireóide, e avaliar a expressão do *DGCR8*.

Foi avaliada por sequenciação a presença da variante c.1552G>A p.(E518K). Para casos com RNA disponível, a expressão de mRNA do *DGCR8* foi avaliada por qPCR. Foi ainda avaliada a imunoexpressão do DGCR8 em tecidos parafinados.

Neste trabalho, descreve-se pela primeira vez a presença da variante c.1552G>A p.(E518K) num carcinoma pouco diferenciado da tireóide. Este caso apresentou níveis de expressão de mRNA do *DGCR8* semelhantes ao valor basal, no entanto, apresentou expressão aberrante a nível proteico. Uma desregulação da expressão de mRNA do *DGCR8* em tumores com padrão folicular foi ainda observada.

Os resultados obtidos sugerem que o *DGCR8* pode estar envolvido na tumorigénese da tireóide, sobretudo em tumores com padrão folicular.

**Palavras-chave:** miRNA; cancro da tireóide; complexo microprocessador; *DGCR8*; p.(E518K).

## Abstract

Deregulation of microRNA (miRNA) processing is a driver event in several tumours including thyroid cancer. DiGeorge Critical Region 8 (*DGCR8*) gene holds a critical role in miRNA biogenesis, as a microprocessor complex component, and in the development of the thyroid. Previous studies identified a *DGCR8* mutation – the variant c.1552G>A p.(E518K) – in cases of thyroid cancer and proposed to cause a familial syndrome characterized by multinodular goitre (MNG) and schwannomatosis. The goal of this study was to characterize the variant p.(E518K) of *DGCR8* in thyroid lesions and evaluate its expression.

A series of thyroid lesions were evaluated by sequencing for the c.1552G>A p.(E518K) variant. When frozen tissue was available, *DGCR8* mRNA expression was analysed by qPCR. Formalin-fixed paraffin-embedded tissues were studied for DGCR8 immunoexpression.

In this work, it is described for the first time the hotspot p.(E518K) mutation in a case of poorly differentiated thyroid carcinoma. This case displayed *DGCR8* mRNA expression comparable to the basal level whereas the protein expression was highly aberrant. Deregulation of follicular-patterned tumours was also observed at the *DGCR8* mRNA level.

The obtained data suggest that *DGCR8* could have a role in thyroid tumorigenesis, particularly in follicular-patterned thyroid tumours.

**Keywords:** miRNA; thyroid cancer; microprocessor complex; *DGCR8*; p.(E518K).

## Summary

## List of abbreviations

**A**

**AGO2**– Argonaute

**ATC**– Anaplastic thyroid carcinoma

**B**

**bp**– Base pairs

**C**

**cDNA**– Complementary DNA

**CHUSJ**– Centro Hospitalar e Universitário de São João

**D**

**DGCR8**– DiGeorge Critical Region 8

**DICER**– Dicer ribonuclease III

**ddNTPs**– Dideoxynucleotides

**dNTPs**– Deoxynucleotides

**DRBM**– Double-stranded RNA binding motif

**DROSHA**– Drosha ribonuclease III

**E**

**Exo**– Exonuclease II

**F**

**FFPE**– Formalin-fixed paraffin-embedded

**FTA**– Follicular thyroid adenoma

**FTC**– Follicular thyroid carcinoma

**G**

**GTEx**– The genome-tissue expression

**I**

**IDT**– Integrated DNA technologies

**IHC**– Immunohistochemistry

**L**

**LOH**– Loss of heterozygosity

**M**

**MAPK**– Mitogen-activated protein kinase

**miRNAs**– MicroRNAs

**MNG**– Multinodular goitre

**mRNAs**– Messenger RNAs

**MTC**– Medullary thyroid carcinoma

**N**

**nt**– Nucleotides

**NT**– Normal tissue

**NTAT**– Non-tumoral adjacent tissue

**P**

**PCR**– Polymerase chain reaction

**PDTC**– Poorly differentiated thyroid carcinoma

**pre-miRNA**– Precursor microRNA

**pri-miRNA**– Primary microRNA

**PTC**– Papillary thyroid carcinoma

**Q**

**qPCR**– Quantitative PCR

**R**

**RISC**– RNA-induced silencing complex

**RNAi**–RNA interference

**S**

**SAP**– Shrimp alkaline phosphatase

**siRNA**– Small interfering RNA

**SNP**– Single nucleotide polymorphism

**T**

**TC**– Thyroid cancer

**TCGA**– The Cancer Genome Atlas

**TPM**– Transcripts per million

**U**

**UTR**– Untranslated region

**W**

**wiFTC**– Widely invasive follicular thyroid carcinoma

**WT**– Wild-type

**X**

**XPO5**– Exportin 5

## List of tables

## List of figures

# 1. Introduction

## 1.1. microRNAs (miRNAs)

Only a small fraction of the human genome (around 2%) codes for protein genes, whereas the majority relates to various noncoding RNAs, including canonical microRNAs (miRNAs) (1, 2). The miRNAs were initially discovered in *Caenorhabditis elegans* (*C. elegans*) as regulatory molecules modulating the developmental timing (3). Canonical miRNAs are a class of small, about 22 nucleotides (nt), non-coding single stranded RNAs essential for animal development that control gene expression at the post-transcriptional level through translational inhibition and destabilization of their target messenger RNAs (mRNAs) (1, 3, 4). The miRNAs along with small interfering RNAs (siRNAs) constitute the RNAs that are involved in RNA interference (RNAi) pathway (5).

The discovery of miRNAs challenged the central dogma of molecular biology that the DNA stored genetic information is transcribed into mRNA and translated to protein, since the miRNAs regulate post-transcription gene silencing (2, 6, 7). Therefore, miRNAs are the major post-transcriptional regulators of mRNA expression, leading to deregulation of protein synthesis by activation of RNAi-mediated pathway, Figure 1 (8, 9).

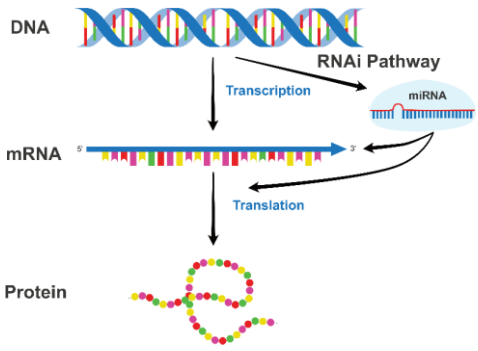



Figure 1: miRNA regulation of gene expression at post-transcription level through the RNAi pathway (8,9).

The miRNA genes may form polycistronic clusters or exist individually and are estimated to account for 2-5% of the human genes being frequently located at fragile sites and genomic regions involved in cancer (2, 9, 10).

It is known that miRNA genes are encoded in the genome of most eukaryotic organisms and transcribed by RNA polymerase II (RNA pol II) into long, poly-adenylated, and capped primary miRNAs (pri-miRNAs) in the nucleus (11-13). These structured RNAs are then processed by the microprocessor complex – a trimeric nuclear complex composed by two DiGeorge Critical Region 8 (DGCR8) proteins binding to one Drosha Ribonuclease III (DROSHA) – and converted in approximately 60 to 70 nt hairpin-

shaped intermediates, the so-called precursor miRNAs (pre-miRNAs), as presented in Figure 2 (11, 13). The pre-miRNAs are then exported to the cytoplasm by the nuclear transport receptor exportin-5 (XPO5), where the Dicer ribonuclease III (DICER) cleaves the base of the loop to generate about 21 to 24 nt double-strand miRNA duplex (11). The duplex is unwound, and one strand is preferentially selected to bind to one of the Argonaute (AGO2) proteins to generate the final and mature form of miRNA (13, 14). This mature miRNA is incorporated into a ribonucleoprotein complex, known as the RISC (RNA-Induced Silencing Complex) (1, 11, 13). The selection of the RNA strand to be incorporated in the RISC is determined by the stability of the base pairs at the 5' end of the duplexes and the remaining strand is degraded (13). The miRNAs regulate gene expression post-transcriptionally, acting as a negative regulator of coding gene expression, by guiding the RISC to their cognate sites at the 3'-untranslated region (UTR) of target mRNAs (15, 16). Although most miRNAs target the 3' UTR of the mRNA targets, some miRNA can also interact with the coding sequence of the target gene (17). The targeted mRNA will be initially subjected either to cleavage or translation repression by inhibiting ribosomal access, depending on whether the miRNA:mRNA pairing is perfect or not, respectively (1, 11, 16, 18).

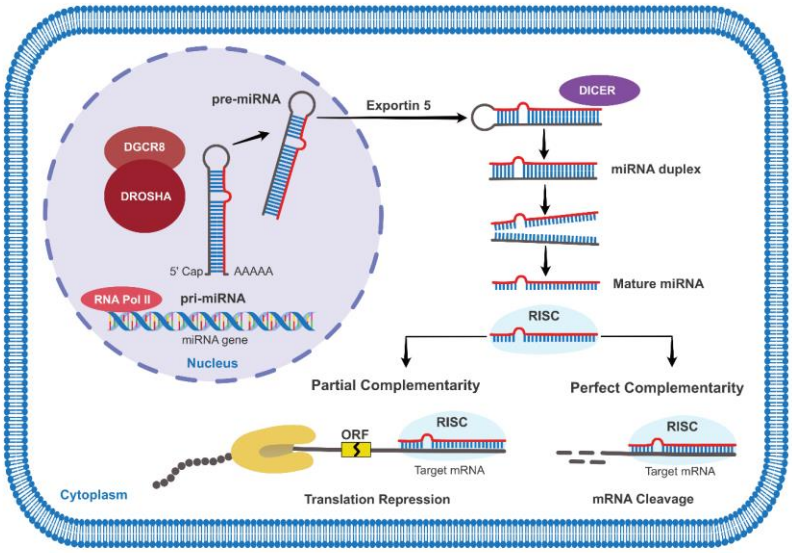



Figure 2: Schematic representation of the miRNA biogenesis. DNA sequences are transcribed into RNA sequences that form a "hairpin" structure pri-miRNA by RNA polymerase II. The endoribonuclease DROSHA and its RNA binding partner DGCR8 cleaves the hairpin from the primary RNA strand into pre-miRNA. Then XPO5 transport pre-miRNA out of the nucleus. In the cytoplasm, the endonuclease DICER processes the pre-miRNA into mature miRNA. miRNA is incorporated into RISC and it will result in translational repression and inhibition of protein synthesis (11,13).

A single miRNA has the potential to regulate more than 200 different genes. By performing bioinformatic prediction for miRNA targeting it is postulated that more than 50% of human mRNAs may be influenced by miRNA-mediated regulation, suggesting that miRNAs may be involved in all biological degrees (1, 6). There are alternatives for canonical miRNA pathway, instead of microprocessor complex there is the presence of a spliceosome, where a small number of miRNAs are derived from pre-miRNA that are produced from introns by the spliceosome and debranching enzymes (19). The introns excised by the spliceosome are directly cleaved and loaded by DICER, being the remaining process comparable to the canonical pathway (20).

### 1.2. miRNAs in cancer

The relevant role of miRNAs in human cancer is probably due to the fact that the majority of miRNA genes are located at chromosomal fragile sites or common break point sites or within regions of deletion or amplification that are generally altered in tumours (13). They are present in the minimal regions of loss of heterozygosity (LOH), in the minimal regions of amplifications or the common breakpoints regions, suggesting that miRNAs may be a new class of genes involved in human tumorigenesis (10). The miRNAs regulate cellular processes such as cell proliferation, apoptosis, growth, senescence, adhesion, invasion, and migration by targeting the mRNAs involved in a variety of cancer-related signalling pathways (6, 8). Mutations in miRNA target sites that lead to an incorrect mRNA recognition may induce severe phenotypic consequences and may promote cancer initiation (20). The miRNAs may affect the initiation, development, and progression of cancer through alteration of the expression levels of their target genes (21). The major evidence that links miRNAs with cancer is their large alteration in expression in malignant cells compared to benign cells (9).

The miRNA expression profile of human tumours is characterized by a general defect in miRNA production resulting in global miRNA downregulation (21). This expression is deregulated in human cancer by through various mechanisms, such as amplification or deletion of miRNA genes, abnormal transcriptional control of miRNAs, epigenetic changes and defects in the miRNA processing machinery (9).

More than 2000 miRNAs have been identified in humans (20). Voorhoeve *et al.* (21) presented miRNAs as "the good, the bad and the ugly": "the good" miRNAs as innocent bystanders in the oncogenic transformation process, whose expression profile can be surrogate for cancer diagnosis and prognosis; "the bad" miRNAs that are causally linked to tumorigenesis and modify tumour suppressor or oncogenic pathways; and, "the ugly" miRNAs representing the ones whose inappropriate loss or gain destabilizes the cellular identity of a tumour, resulting in enhanced phenotypic variability and tumour progression (21).

The analysis of global miRNA expression in cancer patients determined different patterns of miRNA overexpression or downregulation in cancer when compared with their normal counterparts (13, 22). Overexpression of miRNAs – by overexpression, amplification, or loss of epigenetic silencing of a gene encoding an oncogenic miRNA – could result in downregulation of tumour suppressor genes (oncogenic miRNAs); and, downregulation of miRNAs– by deletion, subtle mutation, or epigenetic silencing tumour of the suppressive miRNAs – could result in upregulation of oncogenes (suppressor miRNAs) with subsequent effects on cell proliferation, apoptosis, angioinvasion, and other carcinogenic actions, as represented in Figure 3 (2, 21, 23).

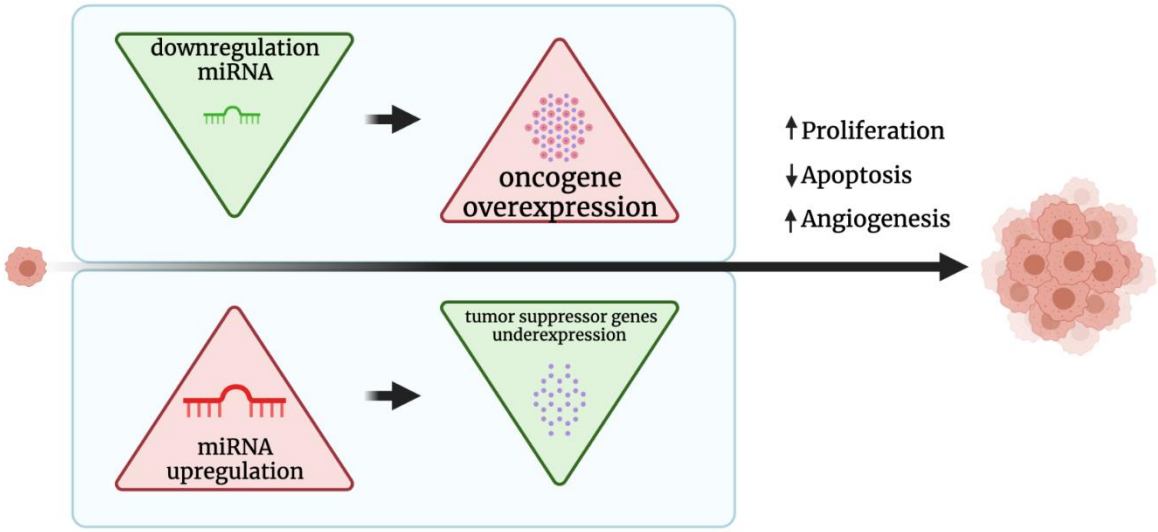



Figure 3: Putative role of suppressor miRNAs and oncogenic miRNAs in carcinogenesis (21,23).

Studies support the role of miRNAs in the initiation and progression of human malignancies (13). An overall downregulation of miRNAs expression is observed in tumoral tissue (6, 24). Moreover, poorly-differentiated tumour samples present a lower miRNAs expression compared to more differentiated ones, which is also observed in human cancer cell lines, being consistent with the hypothesis that a higher global miRNAs expression is associated with cellular differentiation (13). This effect of global miRNAs downregulation was reproduced *in vitro* by knocking-down miRNAs machinery (6).

The regulation of miRNAs machinery components could directly influence expression patterns of various genes by regulating mRNA expression – if any miRNAs machinery component is deregulated, miRNAs may be incompletely matured (25).

The miRNA expression profiling of human tumours has identified signatures associated with diagnosis, staging, prognosis, and response to therapy (26). Hence, manipulating miRNAs expression levels may serve as a potential therapeutic strategy against cancer (21).

### 1.3. miRNAs and thyroid cancer

The miRNAs biogenesis pathway is postulated to hold a key role in early and the proper development of the thyroid gland being miRNAs necessary for accurately establishing thyroid follicles and hormone synthesis (6, 15). As in many types of human cancer, in thyroid carcinomas, miRNAs expression is also deregulated (2).

Thyroid cancer (TC) is the most widespread endocrine malignancy and is a paradigm for the use of molecular information in the clinical field, particularly for diagnosis purposes (23). TC represents an attractive model for studying the events involved in epithelial cell multistep carcinogenesis as these comprise a broad spectrum of lesions with different degrees of malignancy (10). Most thyroid carcinomas originate from thyroid follicular cells being subdivided into well-differentiated papillary thyroid carcinoma(s) (PTC(s)) and follicular thyroid carcinomas (FTC(s)), which accounts for about 75-80% and 10-15%, respectively (20). Both PTC and FTC may progress to poorly differentiated thyroid carcinoma(s) (PDTC(s)), or give rise to anaplastic thyroid carcinoma(s) (ATC(s)) if they completely lose differentiation (7, 26). Follicular thyroid adenoma(s) (FTA(s)) are benign thyroid tumours. Less than 5% of the cells within the thyroid gland are C- cells that give rise to a neuroendocrine tumour- medullary thyroid carcinoma (MTC) (20).

Studies revealed that miRNA mutations can lead to cancer predisposition, as it is the case of a single nucleotide polymorphism (SNP) in miR-146 that was found to contribute to TC (21).

The activation of oncogenes is a known cause of miRNA global deregulation in thyroid cells (6, 7). An important proliferative pathway in thyroid cells is the mitogen-activated protein kinase (MAPK) signalling pathway that when activated by oncogenes, i.e., *RET/PTC*, *RAS*, and *BRAF*, promotes sustained cell proliferation and these mutations are detected in more than 70% of PTC (6, 7). Thyroid malignant transformation by MAPK oncogenes is accompanied by global miRNA changes, with a marked reduction of suppressive miRNAs and activation of oncogenic miRNAs (6). Evidence suggests that deregulated activation of the MAPK cascade can increase the genomic instability of TC cells, thus promoting the acquisition of additional somatic mutations during TC progression (20).

The normal thyroid gland highly expresses some miRNAs that are all downregulated in thyroid carcinomas, indicating that these "normal thyroid miRNAs" are necessary to maintain thyroid follicular cell differentiation and function (2, 6). These studies suggested a vital role for specific miRNAs as key factors in the development and progression of TC (26).

Although thyroid cells do not express a characteristic miRNA, they do express a pool of miRNAs associated with the differentiated status. This miRNA regularly expressed in the normal thyroid gland are downregulated in TC, indicating an important tumour-suppressor action by those miRNAs (6). It is a hypothesis that, among all the deregulated miRNAs, only those which are abundantly overexpressed or strongly downregulated are involved in thyroid tumorigenesis (13). The let-7 miRNA, one of the highly expressed in the normal thyroid gland, targets RAS, a component of the MAPK pathway (6). Let-7 was shown to have complementary binding sites in the 3'-UTR of all 3 *RAS* genes (*HRAS*, *KRAS*, and *NRAS*) and it functions to decrease RAS protein levels (23). Because RAS activation is related to several cancers, let-7 downregulation or deletion could be crucial in tumorigenesis (23).

Previous studies described that many miRNAs are expressed in a tissue-specific manner and it has been shown a markedly different profile of miRNA expression between MTCs and all tumours that derive from follicular cells (7).

The miRNA expression profile presents a significant variability between different kinds of thyroid cancers, even if they originate from the same type of thyroid cell, which could be useful as diagnostic and therapeutic markers (9). The miRNA expression profile is different not only between tumoral and normal tissues but also between different types of tumours, stages, and primary and metastatic tumours (10). Expression profiling of miRNA was found to be able to distinguish tumours with *BRAF* mutation from the other tumour types as well as make a distinction between the more aggressive insular and anaplastic tumours and the classic variant (20, 26).

A study of miRNA expression in PTC with known mutations revealed a strong correlation between the miRNA profile and mutational status. PTC positive for *BRAF*, *RET/PTC*, and *RAS* mutations, and those with no known mutations demonstrated significant differences in the expression of certain miRNAs (2).

The involvement of miRNAs in thyroid tumours has recently changed the paradigm for biomarker discovery in TC, suggesting that these small non-coding RNAs could be used to develop, refine or strengthen strategies for diagnosis and management of TC (20). These miRNAs are frequently deregulated in cancer and constitute a new class of blood-based biomarkers useful for cancer detection and prognosis definition, including in TC (10). Recent studies described the use of some miRNAs as biomarkers in diagnosis, prognosis, and therapeutic targets of TC as different subtypes of TC are associated with specific miRNA profiles (2). It has been described that 4 miRNAs (miR-100, miR-125b, miR-138, and miR-768-3p) distinguished between benign and malignant thyroid tissue samples (27). It is also described an important role of miR-221 in thyroid cell proliferation (22). Consequentially, two miRNA-based molecular tests are already commercially available for indeterminate thyroid nodules on fine-needle aspiration cytology (8, 9).

## 1.4. miRNA machinery components: DGCR8, a microprocessor complex component

A panoply of studies has demonstrated that human disorders, including cancer, are frequently associated with global alterations in the miRNA machinery components, comprising irregular expression and function of the key factors DROSHA, DGCR8, and AGO2 (25, 28). The effect of the up- or down- regulation of the miRNA processing genes seems to be tissue-specific (29).

Altered miRNA expression in cancer is related to the malfunction of DICER and miRNA-machinery associated proteins (6). Data extracted from The Cancer Genome Atlas (TCGA) database shows that truncating mutations in miRNA-machinery components *DICER1* (18 mutations), *DROSHA* (17 mutations), and *DGCR8* (10 mutations) have been detected in a wide variety of tumours investigated (6). However, the impact of these molecules deregulated in the thyroid tumorigenesis is an issue that remains to be addressed. Altered expression of miRNA machinery components may play an important role in the progression of TC (25). It is described that *DGCR8* acts as an oncogene in tumorigenesis of some types of cancer (28).

Alongside the alterations in expression, mutations in the genes involved in the processing of miRNAs are reported both at the somatic and germline levels (29). The discovery of *DICER1* germline mutations identified the first cancer predisposition syndrome that was caused by impaired miRNA biogenesis (15, 16). DICER1 syndrome results in deregulation of miRNA, most likely due to the loss-of-function and altered expression of genes preventing cancer development (tumour suppressors) or in the gain of function of genes that contribute to cancer onset (oncogenes) (16, 30). As a component of DICER1 syndrome, a series of thyroid disorders were identified, in particular, the development of multinodular goitre (MNG) and differentiated follicular and papillary thyroid carcinomas, indicating the relevance of miRNA in human thyroid function (6, 16, 31). The mutations associated with this syndrome result in aberrant cleavage of precursor miRNAs (15).

*DICER1* is the only miRNA biogenesis gene in which germline mutations have been, so far, identified to cause a syndrome. Somatic *DICER1* mutations are reported in follicular-patterned lesions of thyroid (benign and malignant) (29). Other somatic mutations in genes encoding miRNA biogenesis proteins, as in the case of Wilms' tumours, were also reported to be mutated (*DROSHA*, *DGCR8*, *TARBP2*, *XPO5*) (2, 15). SNPs in miRNA biosynthesis genes *DROSHA* and *DGCR8* were indicated to be correlated with cancer risk (32). As the SNP is the most prevalent genetic alteration, in *DROSHA* and *DGCR8* genes it can affect their structure or expression, resulting in incomplete miRNA processing and consequently influence the expression of target genes, thereby acting as a risk factor for diseases such as cancer (32).

It was shown a causal relationship between the presence of *RET* mutation and increase expression of *DICER*, *DGCR8*, and *XPO5* genes (33). It could be hypothesized that expression of *DICER*, *DGCR8*, and *XPO5* is subjected to *RET* regulation in a *RAS*-independent manner (33).

In the case of *DGCR8*, it is suggested that it acts as a critical link between extracellular proliferative cues and reprogramming of the cellular miRNA profile (34).

DGCR8 is a double-stranded RNA binding protein and is reported as a heme-binding protein that functions as the non-catalytic subunit of the microprocessor complex (1). To exert a DGCR8 role in the miRNA biogenesis, two paired DGCR8 proteins bind one DROSHA, forming a trimeric nuclear complex known as the microprocessor (4). Its fidelity is critical for the generation of functional miRNAs (1, 19, 35). DGCR8 plays multiple roles by binding and stabilizing DROSHA, as well as activating the cleavage activity of DROSHA (1, 35). DGCR8 is the microprocessor component that directly interacts with pri-miRNAs and functions as a molecular anchor and directs the cleavage by DROSHA that functions as an endonuclease that cleaves 11 base pairs (bp) away from the junction to release hairpin-shaped miRNAs (14, 35). DGCR8 is capable of specifically binding pri-miRNAs and discriminating against nonspecific nucleic acids in the absence of DROSHA and thus is likely a major contributor to recognition of pri-miRNAs (19). DGCR8 is primarily localized in the nucleus, especially in the nucleolus and small foci adjacent to splicing speckle (19).

Knock-down of *DGCR8* results in, as observed upon *DROSHA* depletion, a pronounced decrease in mature miRNA level. Depletion of both *DROSHA* and *DGCR8* resulted in a substantial accumulation of pri-miRNAs showing the requirement of the microprocessor complex for miRNA processing *in vivo* (11). In another approach, the conditional knock-out of *DGCR8* at early stages of thyroid specification, also leads to severe hypothyroidism with almost undetectable free thyroxine, thyroid tissue disorganization and few follicular structures (6).

*DGCR8* has been found downregulated in several tumours, such as prostate, colorectal, and epithelial skin cancer, contributing to aggressiveness (33, 36). Overexpression of *DGCR8* is described in colorectal and gastrointestinal cancer, and salivary gland pleomorphic adenomas (34, 37).

*DGCR8* expression is reported to be negatively regulated by the DROSHA-DGCR8 complex at the post-transcriptional level through mRNA levels, so if DROSHA and DGCR8 levels are increased in the cell, DROSHA-DGCR8 complex will be cleaved and destabilize the *DGCR8* mRNA, resulting in the reduction of DGCR8 (34). This autoregulatory feedback loop may help maintain the homeostatic control of miRNA biogenesis (34).

Studies revealed that the deletion of *DGCR8* leads to a deficiency in producing all canonical miRNAs (1, 4). Impaired miRNA processing caused by the aberrant expression of miRNA biosynthesis genes *DGCR8* and *DROSHA* can noticeably promote tumorigenesis (32).

### 1.5. *DGCR8* mutation, the p.(E518K)

*DGCR8* gene localizes to chromosome 22 (22q11.2) and it has a critical role in miRNA biogenesis being a component of the microprocessor complex (15). The same region in 22q11.2 is heterozygously deleted in DiGeorge Syndrome and is the most frequent chromosomal deletion found in humans with an incidence of 1 in 2000–4000 live births (15, 19). The 22q11.2 microdeletion leads to upregulation of several pri-miRNAs accompanied by downregulation of a subset of mature miRNAs (19). This 22q deletion is also identified to lead to *DGCR8* haploinsufficiency, resulting in a decrease in microprocessor efficiency and deregulation of miRNA expression which may contribute to a deregulation of their target mRNA (38).

The ploidy of thyroid lesions was extensively studied by several authors, and it is accepted that the presence of aneuploidy is an adverse prognostic factor in thyroid carcinomas, although aneuploidy can be detected in benign and malignant lesions. Our group previously demonstrated that loss of 22q is highly prevalent in widely invasive follicular thyroid carcinomas (wiFTCs) (39). In fact, the 22q is of major interest when dealing with other thyroid defects as is commonly found to be lost in these lesions leading to LOH.

Previous studies identified the mutation c.1552G>A in exon 6 of *DGCR8*, which codes a glutamic acid to lysine substitution in position 518, p.(E518K). This mutation is a somatic hotspot in Wilms' tumours, it has been identified in two PTCs and in the germline of three-generation family with euthyroid MNG and schwannomatosis, and more recently in 2 wiFTCs (15, 39, 40). A biallelic alteration of *DGCR8* is described in all cases: p.(E518K) mutation plus somatic loss of the whole chromosome 22 (15). This combination suggests a critical role for c.1552G>A p.(E518K) in predisposing to tumour development. Somatic loss of chromosome 22, containing the wild-type (WT) allele, appears to be required for tumorigenesis (15).

This mutation disrupts global miRNA production and *DGCR8* mutated tumours show a specific miRNA profile different from *DGCR8*-WT tumours (15). The mutation located in helix 1 of the first of two double-strand RNA-binding (DSRB) domains within DGCR8, is suggested to be translated. The miRNA and RNA profiling demonstrated that this mutation disrupts pre-miRNA production, impacting the population of canonical miRNAs and mirtrons giving rise to a miRNA profile identifiable across different tissue types (15). This reduction of critical miRNAs in tumours is consistent with the observation that the knockdown of *DGCR8* promotes tumour growth (28).

*In silico* modelling predicts that p.(E518K) mutation would likely reduce the affinity of RNA binding to DGCR8 (15). This mutation is in a highly conserved amino acid region, conserved in 95% of the 100 top sequences by UniProt (6, 15, 41). It is also predicted to be pathogenic by multiple algorithms, is expressed at the RNA level, does not affect splicing, and is not subject to nonsense-mediated decay

(15). Indeed, it is easy to conceive that protein overexpression resulting from defective miRNA-based mRNA regulation may compromise normal cell function and cause genetic diseases (11).

When combining the facts that in DICER1 syndrome, affecting the miRNA biogenesis, thyroid defects (MNG and differentiated thyroid carcinomas) are common, that the study of a partner protein (DGCR8) has an important role in miRNA processing and that miRNAs are quite important in the development of thyroid gland, DGCR8 is quite an attractive target and may reveal novel findings in thyroid lesions. Plus, taking in consideration the loss of the remaining 22q in all *DGCR8* mutated cases and the link of 22q with thyroid lesions, *DGCR8* seems to be a susceptibility gene for thyroid lesions.

To this purpose, the main goal of this study was to bridge that gap by characterizing DGCR8 microprocessor defect in a series of thyroid lesions.

Specifically, the aims of this study were:

1. To perform an *in silico* characterization of *DGCR8* gene –mRNA and protein expression;
2. To genotype *DGCR8* hotspot mutation p.(E518K) by direct sequencing in a series of thyroid lesions;
3. To quantify *DGCR8* mRNA expression by real-time PCR (qPCR);
4. To evaluate DGCR8 expression patterns on the tissue samples by immunohistochemistry;
5. To correlate the expression levels of DGCR8 and *DGCR8* mRNA with clinicopathological features of the series of thyroid lesions.

## 2. Material and Methods

### 2.1. *In silico* analysis

*In silico* analysis of *DGCR8* gene, including mRNA expression and protein expression analysis was performed using bioinformatics tools, including databases and portals such as UniPROT (42), GTEx (43), The Human Protein Atlas (44), OncoMX (45), BioMuta (46) and cBioPortal (47).

The data of DGCR8 protein (identifier:Q8WYQ5) was retrieved from UniPROT, a database for protein sequence and functional information (42). For *DGCR8* expression in normal tissues data was extracted from the Genome-Tissue Expression (GTEx) Portal (43). Immunoprofiling data for normal tissue was retrieved from "The Human Protein Atlas" (44) and cancer biomarker-related data from the OncoMX portal (45).

The genomics and transcriptomics data for *DGCR8* were collected from The Cancer Genome Atlas (TCGA) studies that were deposited in the cBioPortal. Data from cBioPortal was filtered from a total of 10967 samples from 10953 patients with mutation and expression information present in 32 studies of TCGA corresponding to the "PanCancer Atlas" studies. For thyroid carcinomas, a cluster of 633 samples from 625 patients was present in 2 studies in cBioPortal, referred to as "Thyroid TCGA Firehose legacy" (48) and "MSKCC, JCI 2016" (49). Alterations in mRNA expression are presented by a Z-score threshold of ± 2, referring to the deviation from the mean in 2 standard deviations, being positive when the samples are above the mean and negative when below (47).

### 2.2. Samples

The samples used in this study were collected at Centro Hospitalar e Universitário de São João (CHUSJ) and retrieved from the Department of Pathology of CHUSJ. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethical Committee of the CHUSJ (approval number CES284-13) (Appendix 1); being an anonymized retrospective study it was exempted from informed consent from patients in accordance with national ethical guidelines.

All clinicopathological data were obtained from the anatomic pathology reports provided by the Department of Pathology from the CHUSJ. For the cases with available formalin-fixed paraffin-embedded (FFPE), tissues were re-evaluated and histological diagnoses were reported according to the strict histomorphological criteria for current World Health Organization guidelines (50). The following clinicopathological parameters were collected from the pathology reports: diagnosis, age at diagnosis, gender, tumour size, presence of tumour capsule, presence of capsular invasion, associated lesions, vascular invasion, lymph node metastasis, extrathyroidal invasion, presence of lymphocytic infiltrate, and other histological observations. According to the availability of adequate tissue,

pathological report and/or clinical information, 226 samples from 209 patients were selected for the subsequent study; this included benign lesions, malignant lesions, and normal adjacent tissue samples. The included samples corresponded to: Normal tissue (NT) (n=5); MNGs (n=15); FTAs (n=86, 2 cases with non-tumoral adjacent tissue (NTAT) available); FTCs (n=22, 6 with NTAT); PTCs (n=80, 2 cases with NTAT); PDTC (n=5, 1 case with NTAT available); and 2 cases of ATC. Regarding the MNGs, 12 cases were composed of germline (blood) DNA of probands from MNGs with familial association. Part of this series was present in a biobank and was previously characterized for *BRAF, RAS,* and *TERT* promoter hotspot mutations, and *RET/PTC* and *PAX8-PPARγ* rearrangements (51).

Complementary to this study, 4 human thyroid cell lines were used. These encompass: a transformed normal human thyroid follicular epithelial cell line, Nthy-ori-3-1; a papillary thyroid carcinoma cell line: TPC-1- with a *CCDC6-RET* fusion; two anaplastic thyroid carcinoma cell lines, 8505C – with a BRAF mutation (V600E) – and T241 – with a *PTEN* mutation p.(D252*fs*) (52-54).

### 2.3. *DGCR8* amplification and genotyping

The primers used for PCR amplification were published by Rivera *et al.* (15) and produced by IDT (Integrated DNA Technologies, IA, USA); The primer sequences for *DGCR8* exon 6 were: forward: GCCCCTAGTTACTGACATGGT, and reverse: CCCTGACCAAAGTTACACCT with expected fragments size of 250 bp. Working solutions of primers to Polymerase Chain Reaction (PCR) primers were prepared at a concentration of 10 μM. Primer temperature optimization was obtained using a temperature gradient (58ºC to 68ºC) for best annealing and product amplification performance. The amplicons of PCR temperature gradient ran in agarose gel by electrophoresis to select the more specific fragments and with greater efficiency.

To screen for *DGCR8* mutations, DNA previously extracted from tumour tissues of all the samples described was analysed (51). DNA quantification and quality analysis of the samples were determined using a spectrophotometer NanoDrop 1000® (Thermo Scientific, Wilmington, USA).

The PCR was performed using the following conditions for each sample: 5 μL of MyTaq HS Mix 2x (BIO-25045, Bioline, London, UK), 0.5 μL of primer mix at 10 μM (composed of forward and reverse primers for the corresponding exons in equal proportions), 3.5 μL of DNAse and RNAse-free water and 50 ng of DNA; the final reaction volume was 10 μL. The following cycling conditions were used for PCR in MyCycler™ Thermal Cycler (Bio-Rad, Hercules, CA, USA): 2 minutes at 95ºC, 35 cycles of: 20 seconds at 95ºC (denaturation), 20 second at 62ºC (annealing), and 20 seconds at 72ºC (extension). To finish the amplification, one final extension at 72ºC for 1 minute was made.

PCR amplification quality control was performed by agarose gel electrophoresis fragments in which amplification, expected size and quality control for contaminations were verified. For sample visualization, 3 μL of amplicon product were mixed with 1 μL of loading buffer - containing

bromophenol blue for colour perception and glycerol for increased fragment density - and RedSafe (21141, iNtRON Biotechnology, WA, USA) - an intercalating DNA agent for fluorescence development in the transilluminator. The fragments ran in 1.5% agarose gel (0.75 g of agarose in 50 mL of SGBT buffer (GB01.0120, GRISP, Porto, Portugal) 0.5%) at 180 Volts. The image was obtained on the GelDoc XR+ (Bio-Rad) transilluminator and analysed in the LabImage® software (Bio-Rad).

Before Sanger sequencing, each product was purified with 1.5 µL of ExoSap enzyme mix - 0.5 µL Exonuclease I (Exo) (20U/µL, EN0581, ThermoFischer, Ma, EUA) and 1 µL of Shrimp Alkaline Phosphatase (SAP) (1U/µL, 783905000UN, ThermoFischer). For enzymes activation and activity, they were set to 37ºC for 30 minutes and a final denaturation step of 85ºC for 15 minutes was completed for enzyme inactivation.

The sequencing reaction generates DNA strands, with different sizes terminated with deoxynucleotides (dNTPs) and dideoxynucleotides (ddNTPs) labelled with different fluorochromes. Each amplicon was sequenced independently by using the corresponding forward and reverse primer. The sequencing reaction was performed using BigDye™ Terminator V3.1 cycle sequencing kit (4337455, Applied Biosystems, CA, USA). For each sequencing reaction 1 µL of the cleaned-up product were mixed with 3.4 µL of sequencing buffer, 3.5 µL of DNAse and RNAse-free water, 0.3 µL of reverse or forward primer (at 10 µM), and 0.25 µL of BigDye™ Terminator enzyme. The cycling conditions used for the reaction were the following: 95ºC for 2 minutes, followed by 40 cycles of: 15 seconds at 94ºC, 15 seconds at 55ºC, and 3 minutes at 60ºC; the final extension was performed at 60ºC for 10 minutes.

Post-sequencing purification was performed by cross-link dextran gel filtration. A tube for each sample was centrifuged at 3400 rpm for 4 minutes in a 750 µL of Sephadex® G-50(G50150, GE Healthcare, IL, USA) solution for column assembly. The sequencing product was then purified using 10 µL of PCR product centrifuged in the Sephadex® column and completed with 15 µL of hi-di formamide (4311320, Applied Biosystem). The purified product was then run in an automated sequencer ABI3130 Genetic Analyzer (Applied Biosystems).

### 2.4. Quantitative PCR analysis

Quantitative real-time PCR (qPCR) was performed when RNA of the frozen tissues was available and converted to cDNA using SuperScript™ IV cDNA Synthesis Kit according to the manufacturer's instructions (18091050, ThermoFischer). For some of the samples, NTAT of the cases was included and analysed to create a pool of non-tumoral tissue.

The qPCR evaluation was carried out in 170 samples, of which 5 were from NT, 28 corresponded to NTAT samples, 62 were FTA, 11 FTC, 61 PTC, and 3 PDTC.

The qPCR was performed using the following conditions for each sample: 0.5 µL of each probe: PrimeTime® std qPCR Assay DGCR8 (Hs.PT.58.1414870 IDT, IA, USA)- exon location 3 to 4- and the human TATA-binding protein (huTBP) gene (no. Hs.PT.39a.22214825, IDT) as endogenous control. The final probe concentration was at 20x, as recommended by the manufacturer, and added 3 µL of water, 5 µL of MasterMix (TaqMan Universal PCR MasterMix No AmpErase ® UNG, Applied Biosystems), and 1 µL of cDNA. The sequence detection system was programmed to an initial step of 10 minutes at 95ºC, followed by 50 cycles of: 95ºC for 15 seconds and 60ºC for 1 minute. The amplification level was detected in a QuantStudio™ 5 Real-Time PCR System (Applied Biosystem).

Relative quantification of target genes was determined using the $\Delta\Delta CT$ method, where similar amplification efficiencies between *DGCR8* mRNA and huTBP were obtained, by Livak's linear regression method (55).

## 2.5. Immunohistochemistry

Immunohistochemistry (IHC) was performed when the FFPE tissues were available and it included 99 FFPE tumour sections, being: 30 FTA, 15 FTC, 50 PTC, and 4 PDTC.

IHC was performed using Ultravision Quanto Detection System HRP (TL-125-QHL, Thermo Scientific), according to the manufacturer's instructions. Briefly, deparaffinized and rehydrated sections were subjected to heat-induced antigen retrieval for 45 minutes at 90ºC in 10 mM sodium citrate buffer (pH 6.0) (AP-9003-500, Thermo Scientific). Endogenous peroxidase activity was blocked with Ultavision Hydrogen Peroxidase Block (TA-125-H202Q, Thermo Scientific) and the nonspecific proteins of the tissue with UltraVision Protein Block (TA-125-PBQ, Thermo Scientific). Sections were incubated overnight at 4°C with anti-DGCR8 polyclonal antibody (PA5-40122, Invitrogen) that recognizes the N-terminus region of human DGCR8 protein at the optimized dilution of 1:250, in diluent solution (TA-125-ADQ, Thermo Scientific). Primary Antibody Amplifier (TL-125-QPB, Thermo Scientific) was used for signal amplification. The detection was performed with Polymer method detection system, HRP Polymer Quanto (TL-125-QPH, Thermo Scientific) followed by 3,3′-diaminobenzidine (DAB) reaction and counterstained with Mayer's hematoxylin. A normal thyroid sample was used as a positive control and the negative control consisted in the omission of the primary antibody.

Slides were evaluated by a Pathologist and semi-quantitatively scored in terms of percentage of tumour-stained cells- extension (0: 0–25%; 1: 25–50%; 2: 50–75%; 3: 75–100%), staining intensity (0—negative; 1—weak; 2—intermediate; 3—strong) and cellular localization (N-nuclear or C-cytoplasmatic). An immunohistochemical score was calculated by multiplying the proportion of positive cells by the intensity of the staining, 9 being the maximum score.

Slides were digitalized using a ZEISS axioscan 7 microscope slide scanner (ZEISS, Oberkochen, Germany) and pictures were treated in ZEISS 3.4. blue edition software (ZEISS).

## 2.6. Statistical methods

The statistical analysis was performed using GraphPad Prism version 9.0 (GraphPad Software, Prism, CA, USA) and IBM SPSS version 25 (IBM, NY, USA). Data was evaluated and tested for outliers' determination and for normal gaussian distributions. Populations were compared with ANOVA, if failed gaussian distributions with Kruskal–Wallis test. Comparison of tumours and NTAT was evaluated by paired t-test. For clinicopathological analysis association with the relative gene expression: gender, age, tumour size, diagnosis, histological characteristics, molecular status and DGCR8 expression were analysed using t-test and Mann–Whitney test. Results were considered statistically significant if $p<0.05$.

# 3. Results

## 3.1. *In silico* analysis of *DGCR8*

*DGCR8* gene is located at 22q11.2 and has 4 transcripts that produce 3 isoforms obtained by alternative splicing (42, 56). Regarding normal tissue expression, data retrieved from GTEx points that the highest expression of *DGCR8* is in the brain–cerebellum with 80.00 transcripts per million (TPM) whereas the thyroid ranks the 5$^{th}$ tissue showing the most expression with a median expression value of 35.11 TPM, Figure 4A (43). Consensus normalized expression (NX) levels for 55 tissue types and 6 blood cell types that combine the data from three transcriptomics datasets (HPA, GTEx and FANTOM5) indicate that there is variable expression within different tissues in the organs, Figure 4B. It is not observed a perfect correlation and consistency between antibody staining and RNA expression data, Figure 4 A and C. In the case of thyroid that has a high RNA expression profile (Figure 4B), the given protein expression score is low. Overall, DGCR8 protein has low tissue specificity and is predicted to have intracellular localization and have a general nuclear expression (44).

When evaluating DGCR8 microprocessor complex subunit protein, it is observed that it presents three domains: one tryptophan–tryptophan (WW) domain and two double–stranded RNA binding motif domains (DRBM)– DRBM1 and DRBM2. The WW domain spans from 301 to 334 bp (exon 3), while the DRBM1 goes from 511 to 578 bp (exon 6–8) and DRBM2 from 620 to 685 bp (exon 9–11), Figure 4D. The WW domains may be repeated up to four times in some proteins and are responsible for protein binding. DRBMs are required for efficient binding to pri–miRNA and the region between residues 276 and 498 has an autoinhibitory function on pri–miRNA processing activity. In contrast to other RNA–binding domains, DRBM are found in several proteins that specifically recognize double–stranded RNAs and are mainly involved in post–transcriptional gene regulation, for e.g., by preventing the expression of proteins or by mediating RNAs localization (56). Multiple DRBMs may act in combination to recognize the secondary structure of specific RNAs (56). The region of 1 to 275 is necessary for nuclear localization and retention. On the other hand, the region 276 to 751 is necessary for heme–binding and pri–miRNA processing and the region 701 to 773 for interaction with DROSHA. There are also three low complexity regions in the regions 165 to 176, 410 to 421 and 737 to 748 (57).

In addition to the other proteins involved in miRNA biogenesis, the DGCR8 protein is predicted to interact with TP53 (57).

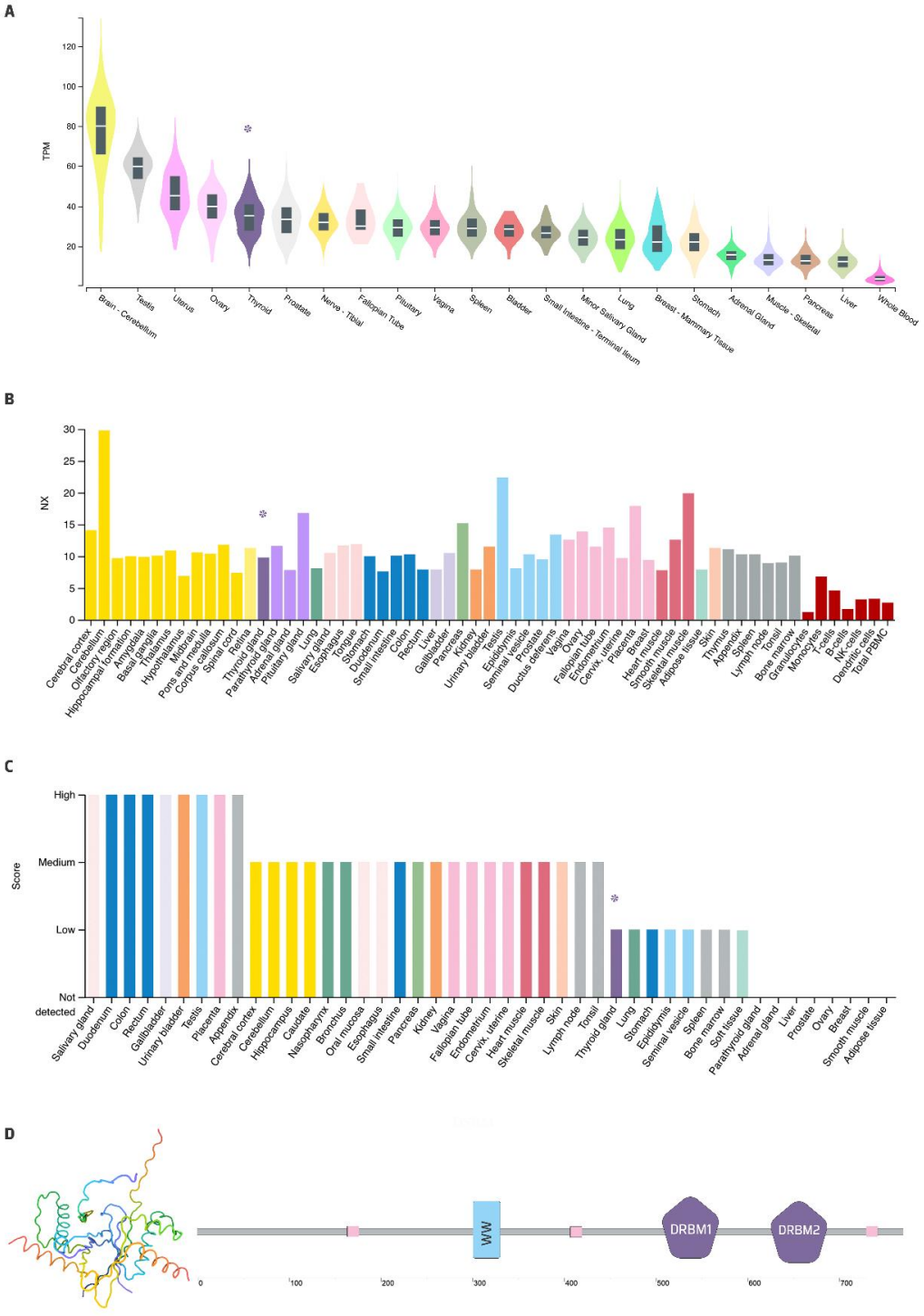



Figure 4: DGCR8 characterization – an *in silico* approach. A: *DGCR8* gene normal tissue expression according to different human organs, data from GTEx (43); B: Transcriptomic overview of *DGCR8* by tissue, figure extracted from The Human Protein Atlas (THPA) (44); C: DGCR8 protein expression in different tissues, figure extracted from THPA (44); and, D: DGCR8 protein principal domains – WW domain and two double–stranded RNA binding domains (DRBM1 and DRBM2), figure adapted from SMART (57).

*: Thyroid gland.

Mutations in *DGCR8* were already detected in different studies and, according to BioMuta(46), for TC only two SNPs are reported for *DGCR8*: one case with a c. 977T>G p.(V326G) and two cases with the c.1552G>A p.(E518K), both probably damaging as evaluated by Polyphen (p=1) (46).

For a broader perception of *DGCR8* mutations in human cancer, data from 10967 samples from 10953 patients available in cBioPortal under the TCGA PanCancer Atlas Studies (47) filter was pulled. Several different types of alterations targeting *DGCR8* gene or its locus are described and include missense mutations - a single nucleotide variation that results in a different amino acid; splice mutations – genetic alterations at the boundary of an exon and an intron (splice site); truncating mutations – a genetic variant that results in a smaller protein; structural variants – copy number variations that could be inversion, deletion or insertions; amplifications – increase in the number of copies of a gene sequence; and deep deletions – deep loss, probably a homozygous deletions (Figure 5A) (47).

In the 10967 cases, 143 mutations were reported. The p.(E518K) mutation is found in two cases of PTC and a truncating mutation (E518*) in one case of melanoma, Figure 5B. A transcriptomic overview, comparing mRNA expression in all mutated and normal cases from those studies points to a slightly higher *DGCR8* expression in mutated cases, Figure 5C.

To perform the same strategic analysis as above for TC, data from cBioPortal that include the sequence studies of Poorly-Differentiated and Anaplastic Thyroid Cancers (MSKCC, JCI 2016) (49) and Thyroid Carcinoma (TCGA, Firehose Legacy) (48) was pulled. The combined studies join 633 samples from 625 patients and fitter data for amplification and missense mutations is available for well-differentiated thyroid carcinomas studies, as represented in Figure 6A (47).

Only one mutation was identified: c.1552G>A p.(E518K), in 2 cases of FV-PTC (Figure 6B). Both cases presented concomitant alterations, being *NRAS* mutated in both cases leading to an altered RTK-RAS pathway. The mRNA expression in mutated and normal cases from those studies were compared, being the expression slightly higher in mutated than in normal cases, as represented in Figure 6C (47).

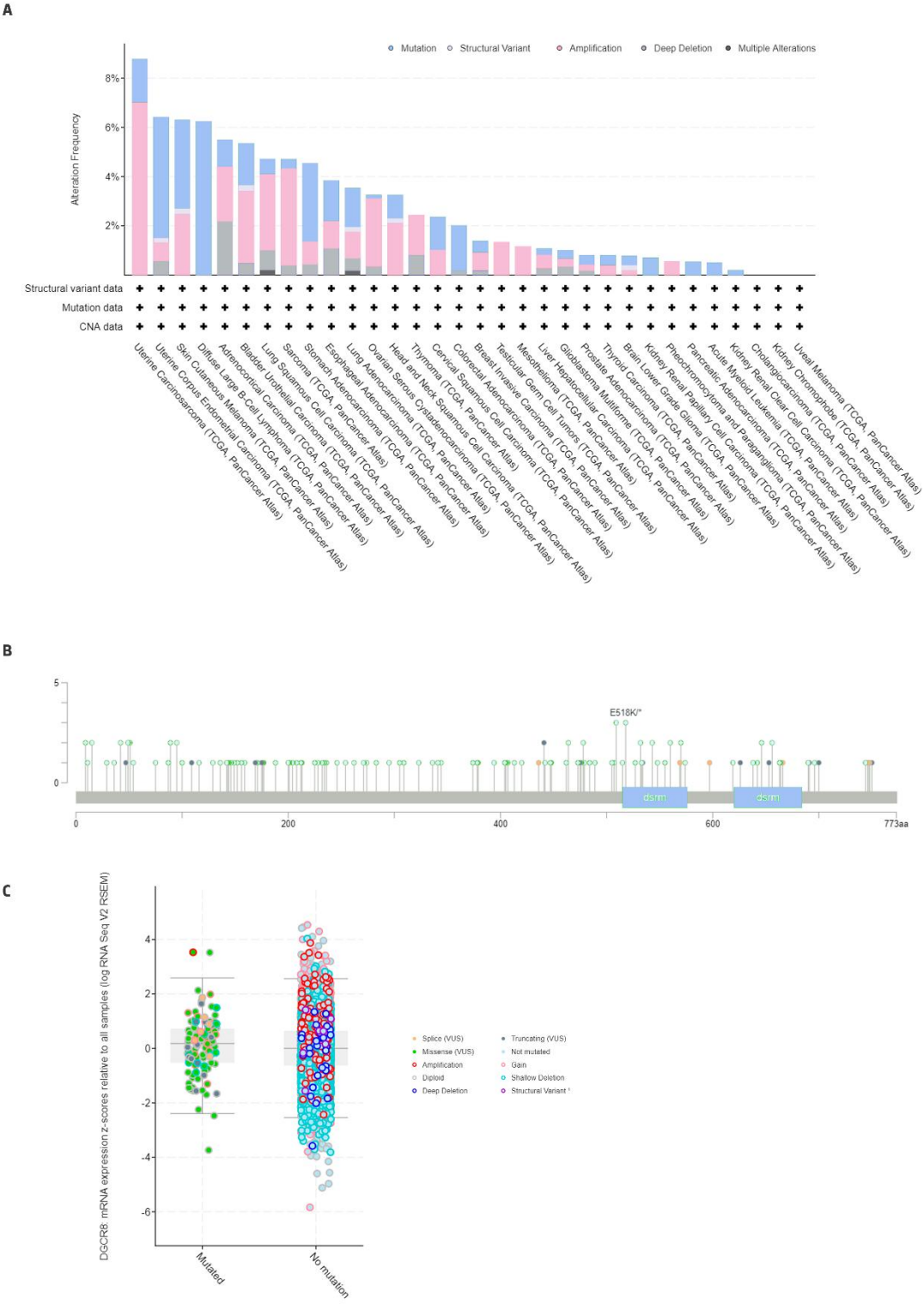



Figure 5: A PanCancer overview of *DGCR8* in cancer. A: *DGCR8* alterations frequency in the 10967 samples of 10953 cases from TCGA PanCancer Atlas Studies; B: Distribution of the genetic alterations (spots) detected in *DGCR8* gene throughout *DGCR8* gene; C: Plot of mRNA expression for mutated and non-mutated cases, -mRNA expression z-scores relative to all samples (log RNA Seq V2 RSEM)- in mutated versus WT cases of TCGA PanCancer Atlas Studies. Figures extracted from cBio Portal for Cancer Genomics (47).

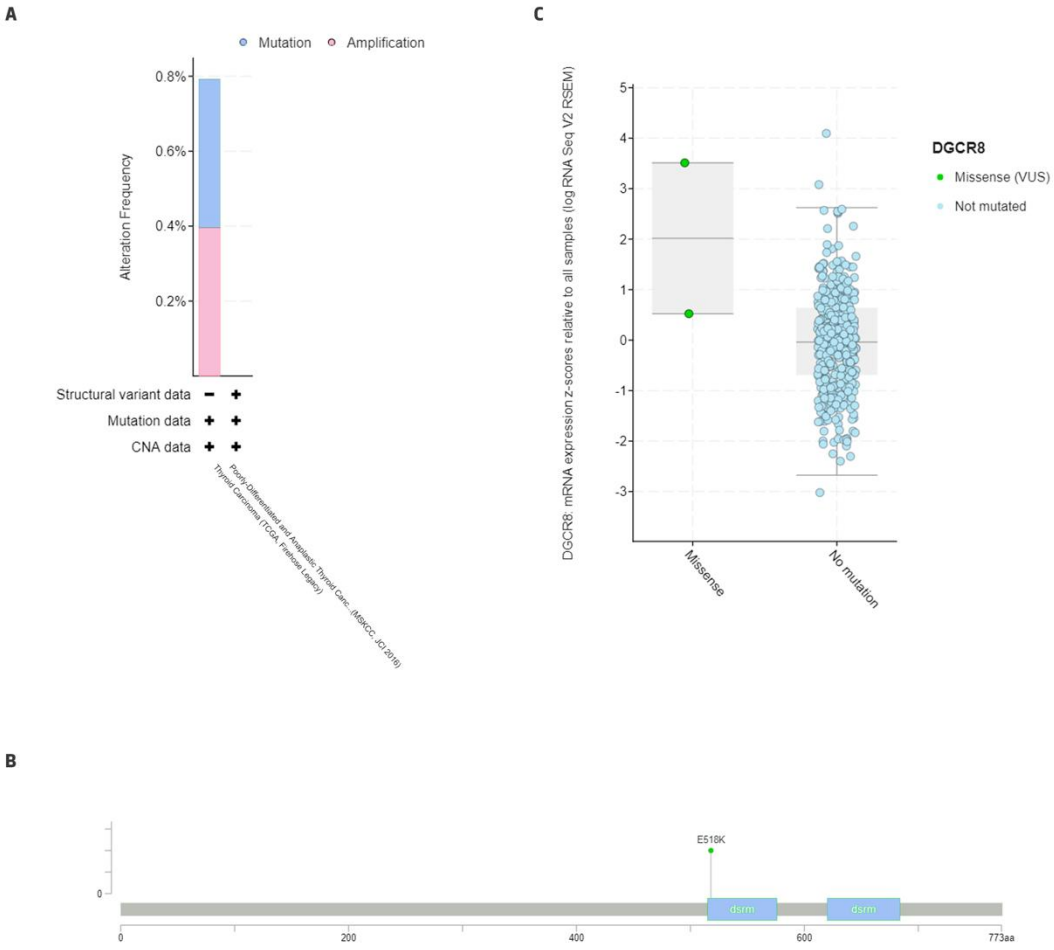



Figure 6: *DGCR8* molecular profiling in thyroid carcinoma. A: Alteration frequency in Thyroid Carcinoma (TCGA, Firehose Legacy) (48) and Poorly-Differentiated and Anaplastic Thyroid Cancers (MSKCC, JCI 2016) (49); B: Distribution of the 2 mutations, that corresponded to the p.(E518K) in FV-PTC cases and were located at the DRBM1 domain; C: Plot of mRNA expression- mRNA expression z-scores relative to all samples (log RNA Seq V2 RSEM)- in mutated (missense) versus normal (no mutation) cases of combined studies of thyroid. All figures were adapted from cBio Portal for Cancer Genomics (47).

### 3.2. c.1552G>A p.(E518K) in a Poorly Differentiated Thyroid Carcinoma

A total of 226 samples from 209 patients were genotyped. The samples included 5 NTs, 15 MNGs, 86 FTAs (2 cases with NTAT), 22 FTCs (6 cases with NTAT), 80 PTCs (2 cases with NTAT), 5 PDTCs (1 case with NTAT) and 2 ATCs. The total of 80 PTCs cases included classical variants of PTC (n= 44; cPTC), follicular variants of PTC (n=22; FV-PTC) and other variants of PTC (n=14; OV-PTC: 3 tall cell PTC, 4 oncocytic PTC, 5 diffuse sclerosing PTC and 2 solid trabecular PTC). In Table 1 are represented the clinicopathological and molecular parameters collected regarding the studied cohort.

Table 1: Clinicopathological and molecular characterization of the series

| | *Histological subtypes* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Clinicopathological variables** | *FTA (n= 86)* | *MNG (n= 15)* | *FTC (n=22)* | *cPTC (n=44)* | *FV-PTC (n=22)* | *OV-PTC (n=14)* | *PDTC (n=5)* | *ATC (n=2)* |
| *Age (mean, y/o)* | 43.2 | 45.3 | 47.8 | 40.2 | 42.3 | 45 | 65.6 | N.D. |
| *Gender n (Females) (%)* | 69 (80.2) | 15 (100.0) | 16 (72.7) | 34 (77.3) | 22 (100.0) | 10 (71.4) | 4 (66.7) | N.D. |
| *Tumour size (mean, mm)* | 35 | 23 | 40 | 25 | 33 | 38 | 51 | N.D. |
| *Lymphocytic infiltrate n (%)* | 20/67 (29.9) | N.D. | 3/14 (21.4) | 17/37 (45.9) | 7/17 (41.2) | 4/13 (30.8) | 0/5 (0.0) | N.D. |
| *Vascular invasion n (%)* | 0/70 (0.0) | N.D. | 8/17 (47.1) | 19/37 (51.4) | 5/18 (27.7) | 6/13 (46.2) | 2/3 (66.7) | N.D.. |
| *Lymph node metastasis n (%)* | 0/1 (0.0) | N.D. | 0/7 (0.0) | 14/22 (63.6) | 5/6 (83.3) | 4/4 (100.0) | 2/3 (66.7) | N.D. |
| *Minimal extrathyroidal extension n (%)* | 0/32 (0.0) | N.D. | 1/14 (7.14) | 15/37 (40.5) | 5/18 (27.7) | 7/12 (58.3) | 2/5 (40.0) | 1/1 (100.0) |
| *Molecular characterization†* | *FTA* | *MNG* | *FTC* | *cPTC* | *FV-PTC* | *OV-PTC* | *PDTC* | *ATC* |
| *TERTp, nm/nt (%)* | 0/84 (0.0) | 0/2 (0.0) | 0/17 (0.0) | 0/41 (0.0) | 0/20 (0.0) | 2/14 (14.3) | 1/4 (25.0) | 0/1 (0.0) |
| *BRAF, nm/nt (%)* | 0/23 (0.0) | 0/2 (0.0) | 0/18 (0.0) | 18/42 (42.9) | 4/20 (20.0) | 4/13 (30.8) | 0/4 (0.0) | 0/1 (0.0) |
| *NRAS, nm/nt (%)* | 4/86 (4.7) | 0/2 (0.0) | 4/21 (19.0) | 2/42 (4.8) | 3/21 (14.3) | 2/12 (16.7) | 0/4 (0.0) | 0/1 (0.0) |
| *RET/PTC, nm/nt (%)* | 0/74 (0.0) | 0/2 (0.0) | 1/13 (7.7) | 5/36 (13.9) | 0/20 (0.0) | 2/14 (14.3) | 0/4 (0.0) | N.D. |
| *PAX8/PPARγ, nm/nt (%)* | 3/76 (3.9) | 0/2 (0.0) | 3/16 (18.8) | 0/41 (0.0) | 1/21 (4.8) | 0/14 (0.0) | 0/4 (0.0) | N.D. |
| *DGCR8, nm/nt (%)* | 0/86 (0.0) | 0/15 (0.0) | 0/22 (0.0) | 0/44 (0.0) | 0/22 (0.0) | 0/14 (0.0) | 1/5 (20.0) | 0/2 (0.0) |

*Notes:* * Not all samples had clinicopathological data available † Not all samples were genotyped.
nm: number of mutated samples; nt: number of total samples genotyped; N.D.: not determined.

A part of the samples used in this series was previously characterized (51) for mutations in *TERT, BRAF, and NRAS* genes and for *RET/PTC* and *PAX8/ PPARγ* rearrangements.

Overall, *TERT* promoter (*TERTp*) mutations were exclusively detected in malignant cases, with a frequency of 14.3% in OV-PTC (two oncocytic carcinomas) and 25.0% in PDTC (one case). The *BRAF* V600E mutation was the most frequent genetic alteration in PTC cases: 42.9% in cPTC (eighteen

cases), 20.0% in FV-PTC (four cases) and 30.8% in OV-PTC (one tall cell, one solid trabecular and two diffuse sclerosing variants). *NRAS* mutations, were detected in benign and malignant lesions: 4.7% in FTA (four cases), 19.0% in FTC (four cases), 4.8% in cPTC and 16.7% in OV-PTC (two cases in each), and 14.3% in FV-PTC (three cases). *RET/PTC* rearrangements were detected in 13.9% of cPTC (five cases), 7.7% of FTC (one case) and 14.3% of OV-PTC (two cases). Finally, the *PAX8/PPARγ* rearrangements, present in follicular-pattern thyroid lesions (58), were detected in 3.9% FTA and 18.8% of FTC (three cases in each) and 4.8% of FV-PTC (one case).

The evaluation of *DGCR8* p.(E518K) was initially performed in a series of familial-associated MNG patients since this (germline) alteration was previously reported to cause a familial syndrome characterized by MNG and schwannomatosis. It was observed that all the familial cases of MNG studied were normal for this exon 6 alteration (where the mutation is located). The hotspot *DGCR8* mutation was then extended to the remaining somatic tissue available in the series.

It was identified for the first time one *DGCR8* mutation – c.1552G>A, p.(E518K)– in a PDTC case (an insular variant) and with a frequency of 20%, Table 1 and Figure 7. This somatic mutation was not detected in the corresponding NTAT, excluding it to be a germline alteration. In terms of other genetic events, no rearrangements *RET/PTC* and *PAX8/ PPARγ*, *TERT*, *NRAS* and *BRAF* mutations were detected in the same case. This mutation (p.(E518K)) occurred in an older patient (82 years-old) with a 10 cm tumour presenting poor prognosis characteristics, such as invasion of capsule and extra-thyroidal invasion.

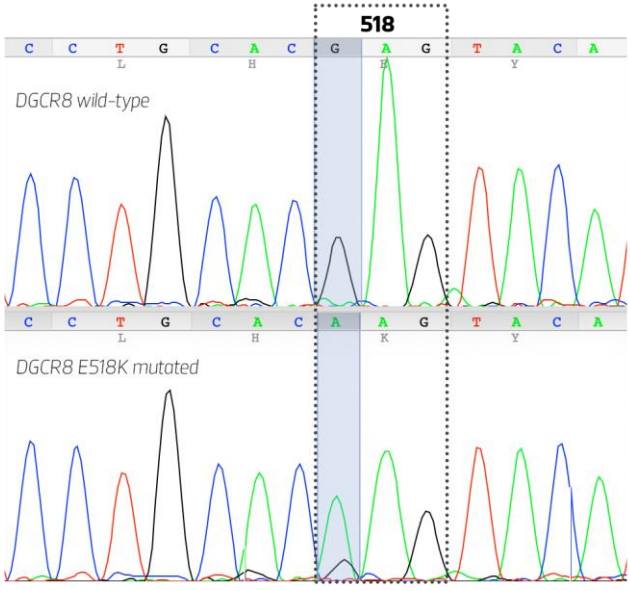
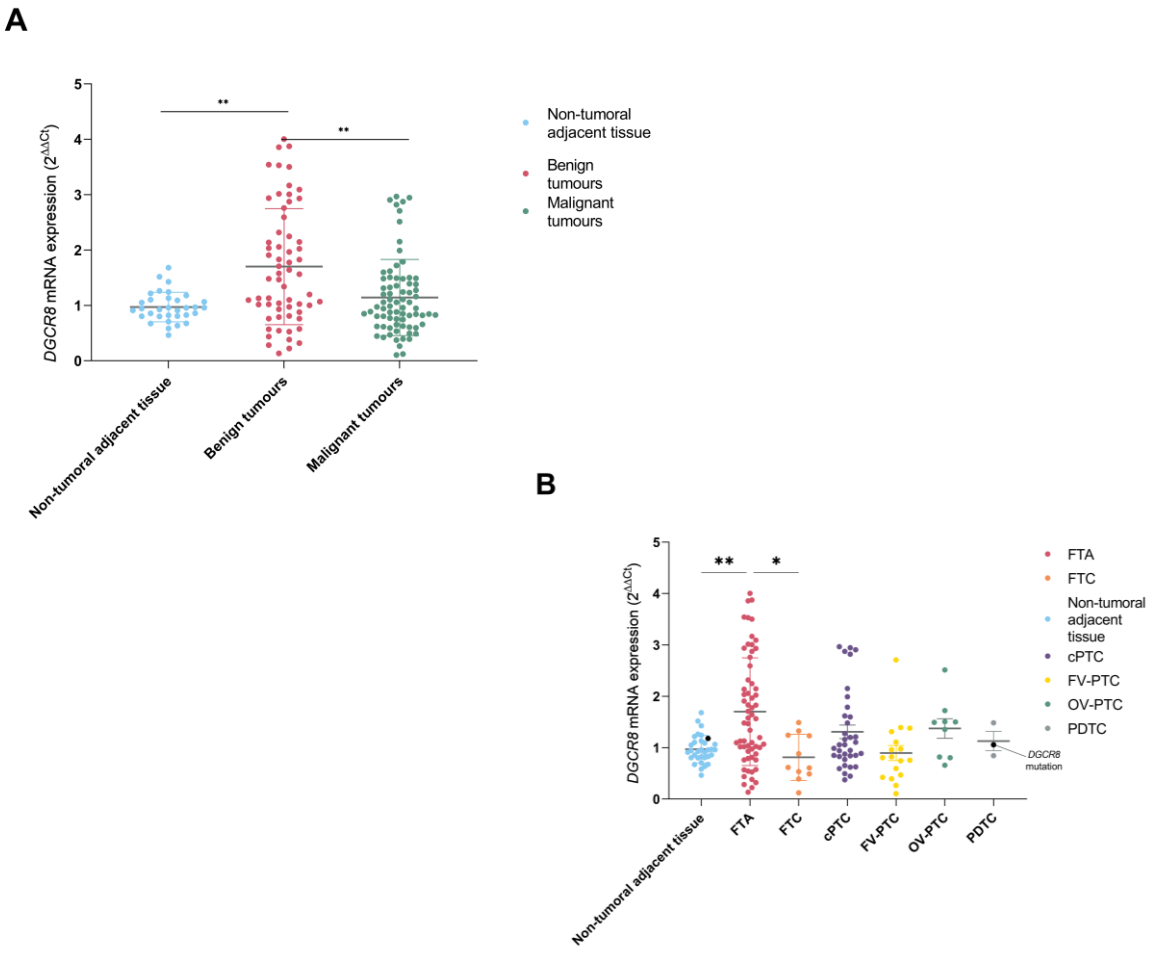



Figure 7: Chromatogram of the c.1552G>A, p.(E518K) DGCR8 mutation in a PDTC. In the shaded area is it possible to observe the guanine to adenine alteration in the amino acid 518, leading the glutamic acid (E) to be substituted by a lysine (K) in the mutated sample.

### 3.3. Deregulation of *DGCR8* mRNA expression in follicular-patterned tumours

The mRNA expression levels of *DGCR8* were available to be quantified in 170 samples by qPCR. The expression of *DGCR8* was significantly different when comparing NTAT and benign tumours (Kruskal-wallis test, p<0.01) and between benign and malignant thyroid tumours (Kruskal-wallis test, p<0.01), Figure 8A. No differences were observed between NTAT and malignant tumours. For cPTCs, FV-PTCs, OV-PTCs and PDTCs follicular cell-derived carcinomas no major differences were detected, Figure 8B. In contrast, for the remaining follicular cell-derived tumours, FTAs and FTCs, significant alterations were detected. The *DGCR8* gene quantification in FTA cases revealed that this histotype presented a higher expression than all subgroups, and most interesting significantly higher than in NTAT (Kruskal-wallis test p<0.01) and in malignant FTC cases (Kruskal-wallis test p<0.05), Figure 8B. For optimal visualization of the changes between the different histotypes, a normalization was conducted with NTAT expression being considered the basal expression and equal to 1. Following normalization, the highest fold-changes were attributed to FTAs (1.75x), followed by cPTCs (1.35x), OV-PTCs (1.42x) and PDTCs (1.16x), Figure 8C. With an inverted behaviour the follicular-pattern carcinomas, were FTCs and FV-PTCs presented a reduction 0.84x and 0.92x, respectively, Figure 8C.



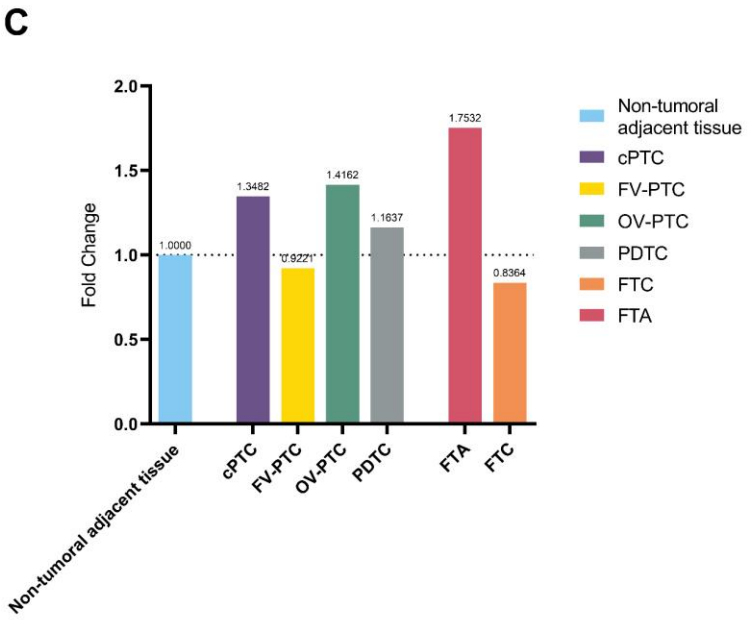



Figure 8: Expression of *DGCR8* mRNA in thyroid. A: Comparison between NTAT and benign tumours (p=0.005), as well as between benign and malignant tumours (p=0.004). B: Comparison of the expression according to the histotype, and overexpression in FTA, being significantly different from FTC (p=0.01) and NTAT (p=0.004). The *DGCR8* mutated case and the correspondent NTAT are marked with black C: Fold-change of *DGCR8* mRNA expression using NTAT as a normalizer (=1) reveals that FV-PTC and FTC are the lesions that present the higher ratios for underexpression, in contrast with the other subtypes, being cPTC and FTA the lesions with a higher gain.

* Kruskal-wallis test statistical significance p<0.05; ** Kruskal-wallis test statistical significance p<0.01

In 28 cases, the *DGCR8* gene expression between the tumour and its respective NTAT was available. The pairwise tumour/NTAT analyses revealed that in 87.5% (7 out of 8) of the cPTC cases, the main finding was overexpression of *DGCR8* in the tumours, with a statistically significant difference in expression (paired t-test, p<0.05), Figure 9A. On the other hand, in FV-PTC cases and contrarily to the previous, underexpression is the main finding in 83.3% (5 out of 6), with statistically significant differences in expression (paired t-test, p<0.05), Figure 9B.

The strong association of *DGCR8* downregulation in follicular-pattern carcinomas is also present in FTCs, where underexpression is again present in most of the cases, 66.7% (4 out 6) and with significant differences (paired t-test, p<0.05), Figure 10. A curious case in this series, was a multifocal PTC in a patient with two subtypes, a cPTC and a FV-PTC with Q61R *NRAS* mutation that in accordance with the previous findings, presented *DGCR8* overexpression and underexpression, respectively, in the different components.

In the PDTC case with p.(E518K) mutation, the expression in tumour tissue was slightly lower than in NTAT, presenting similar patterns of expression with NTAT.

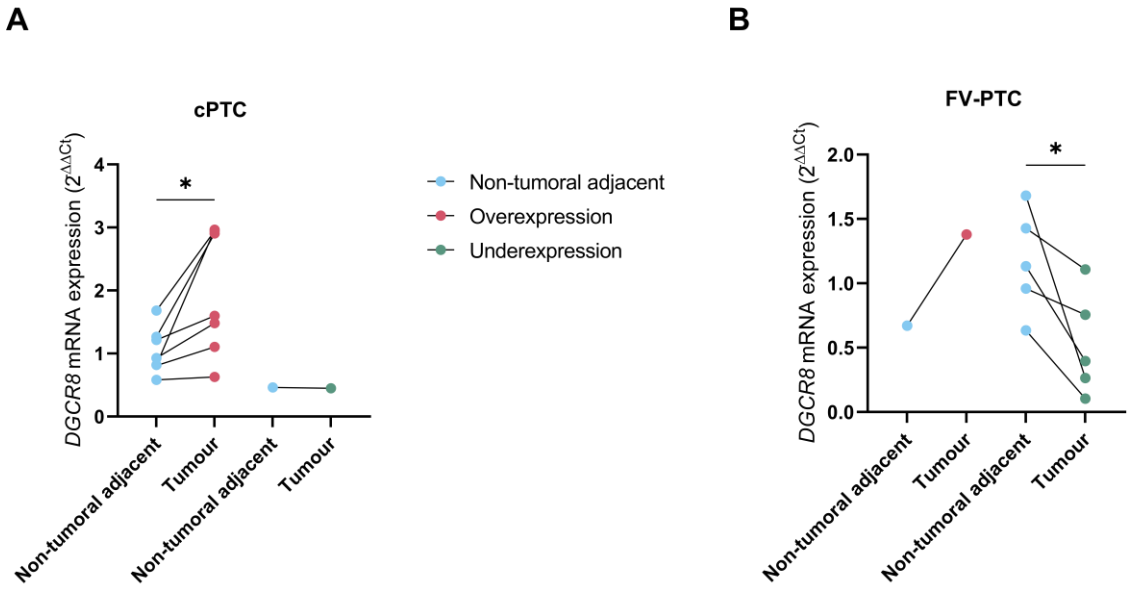



Figure 9: *DGCR8* mRNA pairwise-matched tumour/NTAT analysis in cPTC and FV-PTC. A: cPTC is characterized by overexpression and with significant difference in *DGCR8* expression (p=0.02). B: Contrarily to cPTC, in FV-PTC underexpression is more frequent and *DGCR8* loss of expression is significant between tumours/NTAT (p=0.04).

*Paired t-test statistical significance p<0.05

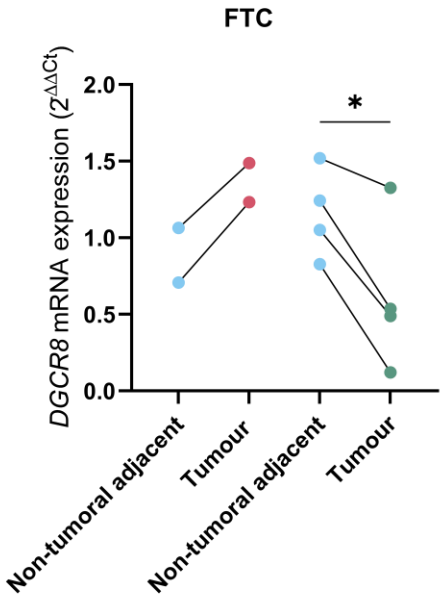



Figure 10: *DGCR8* mRNA pairwise-matched tumour/NTAT expression in FTC cases. Four out of six cases, present a significant reduction in *DGCR8* expression (p=0.02).

*Paired t-test statistical significance p<0.05

The evaluation of *DGCR8* mRNA expression in thyroid cell lines to elucidate expression patterns presented interesting results. The transformed follicular epithelial normal Nthy-ori-3-1 cell line presented the highest *DGCR8* expression. All thyroid, papillary and anaplastic carcinoma cell lines – TPC-1, T241 and 8505C, presented *DGCR8* expression lower than in human normal thyroid cell lines, Figure 11.

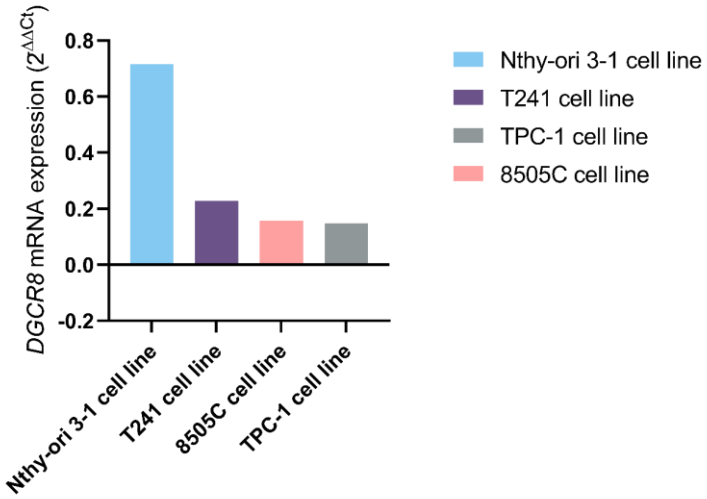



Figure 11: Expression of *DGCR8* mRNA in thyroid cell lines. A underexpression of relative *DGCR8* gene quantification is observed in thyroid tumour cell lines (T241, TPC-1 and 8505c) when compared to normal thyroid cell line (Nthy-ori 3-1).

## 3.4. DGCR8 immunoexpression in thyroid cancer

The expression of DGCR8 protein was performed in 99 FFPE samples and evaluated to create a score reflecting the staining intensity and extension, Table 2. Twenty-two cases (22.2%) were evaluated with a score of 0, however, only 4 cases lacked absolute expression for the DGCR8; the remaining 18 cases had less than 25% of extension. Regarding the other 77.8%, they were distributed throughout the additional score values as presented in Table 2 and a brief representation of part of the scores is presented in Figure 12.

Table 2: Distribution of DGCR8 protein expression score

| | *Staining score* | | | | | | |
|---|---|---|---|---|---|---|---|
| *Frequency* | *0* | *1* | *2* | *3* | *4* | *6* | *9* |
| ***n*** | 22 | 9 | 15 | 5 | 8 | 25 | 15 |
| ***%*** | 22.2 | 9.1 | 15.2 | 5.1 | 8.1 | 25.3 | 15.2 |

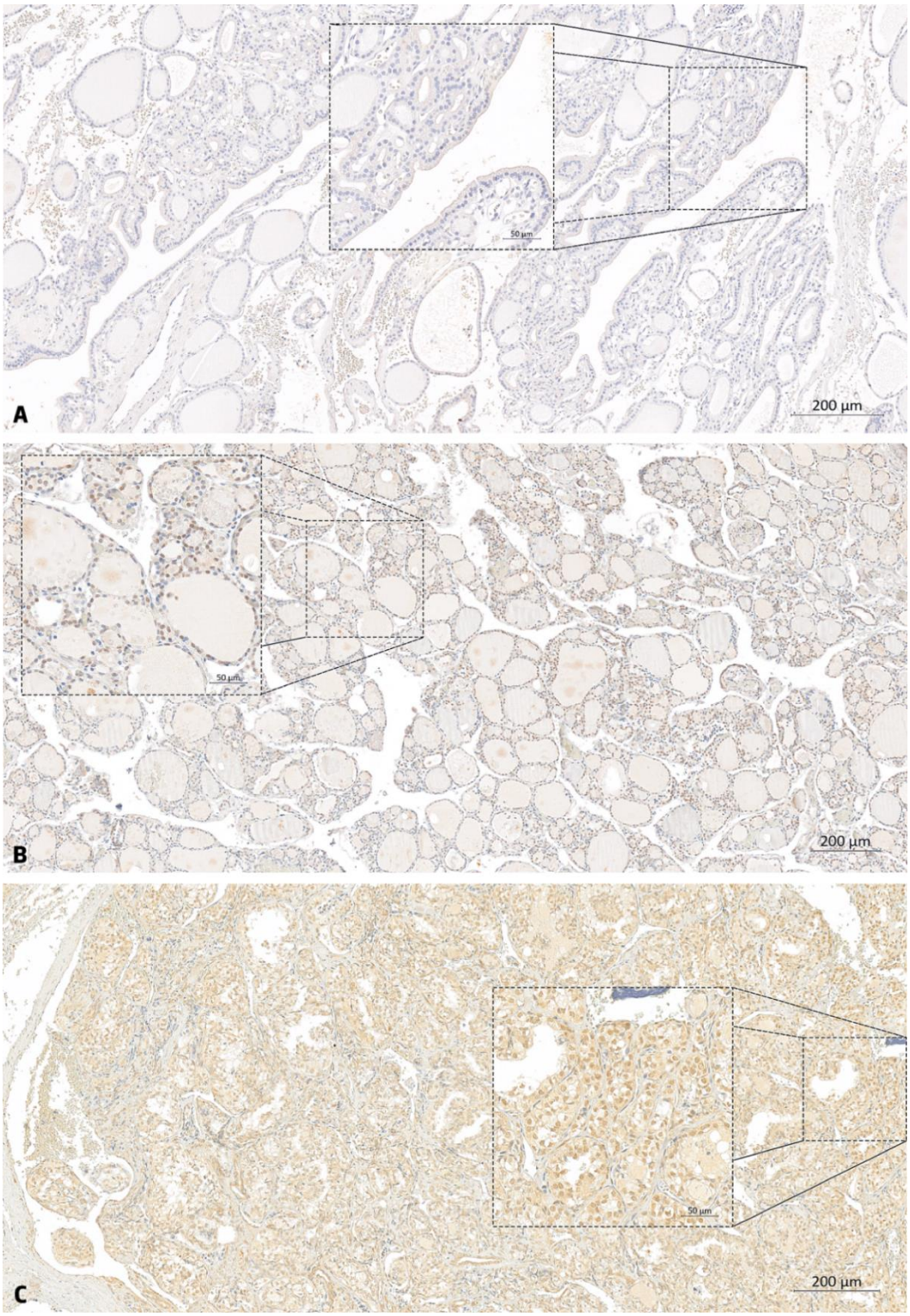



Figure 12: DGCR8 immunoexpression in thyroid lesions. A: A FV-PTC case that was evaluated with a score of 0 where no cells present DGCR8 expression. B: A FTA with an intermediate score (3), where it is observable a clear nuclear expression of DGCR8 in the thyroid follicular cells. C: A cPTC with strong nuclear and cytoplasmatic expression classified with a score of 9. All the drawn insets represent areas of magnification for better visualization of the selected areas.

As exemplified in Figure 12, the immunoexpression of DGCR8 was mainly found in the nucleus, as expected, and with an overall higher expression in lesion areas in comparison with NTAT (data not shown). Stratification by histotype of the DGCR8 scores revealed that tendentially, higher median score patterns were associated with the PTCs (with no statistical significance), Table 3 and Figure 13. The highest score of expression (9) was also commonly detected in PTCs (six cPTCs and one FV-PTC) (7 out 15 cases with a score of 9, 46.7%) and followed by FTAs (5 out 15 cases with score of 9, 33.3%). The PDTCs were the most underrepresented histotype group (n=4), and with three-quarters of the cases presenting low expression scores. Only one PDTC presented strong staining (score=9) and it corresponded to the insular variant of PDTC case with *DGCR8* p.(E518K) mutation, Figure 14; this case presented the highest DGCR8 expression of all the evaluated samples. In some extensive areas of the mutated PDTC, it was noticeable that some nuclei presented loss of DGCR8 protein, Figure 14B (black arrows).

Table 3: Score values of DGCR8 immunoexpression in the different histotypes

| | *Staining score* | | | | | | |
|---|---|---|---|---|---|---|---|
| *Histotype n, (%)* | *0* | *1* | *2* | *3* | *4* | *6* | *9* |
| ***FTA** n=30* | 6 (20.0) | 5 (16.7) | 4 (13.3) | 2 (6.7) | 2 (6.7) | 6 (20.0) | 5 (16.7) |
| ***FTC** n=15* | 3 (20.0) | 0 (0) | 2 (13.3) | 1 (6.7) | 2 (13.3) | 5 (33.3) | 2 (13.3) |
| ***PTC** n=50* | 13 (26.0) | 3 (6.0) | 7 (14.0) | 2 (4.0) | 4 (8.0) | 14 (28.0) | 7 (14.0) |
| ***PDTC** n=4* | 0 (0) | 1 (25.0) | 2 (50.0) | 0 (0) | 0 (0) | 0 (0) | 1 (25.0) |

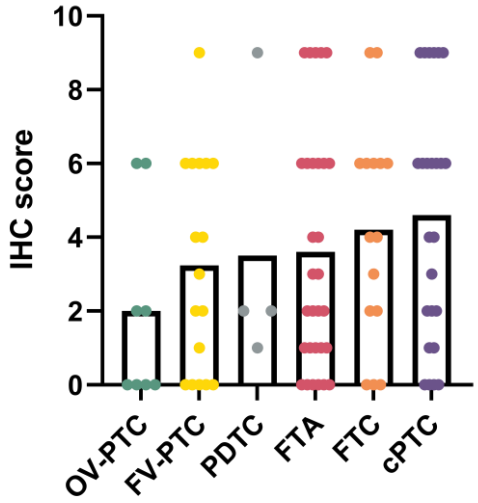



Figure 13: Distribution of the immunohistochemical expression scores according to the histotypes of the lesions.

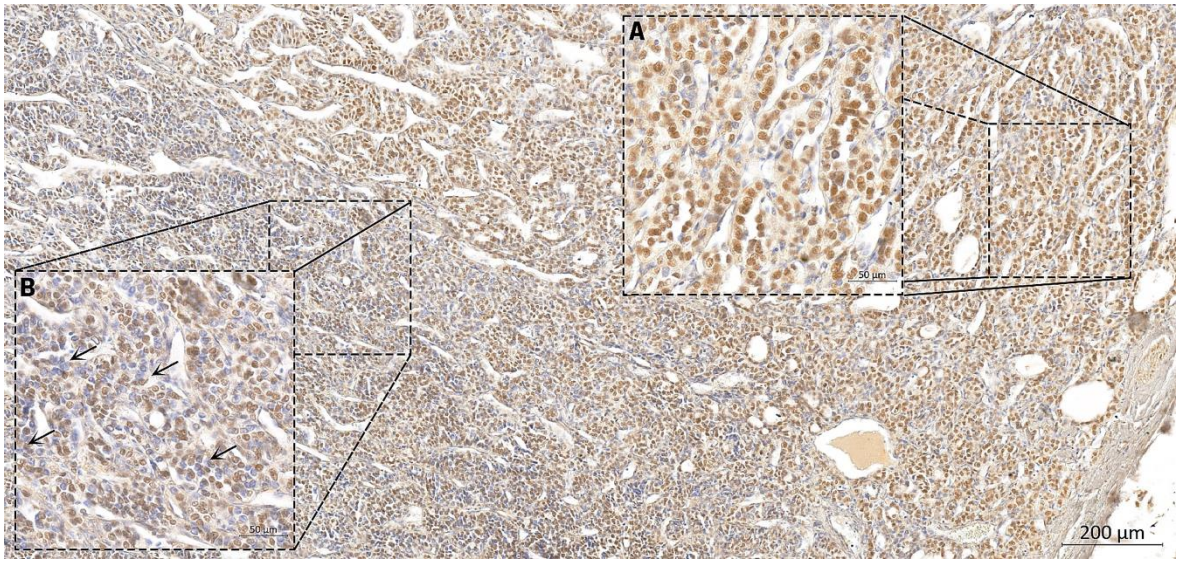



Figure 14: DGCR8 expression in PDTC case with p.(E518K) mutation. A: Overexpression of DGCR8 protein, with the highest expression in the series (9). B: Some areas presented extensive loss of expression in some nuclei (black arrows) possibly due to LOH (as previously described in cases with p.(E518K) mutation).

### 3.5. *DGCR8* mRNA and protein expression, and clinicopathological associations

*DGCR8* mRNA/protein expression was highly discordant and, overall, tendentially, it behaved contrarywise, i.e., higher levels of *DGCR8* mRNA expression associated with lower protein scores. For absent or low immunoexpression (scores 0–3), a median of 1.41 *DGCR8* mRNA expression was observed; in medium immunoexpression (scores 4–6) a median of 1.31, and in the high immunoexpression group (9) a median of 1.09 was observed; this data failed to be statistically significant (not shown).

*DGCR8* mRNA expression was additionally compared to age at diagnosis (less and more than 50 years), tumour number, vascular invasion (yes, no), lymphocytic infiltrate (yes, no), and mutation/rearrangement (presence/positive, absent/negative, respectively) but no significant associations were detected. The only significant association was verified between the *DGCR8* mRNA expression by tumour size (<2 and ≥ 2 cm), with larger tumours presenting higher expression, but the association was lost when histotype stratification was performed.

The association of the mutation presence with clinicopathological data was not performed since the low number of events (only one *DGCR8* mutated case) precluded this analysis.

## 4. Discussion

The *in silico* analysis of *DGCR8* was an initial approach to understand the gene and the protein expression in thyroid, as well as the consequences in expression upon the presence of mutations. In the combined study of thyroid, with well-differentiated and poorly-differentiated carcinomas, it was observed that in 663 samples of TC only one *DGCR8* mutation is described – c.1552G>A p.(E518K) mutation –, in two FV-PTCs (47). Remarkably, this recurrent hotspot mutation, p.(E518K), was initially described in Wilms' tumours, the most common renal paediatric malignancy, in a familial syndrome characterized with MNG and schwannomatosis (15, 40) and more recently, in a study evaluating FTCs, where two widely-invasive carcinomas were mutated (40).

The fact that our group previously demonstrated that loss of 22q was highly prevalent in wiFTC (39), that 22q is of major interest in thyroid lesions, and more recently the description of *DGCR8* (a gene located in 22q) mutation in familial MNG forms and sporadic thyroid carcinomas, presented *DGCR8* as an optimal candidate for studying it in a series of thyroid lesions.

The evaluation of *DGCR8* p.(E518K) was initially performed in familial-associated MNG patients that did not present alterations; germline DNA evaluation of 12 index, one for each family. Next, 214 samples from sporadic cases were evaluated and one case was found mutated. It corresponded to an insular variant of PDTC with the recurrent somatic missense mutation: c.1552G>A p.(E518K). This finding established the mutation frequency in 20.0% for the PDTC group and in 0.9% (1 out 109) for all the carcinomas in the series.

In the previously reported (FV-PTCs and FTCs with p.(E518K)-mutation) it was commonly detected *NRAS* mutations concomitantly; this is in accordance with follicular-patterned tumours where *NRAS* is frequently found mutated. It was advanced the *DGCR8* p.(E518K) mutation could influence the tumour progression or invasive behaviour without driving the tumour *per se*, since the tumours with this mutation are always described to bear additional genetic events (40). No other molecular alterations were detected in the mutated case (*TERTp, BRAF* and *NRAS* hotspot mutations and *RET/PTC* and *PAX8/PPARγ* rearrangements), contrasting with the previous observation, or pointing that other event (non-classical) may be involved; this is the first report, so far, of a *DGCR8*-mutated PDTC.

Mechanistically, the *DGCR8* p.(E518K) mutation is located in exon 6 and may affect DRBM1 protein domain, responsible for RNA binding potential, and to have an impact on miRNA production, leading to incomplete maturation and therefore to have a role in carcinogenesis (40). PDTC is characterized by the accumulation of several different oncogenic alterations, being important to study new biomarkers that justify the progression of follicular cells up to this poorly-differentiated carcinoma, working as predictive and prognostic markers (59). By *in* silico analysis (STRING, data not shown), one

of the *DGCR8* partners is the tumour suppressor p53 (*TP53*), a highly frequent altered gene in PDTC and ATC (60). A link between miRNA machinery components and *TP53* is described in the thyroid gland since miRNAs specifically deregulated in *DICER1*-mutant lesions of the thyroid regulate crucial pathways in cancer such as the p53 pathway (29). Results suggest that miRNAs and transcription factors, such as p53, may cooperate in complex transcriptional and post-transcriptional loops that control major cellular processes such as cell proliferation and apoptosis (21). Most cancer associated *TP53* mutations are in a domain that is required for miRNA processing and transcriptional activity. Thus, loss of p53 transcription and miRNA abnormal processing may contribute to thyroid tumour progression towards undifferentiation (21).

The expression of *DGCR8* in benign tumours was significantly different from NTAT and malignant tumours. In the findings by Paulsson *et al.* (40) a downregulation of *DGCR8* gene expression in FTCs in comparison to FTAs was reported, still, there was no data regarding the normal (or NTAT). We obtained similar results; however, we report that FTCs change in expression is only significant when compared to FTAs but not NTAT. Aberrant expression of miRNA biosynthesis genes *DGCR8* and *DROSHA* is described to promote tumorigenesis and, recent studies suggest that aberrant miRNA expressional pattern could be at play even at the level of tumour initiation (40). These results suggest that overexpression of *DGCR8*, especially in FTA (the highest *DGCR8* expression), could be at play to force maintenance of "normal" thyroid differentiation, in particular, of the follicular differentiation and structure. On the other hand, follicular-pattern carcinomas of the thyroid (FTC and FV-PTC) displayed lower gene expression than NTAT, suggesting that not only mutations but also deregulation in expression takes part in tumorigenesis of thyroid follicular-patterned carcinomas as loss of differentiation occurs. When tumoral and NTAT matched-paired study was conducted, the follicular-pattern carcinomas (FTC and FV-PTC) presented more frequently a downregulation of *DGCR8* in comparison to their normal (NTAT) counterparts. The dichotomy cPTC/high *DGCR8* expression versus FV-PTC/low *DGCR8* expression was also present in the case of a patient with multifocal PTC and with two subtypes, a cPTC and a FV-PTC that presented overexpression and underexpression, respectively. Beyond *DGCR8*, somatic *DICER1* mutations are reported in follicular-patterned lesions of thyroid (benign and malignant) which underlines the importance of the miRNA processing genes in follicular-patterned lesions (61, 62).

In the PDTC with p.(E518K) mutation, the mRNA expression in tumour tissue was slightly lower than in NTAT with comparable expression to normal tissues but with a high protein expression as evaluated by IHC, in the tumour. Contrarily to what was observed, it was described that the p.(E518K) mutation could elevate *DGCR8* mRNA expression level, probably through interrupting miRNA binding (32, 33) but we detected an expression comparable to the basal level. In the cell lines, the *DGCR8* expression in the transformed follicular normal cell line (Nthy-ori 3-1) was higher in comparison with

differentiated and undifferentiated carcinoma cell lines expression. The latter findings were not in total agreement with the observations done in our series but since the carcinoma cell lines start losing differentiation, proper thyroid tissue organization and follicle structure, DGCR8 may be impacted; it must also be taken into consideration the genomic background of the undifferentiated cell lines and its impact in *DGCR8* expression.

DGCR8 immunoprofiling was done by immunohistochemistry in 99 FFPE samples and as expected, it was mainly found in the nucleus. A tendency to more intense patterns were associated with PTC followed by FTA and FTC. PDTC cases presented low expression scores, except for PDTC case with *DGCR8* p.(E518K) mutation that presented strong staining (score=9), the highest DGCR8 expression of all the evaluated samples. In areas of the mutated PDTC, it was noticeable that some nuclei presented loss of DGCR8 protein, possibly due to LOH reported in all *DGCR8* p.(E518K) mutated cases. It would be interesting to determine if loss of the locus 22q is a second event in these cells that are losing the nuclear expression. Contrarily to *DICER*, where a positive correlation between expression at protein and RNA level is described (63), *DGCR8* mRNA/protein expression was highly discordant and, in general, it behaved contrariwise, being higher levels of *DGCR8* mRNA expression associated with lower protein scores. A perfect example was the PDTC with the mutation that had a high protein staining and low mRNA expression. This could be explained by the fact that when DROSHA and DGCR8 levels are elevated in the cell, the microprocessor cleaves and destabilizes the *DGCR8* mRNA to reduce DGCR8 levels. It is also described that the knockdown of *DROSHA* leads to upregulation of *DGCR8* expression at mRNA and protein levels, suggesting that not only alterations in *DGCR8* but also alterations in other genes involved in miRNA biogenesis could alter the DGCR8 protein expression in thyroid lesions (64). As this is the first study evaluating protein expression, further studies will help to clarify the mRNA and protein expression for DGCR8 in TC.

Correlation between *DGCR8* mRNA expression levels and clinicopathological parameters including age, histological grade, tumour size and lymph node metastasis presented significant results in previous studies (64). However, in our study no relationship between the gene expression and clinicopathological features were revealed, unless for *DGCR8* mRNA expression by tumour size (<2 and ≥ 2 cm), with larger tumours presenting higher expression but the association is lost when histotype stratification was performed.

As miRNAs are important to stabilize thyroid follicles and hormone production, it is perceivable that alterations in genes involved in miRNA biogenesis may have repercussions at follicular level, having a role in tumorigenesis of follicular patterned tumours.

The findings from this study strengthen the association between abnormal miRNA processing and the development and/or progression of TC, especially in follicular-patterned carcinomas, suggesting that not only mutations but also alterations in *DGCR8* mRNA/protein expression may be

important in thyroid tumorigenesis. Succeeding DICER1 alterations in the susceptibility of thyroid disease, we reaffirm DGCR8 as a novel player of the miRNA microprocessor complex team.

As future perspectives, it would be interesting to study the loss of 22q in the mutated PDTC, as in other p.(E518K)-mutated cases it is described loss of the entire 22 chromosome. This would also clarify if the loss of expression observed in some nuclei were due to LOH. It will be also important to extend the characterization of the 22q and determine if LOH or other mechanisms (e.g., methylation) is responsible for the changes we observed in expression. Additional molecular characterization would be also useful to determine if other events cooperated with *DGCR8* mutation, in particular *TP53* loss, since it is a common alteration in PDTCs and it is important in TC progression. Finally, but not less important, once we are dealing with a miRNA microprocessor complex, miRNAs expression changes and miRNA patterns of DGCR8-altered lesions could bring interesting results.

## 5. Conclusion

The study of the miRNA biogenesis gene *DGCR8* in this series of thyroid lesions allowed for the first time the identification of the hotspot *DGCR8* p.(E518K) in a PDTC case. The *DGCR8* mRNA expression revealed a deregulation in follicular-pattern lesions of thyroid. Overexpression of *DGCR8* mRNA in FTA cases and an underexpression in follicular pattern carcinomas (FTC and FV-PTC) was observed and these results are in agreement with a pairwise study between tumoral and NTAT.

At the protein level, DGCR8 expression showed a higher score of expression in PTC and FTA cases, being observed an inverse tendency between RNA and protein expression. The PDTC p.(E518K) mutated stained at the highest score, and in some areas some nuclei were detected with loss of expression, possibly due to LOH, as reported in the literature in cases with this mutation.

Our findings strengthen the association between miRNAs processing machinery interference in the development and progression of TC, especially in follicular-pattern carcinomas, suggesting that not only mutations but also alterations in *DGCR8* mRNA/protein expression could have an impact on thyroid tumorigenesis.

## References


1. Guo WT, Wang Y. Dgcr8 knockout approaches to understand microRNA functions in vitro and in vivo. Cell Mol Life Sci. 2019;76(9):1697–711.
2. Nikiforova MN, Chiosea SI, Nikiforov YE. MicroRNA expression profiles in thyroid tumors. Endocr Pathol. 2009;20(2):85–91.
3. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993;75(5):843–54.
4. Yi R, Pasolli HA, Landthaler M, Hafner M, Ojo T, Sheridan R, et al. DGCR8-dependent microRNA biogenesis is essential for skin development. Proc Natl Acad Sci U S A. 2009;106(2):498–502.
5. Shiohama A, Sasaki T, Noda S, Minoshima S, Shimizu N. Nucleolar localization of DGCR8 and identification of eleven DGCR8-associated proteins. Exp Cell Res. 2007;313(20):4196–207.
6. Fuziwara CS, Kimura ET. MicroRNAs in thyroid development, function and tumorigenesis. Mol Cell Endocrinol. 2017;456:44–50.
7. Nikiforova MN, Tseng GC, Steward D, Diorio D, Nikiforov YE. MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility. J Clin Endocrinol Metab. 2008;93(5):1600–8.
8. Park JL, Kim SK, Jeon S, Jung CK, Kim YS. MicroRNA Profile for Diagnostic and Prognostic Biomarkers in Thyroid Cancer. Cancers (Basel). 2021;13(4).
9. Santiago K, Chen Wongworawat Y, Khan S. Differential MicroRNA-Signatures in Thyroid Cancer Subtypes. J Oncol. 2020;2020:2052396.
10. Leonardi GC, Candido S, Carbone M, Colaianni V, Garozzo SF, Cina D, et al. microRNAs and thyroid cancer: biological and clinical significance (Review). Int J Mol Med. 2012;30(5):991–9.
11. Ouellet DL, Perron MP, Gobeil LA, Plante P, Provost P. MicroRNAs in gene regulation: when the smallest governs it all. J Biomed Biotechnol. 2006;2006(4):69616.
12. Wen J, Lv Z, Ding H, Fang X, Sun M. Association of miRNA biosynthesis genes DROSHA and DGCR8 polymorphisms with cancer susceptibility: a systematic review and meta-analysis. Bioscience reports. 2018;38(3).
13. Marini F, Luzi E, Brandi ML. MicroRNA Role in Thyroid Cancer Development. J Thyroid Res. 2011;2011:407123.
14. Macias S, Cordiner RA, Caceres JF. Cellular functions of the microprocessor. Biochem Soc Trans. 2013;41(4):838–43.
15. Rivera B, Nadaf J, Fahiminiya S, Apellaniz-Ruiz M, Saskin A, Chong AS, et al. DGCR8 microprocessor defect characterizes familial multinodular goiter with schwannomatosis. J Clin Invest. 2020;130(3):1479–90.

16. Robertson JC, Jorcyk CL, Oxford JT. DICER1 Syndrome: DICER1 Mutations in Rare Cancers. Cancers (Basel). 2018;10(5).

17. Boufraqech M, Klubo-Gwiezdzinska J, Kebebew E. MicroRNAs in the thyroid. Best Pract Res Clin Endocrinol Metab. 2016;30(5):603-19.

18. Slade I, Bacchelli C, Davies H, Murray A, Abbaszadeh F, Hanks S, et al. DICER1 syndrome: clarifying the diagnosis, clinical features and management implications of a pleiotropic tumour predisposition syndrome. J Med Genet. 2011;48(4):273-8.

19. Guo F. Drosha and DGCR8 in MicroRNA Biogenesis. Eukaryotic RNases and their Partners in RNA Degradation and Biogenesis, Part B. The Enzymes2012. p. 101-21.

20. Forte S, Pecce V, Rosignolo F, Memeo L. The role of micrornas in thyroid carcinomas. 2015.

21. Markopoulos GS, Roupakia E, Tokamani M, Chavdoula E, Hatziapostolou M, Polytarchou C, et al. A step-by-step microRNA guide to cancer development and metastasis. Cell Oncol (Dordr). 2017;40(4):303-39.

22. Visone R, Pallante P, Vecchione A, Cirombella R, Ferracin M, Ferraro A, et al. Specific microRNAs are downregulated in human thyroid anaplastic carcinomas. Oncogene. 2007;26(54):7590-5.

23. Perdas E, Stawski R, Nowak D, Zubrzycka M. The Role of miRNA in Papillary Thyroid Cancer in the Context of miRNA Let-7 Family. Int J Mol Sci. 2016;17(6).

24. He H, Jazdzewski K, Li W, Liyanarachchi S, Nagy R, Volinia S, et al. The role of microRNA genes in papillary thyroid carcinoma. Proc Natl Acad Sci U S A. 2005;102(52):19075-80.

25. Kim S. The dysregulation of microrna machinery components in papillary thyroid cancer.

26. Li X, Abdel-Mageed AB, Mondal D, Kandil E. MicroRNA expression profiles in differentiated thyroid cancer, a review. Int J Clin Exp Med. 2013;6(1):74-80.

27. Vriens MR, Weng J, Suh I, Huynh N, Guerrero MA, Shen WT, et al. MicroRNA expression profiling is a potential diagnostic tool for thyroid cancer. Cancer. 2012;118(13):3426-32.

28. Hata A. Dysregulation of microRNA biogenesis machinery in cancer. Critical Reviews in Biochemistry and Molecular Biology. 2016(51(3)):121-34.

29. Poma AM, Condello V, Denaro M, Torregrossa L, Elisei R, Vitti P, et al. DICER1 somatic mutations strongly impair miRNA processing even in benign thyroid lesions. Oncotarget. 2019;10(19):1785-97.

30. Schultze-Florey R, Graf N, Vorwerk P, Koscielniak E, Schneider D, Kratz C. DICER1 syndrome: a new cancer syndrome. Klin Padiatr. 2013;225(3):177-8.

31. Schultz KAP, Rednam SP, Kamihara J, Doros L, Achatz MI, Wasserman JD, et al. PTEN, DICER1, FH, and their associated tumor susceptibility syndromes: clinical features, genetics, and surveillance recommendations in childhood. Clinical Cancer Research. 2017;23(12):e76-e82.

32. Wen J, Lv Z, Ding H, Fang X, Sun M. Association of miRNA biosynthesis genes DROSHA and DGCR8 polymorphisms with cancer susceptibility: a systematic review and meta-analysis. Biosci Rep. 2018;38(3).

33. Puppin C, Durante C, Sponziello M, Verrienti A, Pecce V, Lavarone E, et al. Overexpression of genes involved in miRNA biogenesis in medullary thyroid carcinomas with RET mutation. Endocrine. 2014;47(2):528-36.

34. Fardmanesh H, Shekari M, Movafagh A, Alizadeh Shargh S, Poursadegh Zonouzi AA, Shakerizadeh S, et al. Upregulation of the double-stranded RNA binding protein DGCR8 in invasive ductal breast carcinoma. Gene. 2016;581(2):146-51.

35. Nguyen TA, Jo MH, Choi YG, Park J, Kwon SC, Hohng S, et al. Functional Anatomy of the Human Microprocessor. Cell. 2015;161(6):1374-87.

36. Kim J, Park WJ, Jeong KJ, Kang SH, Kwon SY, Kim S, et al. Racial Differences in Expression Levels of miRNA Machinery-Related Genes, Dicer, Drosha, DGCR8, and AGO2, in Asian Korean Papillary Thyroid Carcinoma and Comparative Validation Using the Cancer Genome Atlas. Int J Genomics. 2017;2017:5789769.

37. Kim B, Lee JH, Park JW, Kwon TK, Baek SK, Hwang I, et al. An essential microRNA maturing microprocessor complex component DGCR8 is up-regulated in colorectal carcinomas. Clin Exp Med. 2014;14(3):331-6.

38. Sellier C, Hwang VJ, Dandekar R, Durbin-Johnson B, Charlet-Berguerand N, Ander BP, et al. Decreased DGCR8 expression and miRNA dysregulation in individuals with 22q11.2 deletion syndrome. PLoS One. 2014;9(8):e103884.

39. Castro P, Eknæs M, Teixeira MR, Danielsen HE, Soares P, Lothe RA, et al. Adenomas and follicular carcinomas of the thyroid display two major patterns of chromosomal changes. The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland. 2005;206(3):305-11.

40. Paulsson JO, Rafati N, DiLorenzo S, Chen Y, Haglund F, Zedenius J, et al. Whole-genome sequencing of follicular thyroid carcinomas reveal recurrent mutations in microRNA processing subunit DGCR8. J Clin Endocrinol Metab. 2021.

41. Walz AL, Ooms A, Gadd S, Gerhard DS, Smith MA, Guidry Auvil JM, et al. Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. Cancer Cell. 2015;27(2):286-97.

42. UniProtKB - Q8WYQ5 (DGCR8_HUMAN) [Online] Available from: https://www.uniprot.org/dataset/identifier.

43. Bulk tissue gene expression for DGCR8 GTEx Portal [Online] Available from: https://www.gtexportal.org/home/gene/ENSG00000128191.

44. The Human Proten Atlas: DGCR8 [Online] Available from: https://www.proteinatlas.org/ENSG00000128191-DGCR8/tissue.

45. OncoMX [Online] Available from: https://www.oncomx.org/#hero.

46. BioMuta: Q8WYQ5: BioMuta; [Online] Available from: https://hive.biochemistry.gwu.edu/biomuta/proteinview/Q8WYQ5.

47. cBioPortal for Cancer Genomics [Online] Available from: www.cbioportal.org.

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

49. Landa I, Ibrahimpasic T, Boucai L, Sinha R, Knauf JA, Shah RH, et al. Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. J Clin Invest. 2016;126(3):1052-66.

50. Lloyd R, Osamura R, Klöppel G, Rosai J. WHO Classification of Tumours of Endocrine Organs, 4th edn, vol. 10. IARC Press, Lyon; 2017.

51. Pestana A, Batista R, Celestino R, Canberk S, Sobrinho-Simoes M, Soares P. Comprehensive Assessment of TERT mRNA Expression across a Large Cohort of Benign and Malignant Thyroid Tumours. Cancers (Basel). 2020;12(7).

52. Cheng Q, Zhang X, Xu X, Lu X. MiR-618 inhibits anaplastic thyroid cancer by repressing XIAP in one ATC cell line. Ann Endocrinol (Paris). 2014;75(4):187-93.

53. Landa I. Comprehensive Genetic Characterization of Human Thyroid Cancer Cell Lines: A Validated Panel for Preclinical Studies. Translational Cancer Mechanisms and Therapy. 2019.

54. Ribeiro FR, Meireles AM, Rocha AS, Teixeira MR. Conventional and molecular cytogenetics of human non-medullary thyroid carcinoma: characterization of eight cell line models and review of the literature on clinical samples. BMC Cancer. 2008;8:371.

55. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-8.

56. Gene DGCR8: NCBI; [Online] Available from: https://www.ncbi.nlm.nih.gov/gene/54487#gene-expression.

57. SMART: Domains within Homo sapiens protein DGCR8_HUMAN (Q8WYQ5) [Online] Available from: https://smart.embl.de/smart/show_motifs.pl?ID=Q8WYQ5.

58. Soares P, Trovisco V, Rocha AS, Lima J, Castro P, Preto A, et al. BRAF mutations and RET/PTC rearrangements are alternative events in the etiopathogenesis of PTC. Oncogene. 2003;22(29):4578-80.

59. Takano T. Fetal cell carcinogenesis of the thyroid: theory and practice. Semin Cancer Biol. 2007;17(3):233-40.

60. Soares P, Lima J, Preto A, Castro P, Vinagre J, Celestino R, et al. Genetic alterations in poorly differentiated and undifferentiated thyroid carcinomas. Current genomics. 2011;12(8):609-17.

61. Wasserman JD, Sabbaghian N, Fahiminiya S, Chami R, Mete O, Acker M, et al. DICER1 Mutations Are Frequent in Adolescent-Onset Papillary Thyroid Carcinoma. The Journal of clinical endocrinology and metabolism. 2018;103(5):2009-15.

62. Canberk S, Ferreira JC, Pereira L, Batista R, Vieira AF, Soares P, et al. Analyzing the Role of DICER1 Germline Variations in Papillary Thyroid Carcinoma. Eur Thyroid J. 2021;9(6):296-303.

63. Ma Z, Swede H, Cassarino D, Fleming E, Fire A, Dadras SS. Up-regulated Dicer expression in patients with cutaneous melanoma. PLoS One. 2011;6(6):e20494.

64. Triboulet R, Chang HM, Lapierre RJ, Gregory RI. Post-transcriptional control of DGCR8 expression by the Microprocessor. RNA. 2009;15(6):1005-11.

# APPENDIX

## Appendix 1

CES 284-13


**Exmo. Senhor**
**Presidente do Conselho de Administração do**
**Centro Hospitalar de S. João – EPE**

**Assunto:** Pedido de autorização para realização de estudo/projecto de investigação

**Nome do Investigador Principal:** Paula Soares

**Título do projecto de investigação:** Avaliação de alterações genéticas e de expressão proteica em tumores da tiroide

Pretendendo realizar nos Serviços de Anatomia Patológica do Centro Hospitalar de S. João – EPE o estudo/projecto de investigação em epígrafe, solicito a V. Exa., na qualidade de Investigador/Promotor, autorização para a sua efectivação.
Para o efeito, anexa toda a documentação referida no dossier da Comissão de Ética do Centro Hospitalar de S. João respeitante a estudos/projectos de investigação, à qual endereçou pedido de apreciação e parecer.

Com os melhores cumprimentos.

Porto, _9___ / _____Setembro / 2013__

O INVESTIGADOR/PROMOTOR

Paula Soares

Comissão de Ética para a Saúde do Centro Hospitalar de S. João – EPE
Modelo CES 01

SÃO JOÃO

**Comissão de Ética para a Saúde – Centro Hospitalar São João**

**Parecer**

**Título do Projecto**: Avaliação de alterações genéticas e de expressão proteica em tumores da tireóide.

**Nome do Investigador Principal**: Profª Doutora Paula Soares

**Local onde sera realizado o estudo**: Serviço de Anatomia Patológica do Centro Hospitalar de São João, havendo autorização da respectiva Directora de Serviço. Apresenta como Elo de Ligação, o Dr. João Magalhães.

**Objectivo do estudo**:
Caracterizar as alterações genéticas de tumores da tireóide e avaliar qual a sua contribuição no desenvolvimento, progressão e prognóstico da doença oncológica da tireóide.

**Período previsto de conclusão**: 2018.

**Benefício**: N/A

**Riscos:** N/A



CENTRO HOSPITALAR DE **SÃO JOÃO** E.P.E
**PORTO** (Sede) Alameda Professor Hernâni Monteiro 4200-319 Porto T • 351 225 512 100 geral@hsjoao.min-saude.pt
**VALONGO** Rua de Misericórdia 4440-563 Valongo T • 351 224 220 019 www.hsjoao.min-saude.pt

SÃO JOÃO

**Respeito pela liberdade e autonomia do sujeito do ensaio**: Consentimento informado do Banco de Tecidos e Tumores do Centro Hospitalar de São João.

**Confidencialidade dos dados**: está garantida a confidencialidade dos dados e esta informação será restrita à investigadora principal.

Não está previsto a realização de questionários.

A Investigadora Principal dispõe de competência técnica e científica para a realização do estudo.

**Custos**: O estudo não prevê custos acrescidos para o CHSJ. O estudo será financiado pelo IPATIMUP.

**Parecer**: Em face da análise do protocolo de estudo, proponho a sua aprovação pela CES do CHSJ.

Porto, CHSJ, 22 de outubro de 2013

O Relator

Dr. John Preto



CENTRO HOSPITALAR DE **SÃO JOÃO** E.P.E.
**PORTO** (Sede) Alameda Professor Hernâni Monteiro 4200-319 Porto T + 351 225 512 100 geral@hsjoao.min-saude.pt
**VALONGO** Rua de Misericórdia 4440-563 Valongo T + 351 224 220 019 www.hsjoao.min-saude.pt