Biomarkers of Aggressive Thyroid Cancer

By

George Jiajie Xu

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Molecular Pathology and
Immunology

December 16th, 2023

Nashville, Tennessee

Approved by:

Jeff Rathmell, PhD

Vivian Weiss, MD, PhD

Ethan Lee, MD, PhD

Fei Ye, PhD, MSPH

Thomas P. Stricker, MD, PhD

ACKNOWLEDGEMENTS

cancer pathology and surgical oncology was indispensable for our project, and we are so grateful to have you as our guide.

On a personal note, I would like to thank Jeff Jian, Sabrina van Ravenstein, Baltazar Zuniga, Mirazul Islam, and Gabriel Garcia. Your friendships have been the highlights of my graduate school experience, and I will always cherish the time and memories we shared in Nashville together. I would also like to thank Elaine Do; I am so grateful for the last four years we've known each other, and your love and kindness constantly inspires me to be the best I can be. I would also like to thank my parents, Xilin and Guangdou, as well as my siblings Linda, Sarah, and David; since the beginning, you all have been my greatest supporters, and I know I wouldn't have gotten this far without you.

TABLE OF CONTENTS

Chapter

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

ATC	Anaplastic Thyroid Carcinoma
BRS	*BRAF-RAS* Score
CAF	Cancer-associated Fibroblast
ECM	Extracellular Matrix
EFVPTC	Encapsulated Follicular Variant Papillary Thyroid Carcinoma
EMT	Epithelial-Mesenchymal Transition
FA	Follicular Adenoma
FFPE	Formalin-fixed Paraffin-embedded
FTC	Follicular Thyroid Carcinoma
HT	Hashimoto Thyroiditis
IFVPTC	Infiltrative Follicular Variant Papillary Thyroid Carcinoma
MAP Score	Molecular Aggression and Prediction Score
MNG	Multinodular Goiter
NIFTP	Noninvasive Follicular Thyroid Neoplasm with Papillary-like Nuclear Features
OA	Oncocytic Adenoma
OTC	Oncocytic Thyroid Carcinoma
PDTC	Poorly Differentiated Thyroid Carcinoma
PTC	Papillary Thyroid Carcinoma
TCGA	The Cancer Genome Atlas

CHAPTER 1

Title of Chapter: Background and Research Objectives

1.1 Thyroid Cancer

1.1.1 Prevalence

Thyroid cancer is extremely common, with some estimates projecting thyroid cancer to be the 4th leading cancer by diagnosis in 2030.[1] In the United States alone, there are about 50,000 new thyroid cancer patients diagnosed each year.[2] Thyroid cancer also disproportionately affects women, with rates of diagnosis approximately three times higher in women than in men. While thyroid cancer is common, it is also typically indolent, meaning that the disease does not cause significant morbidity, and causes practically no mortality. Most tumors can be successfully treated with surgical resection of the thyroid followed by radioactive iodine.[3,4] However, thyroid cancer on average appears earlier than most other solid tumors; for example, thyroid cancer is currently the second most common type of cancer in adolescents and young adults (ages 15–39).[5] In addition, thyroid cancer can recur aggressively many years after initial diagnosis and treatment. The median age of new cases is 51 years, while the median age of death from thyroid cancer is 74 years.[5] This 23-year difference is much longer than that of most other cancer types and provides a wide window of time for potential recurrence. With increased time and age also comes increased risk of distant metastasis or transformation to poorly differentiated thyroid cancer (PDTC) or anaplastic thyroid cancer (ATC)[6], the later diagnosis having a median survival of just 3-5 months.[7] The current lack of molecular testing to predict future aggressive thyroid cancer results in challenges in patient management, such as determining who should seek radioactive iodine treatment. Without a clear understanding of how advanced disease develops, many people continue to get thyroid cancer recurrence or die from the disease.

1.1.2 Molecular Alterations

Most thyroid cancers can be classified by common alterations in the mitogen-activated protein kinase (MAPK) pathway and the PI3K-AKT-mTOR (PI3K) pathway (Figure 1-1). These pathways play important roles in cell proliferation and survival, and abnormal upregulation of these pathways is important in many cancers. The MAPK pathway includes RAS, which is commonly mutated in follicular thyroid cancer subtypes, as well as BRAF, which is downstream in this pathway and is characteristic of papillary thyroid carcinoma (PTC).[8,9] Mutations in *RAS* and *BRAF* are mutually exclusive. *BRAF* mutations have been reported to be in around 45% of PTCs, and even more frequently in more aggressive PTC subtypes.[10] The most common *BRAF* mutation in thyroid cancer, as well as several other cancers, is the *BRAF* V600E point mutation. This mutation, accounting for up to 99% of all thyroid cancer *BRAF* mutations, results in constitutive activation of BRAF activity, driving signaling that enhances cell proliferation and supports cancer progression.[8,11] Both *RAS* and *BRAF* mutations can also be found in de-differentiated thyroid cancers such as PDTC and ATC.[8] In the PI3K pathway, a common thyroid cancer alteration is the *PAX8/PPARG* fusion, which is seen most commonly in follicular thyroid cancers.[12] Less common alterations in the PI3K pathway included those affecting PI3K and PTEN.[12]
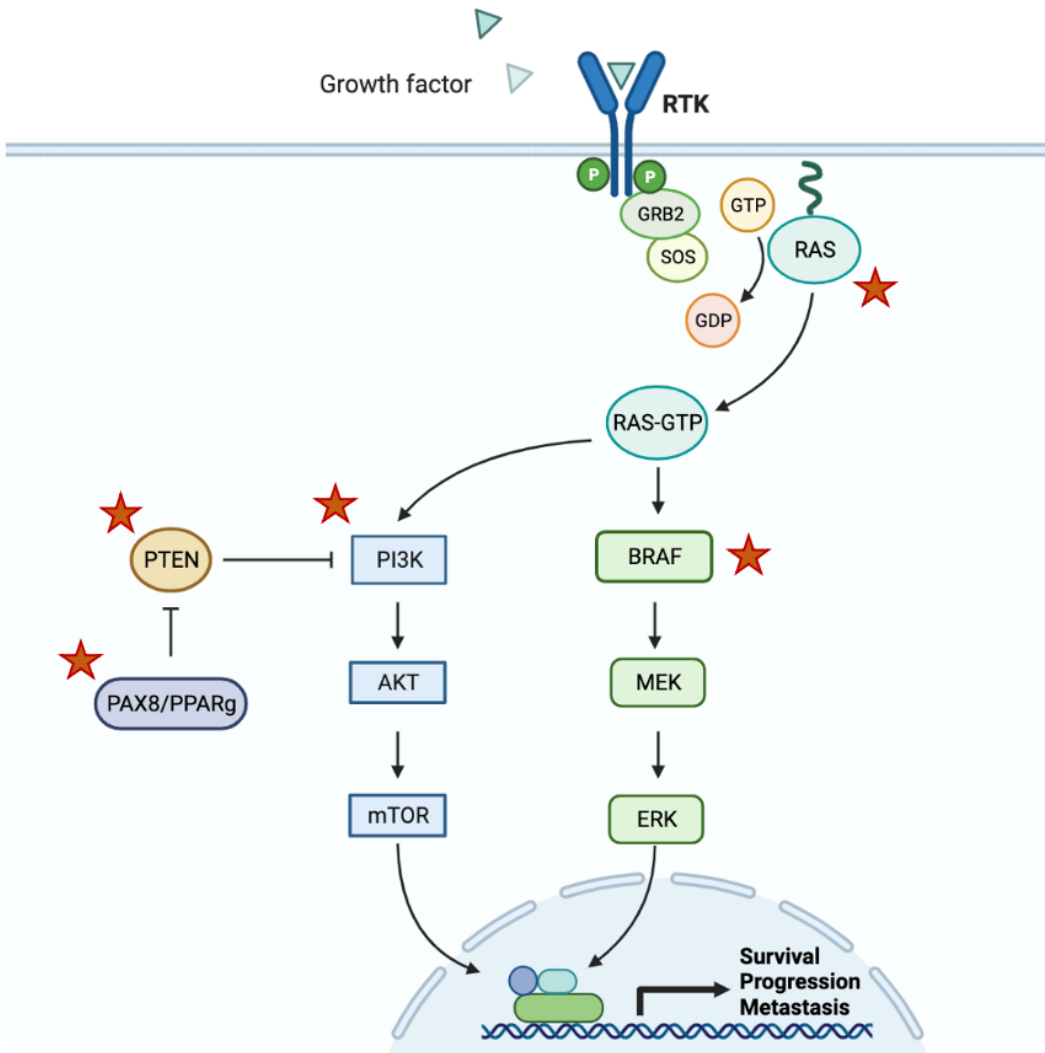
Figure 1-1: Common thyroid cancer mutations in the MAPK and PI3K pathway
Summary of the MAPK and PI3K pathways, with proteins commonly altered in thyroid cancer marked by red stars. This figure was generated using BioRender.

Mutations associated with aggressive and de-differentiated thyroid cancer include alterations in the genes *TP53*, *PIK3CA*, and the *TERT* promoter.[13] *TP53* is a well-established tumor suppressor that plays a critical role in functions such as arresting the cell cycle and initiating DNA repair when damage is detected, initiating senescence in response to low telomere length, and initiating apoptosis if DNA damage is unable to be repaired.[14] While mutations have been recorded in all coding exons of *TP53*, they are most commonly found in exons 4-9, in the DNA-binding domain; about 30% of these mutations occur within 6 "hotspot" residues: R175, G245, R248, R249, R273, and R282.[15] DNA-binding domain mutations in *TP53* promote cancer by destroying the ability of the p53 protein to bind to its target DNA sequences, thus preventing transcriptional activation of genes involved in tumor suppression.

PIK3CA is another gene commonly mutated in many cancers, and codes for the catalytic subunit of PI3K. As part of the PI3K pathway, PI3K is involved in regulating a wide range of cellular processes including cell cycle progression, proliferation, adhesion, migration, survival, and differentiation.[16] Of the common thyroid cancer alterations involving the PI3K pathway, alterations in *PIK3CA* are the most common for transformed subtypes PDTC and ATC, with some studies reporting *PIK3CA* mutations in 10-20% of these tumors.[8] Alterations in *PIK3CA* include mutations affecting the kinase domain in exon 20, as well as copy number gains; these alterations activate the activity of PI3K, supporting tumor growth and progression.[17,18]

TERT codes for telomerase reverse transcriptase, the catalytic subunit of telomerase. Telomerase functions to lengthen telomeres, protecting the ends of DNA strands from the shortening that occurs with each division. Two common *TERT* promoter mutations are C228T and C250T, which generate a consensus binding site in the promoter for E-twenty-six (ETS) transcription factors, conferring the *TERT* promoter with increased transcriptional activities.[19] These two mutations have been reported to occur in around 40% of PDTC and ATC tumors.[19] *TERT* promoter mutations have been thought to support cancer by upregulating telomerase activity, allowing cells to achieve replicative immortality. However, studies have also shown that

telomeres are shorter in thyroid cancer cells with high telomerase activity,[20,21] implying other roles of these mutations, such as modulating changes in growth, angiogenesis, metastasis, inflammation, and immune surveillance.[22]

      To summarize how some common mutations fit into the evolution of a thyroid tumor, we start with a normal thyroid, composed largely of thyroid follicular cells organized into macrofollicles. These cells function to produce and secrete the thyroid hormones thyroxine (T4) and triiodothyronine (T3). During tumorigenesis, thyroid follicular cells, or thyrocytes, commonly acquire either a *RAS* or *BRAF* driver mutation (Figure 1-2A). This initial driver mutation can lead to the development of a differentiated thyroid tumor, such as follicular thyroid cancer (FTC), which is associated with *RAS* mutations, or PTC, which is associated with *BRAF* mutations. PTC is the most common form of thyroid cancer and accounts for between 80-85% of all thyroid cancers.[23] Well-differentiated tumors may then undergo a secondary hit in a gene such as *TP53* or in the *TERT* promoter, which may lead the thyroid tumor to undergo transformation to a de-differentiated cancer, such as PDTC or ATC.[8,24] This progression from normal cells to tumor and then transformed tumor is based on the multistep model of colorectal tumorigenesis postulated by E.F. Fearon and B. Vogelstein, in which cancer arises and advances towards increasingly advanced disease through a series of accumulated genetic alterations.[25] Transformation in thyroid cancer involves further loss of thyroid histologic organization, as well as loss of expression of thyroid-specific proteins such as thyroid peroxidase and thyroglobulin.[26] A complete summary of all 12 thyroid lesion subtypes is shown in Figure 1-2B, showing proposed paths of development towards increasingly aggressive disease from left to right.
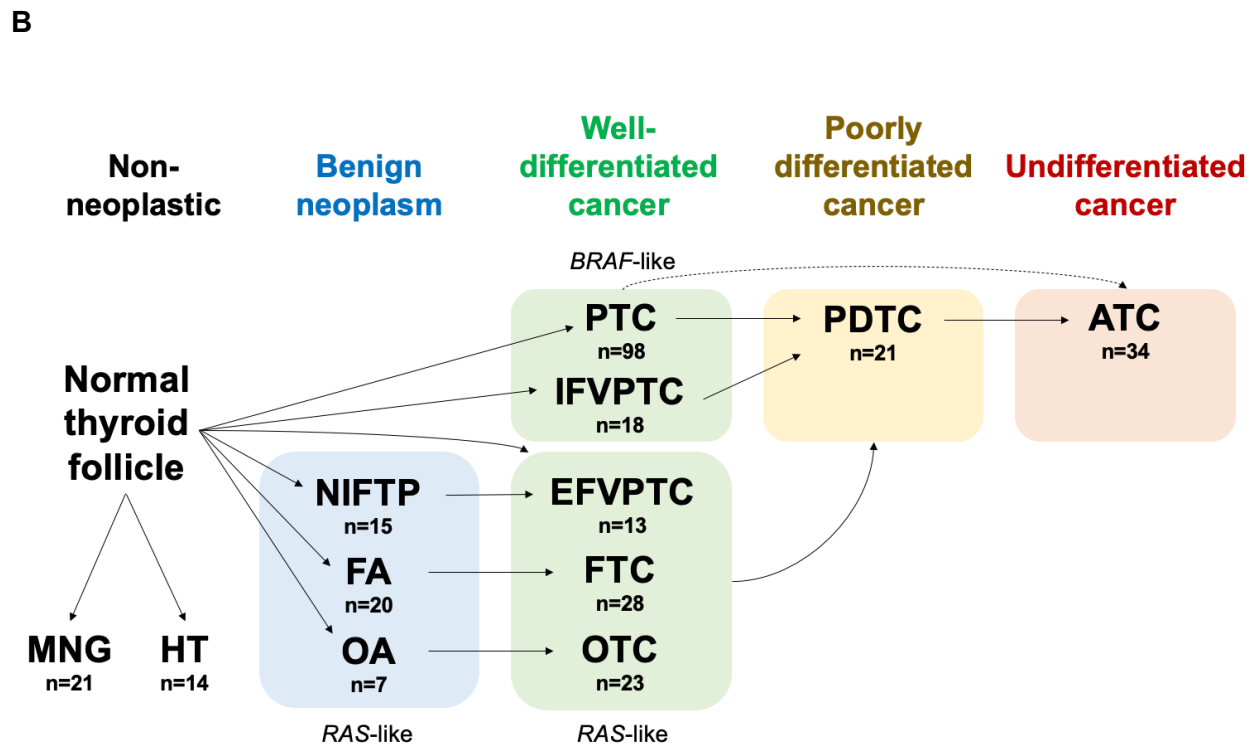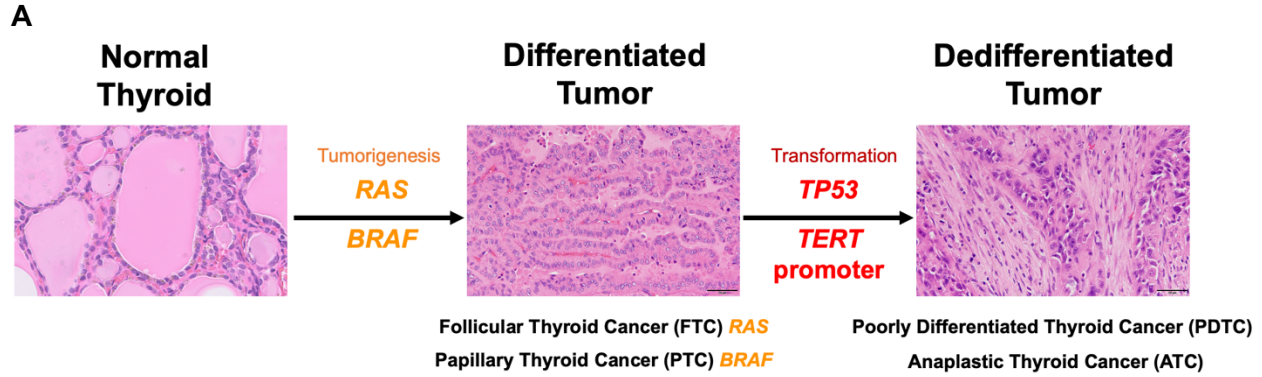
Figure 1-2: A model of thyroid cancer development

A) Diagram showing example pathway of thyroid cancer development, with representative H&E histology images for normal thyroid,[27] differentiated tumor, and dedifferentiated tumor. Commonly mutated genes thought to support each transition and example thyroid lesion subtypes are included.

B) Summary of the 12 thyroid lesion subtypes present in our cohort. For each subtype, the total number of cohort samples (n) is shown. Potential paths of evolution between different subtypes are indicated by arrows, with dotted arrows representing uncertain paths. Abbreviations: MNG = Multinodular Goiter; HT = Hashimoto Thyroiditis; FA = Follicular Adenoma; OA = Oncocytic Adenoma; NIFTP = Noninvasive Follicular Thyroid Neoplasm with Papillary-like Nuclear Features; EFVPTC = Encapsulated Follicular Variant Papillary Thyroid Carcinoma, OTC = Oncocytic Thyroid Carcinoma; FTC = Follicular Thyroid Carcinoma; PTC = Papillary Thyroid Carcinoma; IFVPTC = Infiltrative Follicular Variant Papillary Thyroid Carcinoma, PDTC = Poorly Differentiated Thyroid Carcinoma; ATC = Anaplastic Thyroid Carcinoma.

Molecular changes involved in thyroid cancer can also be found in the canonical Wnt pathway (Figure 1-3). In both normal development as well as in many different types of cancers, the Wnt pathway is involved in controlling important cellular functions such as proliferation, migration, and differentiation. Canonical Wnt signaling involves regulation of β-catenin, whose accumulation or degradation is central to Wnt signaling upregulation or downregulation. When the canonical Wnt pathway is turned on, Wnt ligand binds to the Frizzled receptor which leads to inactivation of the destruction complex and β-catenin accumulation and translocation to the nucleus. There, it binds to transcription factors of the T cell factor and lymphocyte enhancer factor (TCF/LEF) family, ultimately upregulate the expression of Wnt pathway target genes. When β-catenin is not present, TCF/LEF interacts with the transcriptional corepressor Groucho (Gro) in invertebrates or Transducin-like enhancer (TLE) in vertebrates. Mutations in the Wnt pathway that stabilize β-catenin or disrupt its destruction complex lead to abnormal upregulation of genes driving cell growth, proliferation, stemness, and other behaviors that may promote cancer progression. While Wnt pathway mutations may act as drivers of carcinogenesis in some cancers, such as *APC* mutations in colorectal cancer,[28] Wnt pathway mutations in thyroid cancer have been associated with more aggressive outcomes. Previous research has shown that as many as 60-80% of transformed thyroid tumors may carry Wnt pathway mutations.[29-31] One study of 22 ATC tumors found about 82% of tumors with an AXIN1 mutation, but less than 10% of tumors having a *CTNNB1* (β-catenin gene) or *APC* mutation.[29] In contrast, other studies involving ATC have found *CTNNB1* mutations to be much more common. Stabilization of β-catenin may occur through activating mutations in exon-3 at the phosphorylation sites for ubiquitination and degradation of beta-catenin. In one study of 31 ATC patients, researchers performed sequencing of exon-3 and found *CTNNB1* mutations in about 61% of the patients analyzed.[30] Another study involving 29 ATCs found about 66% of cases showing *CTNNB1* exon 3 mutations.[31]
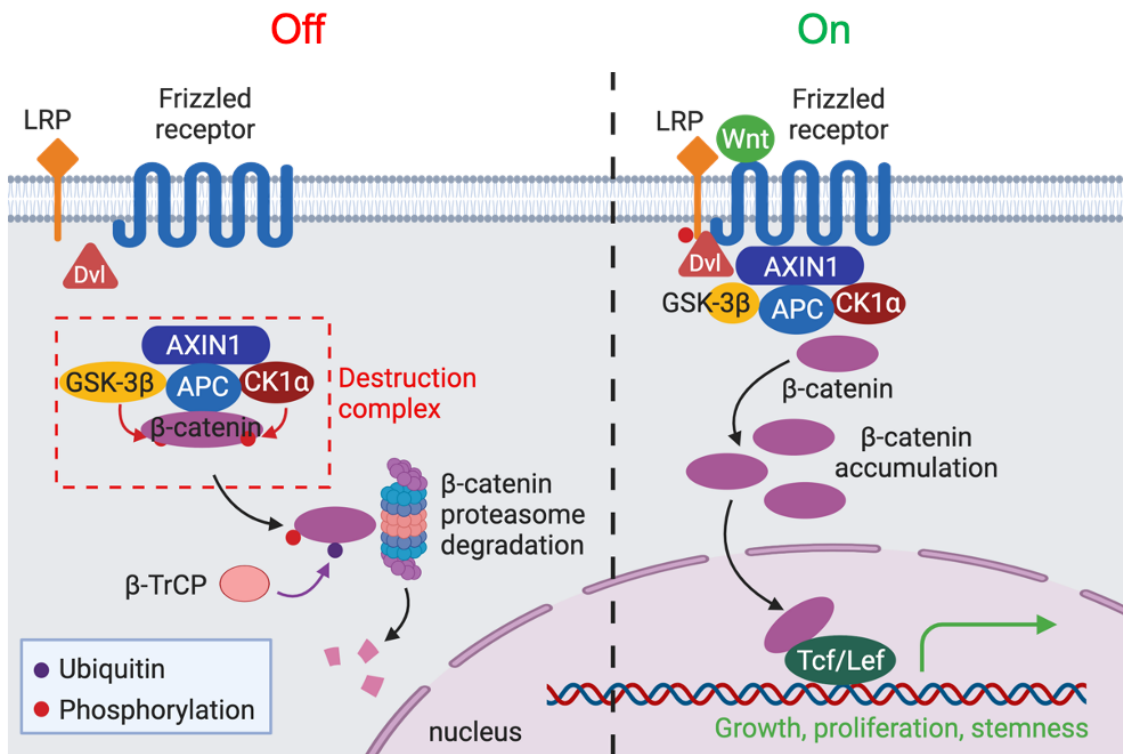
Figure 1-3: Canonical Wnt Signaling
Summary of canonical Wnt signaling pathway in either the off (left) or on (right) state. This figure was generated using BioRender.

In addition to mutations driving thyroid cancer, other alterations are known to be important in thyroid cancer pathogenesis, such as loss of heterozygosity (LOH) alterations, copy number variants (CNVs), and gene fusions. LOH alterations are more common in oncocytic thyroid carcinomas (OTC), also known as Hürthle cell thyroid carcinomas. LOH alterations can result in the loss of multiple tumor suppressor genes, and tumors with these alterations can have enhanced genetic instabilities and tend to be activated with cyclin-dependent kinase signaling.[32] A few studies have also reported LOH alterations to occur more commonly in follicular thyroid cancer subtypes than in PTCs,[33] and have also reported LOH alterations in ATCs on chromosome 16p and 18q.[34] CNVs can be found in follicular thyroid cancer subtypes, PTC, and ATC.[35] CNVs in PTC are likely common but usually involve gains, while CNVs in follicular carcinoma may be more common and often involve deletions, such as the loss of chromosome 22. ATCs may commonly contain CNVs involving both gains and deletions.[35] Regarding thyroid cancer fusions, in addition to the previously mentioned *PAX8/PPARG* fusion commonly found in follicular thyroid cancers,[12] other common fusions include those involving the receptor tyrosine kinases (RTK), such as *RET, NTRK1, NTRK3*, and *ALK*.[36] Of these fusions, *RET* fusions are the most common, particularly in PTC; these fusions sometimes referred to as *RET/PTC* fusions. Fusions involving RTK genes that drive cancer can occur when the 3′ portion of the RTK gene, coding for the kinase domain, becomes fused with the 5′ portion of a partner gene, usually carrying with it ubiquitous expression which replaces normal regulation of the original RTK gene.[36]

1.1.3 Tumor Microenvironment

In addition to mutations and gene expression changes, the tumor microenvironment may also influence tumor progression as well as response to immunotherapy. The tumor microenvironment includes the extracellular matrix (ECM), signaling molecules, and non-tumor cells, such as immune cells and stromal cells (Figure 1-4). While many thyroid tumors are generally considered to be immune cold, recent research has shown that neutrophils may be

associated with larger thyroid tumor size.[37] In addition, thyroid cancer-derived soluble factors have been found to activate neutrophils and induce pro-tumorigenic activities; for example, thyroid cancer-derived CXCL8/IL-8 can act on CXCR1 and CXCR2 expressed on neutrophil to promote chemotaxis, while granulocyte-macrophage colony-stimulating factor (GM-CSF) can prolong neutrophil survival.[37] These activated neutrophils may then go on to express pro-tumorigenic and angiogenic factors such as VEGF-A, CXCL8/IL-8, and MMP-9, which are known contributors of cancer-related inflammation.[37]

In addition, tumor associated macrophages (TAMs) may also be associated with thyroid cancer aggression. For example, TAMs have been shown to be correlated with lymph node metastasis in PTC, potentially supported by TAM production of CXCL8/IL-8 as demonstrated by in-vivo studies.[38] Another study involving high-throughput single-cell RNA-Seq (scRNA-Seq) analysis of PTCs and ATCs found that both subtypes may have high myeloid cell populations.[39] Their findings also suggested that during ATC progression, macrophages may be reprogrammed from an M1 state, which is associated with antitumor functions such as promoting inflammation and mediating cytotoxicity against tumor cells, toward an M2 state, which is associated with aggression-promoting behaviors such as suppressing immune responses and promoting angiogenesis and metastasis.[39,40]

Cancer associated fibroblasts, or CAFs, are another key component of the tumor microenvironment. CAFs are a heterogenous and highly plastic group of cells whose roles include secretion of growth factors, inflammatory ligands, and extracellular matrix (ECM) proteins. CAFs may also promote tumor proliferation, metastasis, immune exclusion, and therapy resistance.[41,42] CAFs may be recruited by factors such as TGFβ, PDGF, and FGF2, produced by cancer cells, stromal cells, or immune cells.[41] As CAFs are active players during wound healing, and tumors are sometimes considered wounds that do not heal, CAFs recruited to tumors may be perpetually activated and drive increasing stromal content, which can include collagens, laminins, fibronectins, proteoglycans, periostin, and tenascin C.[41] Activated fibroblasts can also produce

ECM-degrading proteases, such as the MMPs, supporting aggression by promoting epithelial to mesenchymal transition (EMT), motility, and invasion.[41] Recent studies exploring the clinical and therapeutic relevance of CAFs have highlighted the difficulty of targeting this diverse and complex class of cells and have emphasized the importance of understanding both the tumor-promoting and tumor-restraining subtypes of CAFs.[42] While recent studies have begun to characterize CAFs in other cancers, such as pancreatic cancer and breast cancer, less is known about the role of CAFs in thyroid cancer. Some recent research has found that CAFs may be associated with thyroid cancer dedifferentiation and aggressive outcomes.[43]
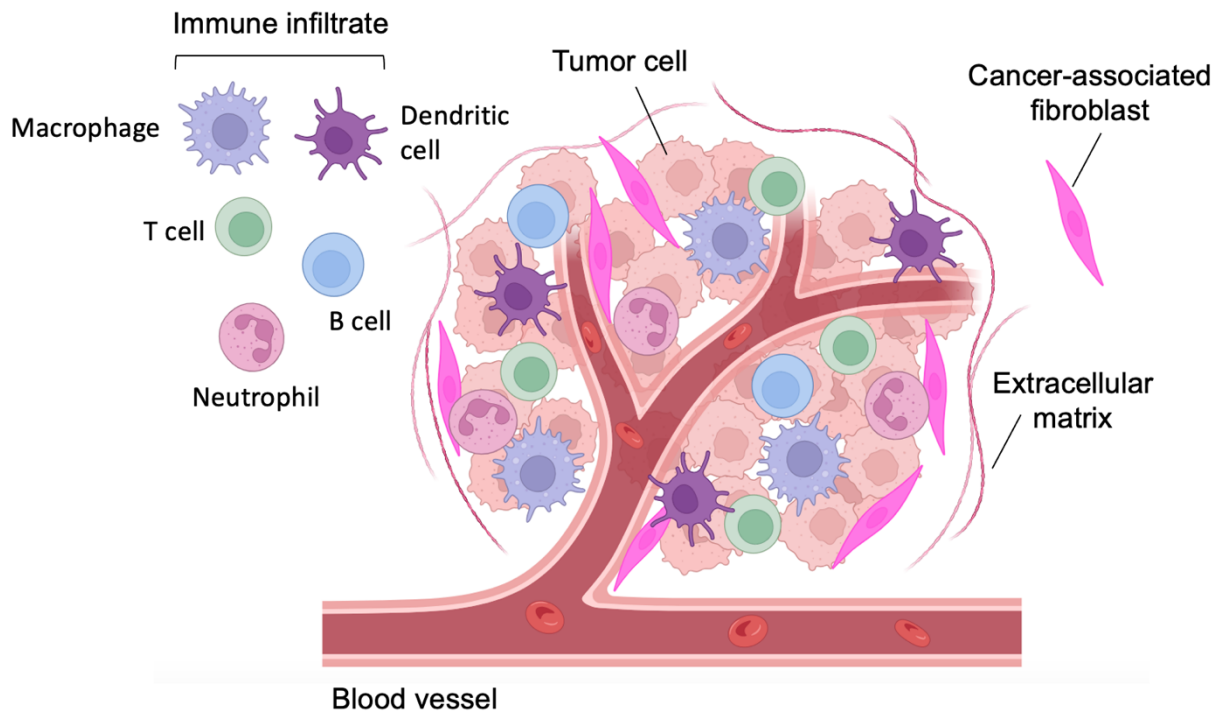
Figure 1-4: The tumor microenvironment
Summary of the tumor microenvironment. This figure was generated using BioRender.

1.1.4 Prediction

Accurate molecular testing for cancer prediction is critical for both preventing cancer and guiding therapy. Thyroid cancer typically occurs in thyroid nodules, which are highly common and can be detected by palpation and imaging. As the majority of thyroid nodules are benign, testing is required to determine which nodules contain cancer and require further treatment.[8,9,44-47] A widely used malignancy testing method for thyroid cancer is fine-needle aspiration (FNA) cytology, a minimally invasive biopsy technique that uses a needle and syringe to extract a sample of cells from the thyroid nodule, followed by cytologic examination of collected cells. While FNA cytology is able to diagnose of a malignant or benign nodule in most cases, as high as 25% of nodules may have an inconclusive diagnosis using this method.[8,9,44-47] Such nodules may be diagnosed conclusively by diagnostic lobectomy; however, this method is more invasive and unfavorable both for patients with malignant disease, who are required to undergo additional surgery for completion thyroidectomy, and for patients with benign disease, who had surgery performed for benign nodules.[8,9,44-47]

Therefore, recent research to predict thyroid nodule malignancy at the time of initial biopsy has focused on molecular testing. The majority of molecular tests developed for thyroid cancer for malignancy prediction typically involve screening for common driver alterations in well-differentiated thyroid cancers; these include mutations in the MAPK and PI3K–AKT signaling pathways, such as *BRAF* and *RAS* mutations. Classes of alterations that may be used in such tests include point mutations, insertions/deletions, gene fusions, copy number alterations, or gene expression alterations.[8,9,44-47] For example, ThyroSeq v3 is a 112-gene test that uses both DNA- and RNA-based next-generation sequencing to look at a broad range of thyroid cancer-related point mutations, insertions/deletions, gene fusions, copy number alterations, and gene expression alterations.[45,48] Another commercial molecular test for thyroid malignancy is the Afirma Genomic Sequencing Classifier (GSC), which uses analysis of RNA sequencing data to look at 905 genomic variants and 235 fusion pairs from 593 genes to rule out malignancy and reclassify

cytologically indeterminate nodules to molecularly benign or suspicious.[44] A third commercial test, Interpace Diagnostics ThyGeNEXT + ThyraMIR, uses both a mutation panel and a microRNA panel to diagnose cytologically indeterminate thyroid lesions.[49,50]

Current molecular tests are capable of distinguishing malignant from benign thyroid lesions with high sensitivity, specificity, and accuracy.[44,45,48] However, there are currently only a few established molecular tests for predicting aggressive thyroid cancer. Aggressive thyroid cancer can include recurrence of disease, transformation to PDTC or ATC, and distant metastasis. While ThyroSeq is used for malignancy prediction, it also provides pre-operative assessment of cancer recurrence risk.[45,48] However, while this test has been evaluated in published studies, it has not yet been widely utilized to confirm its the real-world utility and performance. Afirma also states that its assay has the potential to predict tumor behavior, such as lymph node metastasis and extrathyroidal extension, with inclusion of its newer test, XA. XA uses RNA seq data to identify variants and fusions instead.[44] However, similar to ThyroSeq's risk of recurrence score, the real-world utility of Afirma's test has not yet been evaluated.

In addition to improving aggressive thyroid cancer prediction, molecular tests also have the potential to further improve other aspects of patient care. Importantly, such molecular tests could offer is the ability to inform specific targetable alterations or therapies for each patient. For example, testing to detect molecular markers in a patient's tumor microenvironment could potentially inform their response to immunotherapy. These advances could be particularly important for treating the most aggressive thyroid cancer subtype, ATC, which currently has no curative therapy. However, regardless of the markers used or information gained, future molecular tests must also demonstrate clear advantages over existing methods. For example, some studies have suggested that current diagnostic molecular tests may not be cost-effective compared to diagnostic lobectomy.[51] Addition research is needed to ensure that future molecular tests significantly improve patient care and outcomes when compared to established methods.

1.2 Research objectives

To study the molecular markers of thyroid cancer aggression, our lab has collected and curated a large thyroid lesion cohort with patient samples from Vanderbilt University Medical Center (VUMC) and the University of Washington Medical Center (UWMC). In other thyroid cancer studies, cohorts can be limited to one or a few specific thyroid lesion subtypes, limiting the ability to comprehensively assess alteration patterns across different diagnoses. In addition, as aggressive thyroid cancer is rare in comparison to the numerous cases of well-differentiated subtypes such as PTC, thyroid cancer cohorts typically contain smaller numbers of aggressive thyroid lesion samples. These cohorts reflect real-world populations but may also limit the study of aggressive thyroid cancer. To try and maximize the value of our cohort, we collected a wide range of thyroid lesion subtypes spanning 12 unique diagnoses, and deliberately enriched our cohort for aggressive thyroid cancer cases. Our first objective was to analyze DNA sequencing data from this cohort to identify mutations driving aggressive thyroid cancer behaviors, such as distant metastasis, recurrence, and transformation to dedifferentiated subtypes such as poorly differentiated thyroid cancer (PDTC) and anaplastic thyroid cancer (ATC).

Our next research objective was to look at gene expression changes associated with thyroid cancer aggression. One key pathway of interest in our lab has been the canonical Wnt pathway, which has been implicated in the development and progressions of many cancers including thyroid cancer. We aimed to use bulk RNA sequencing data to analyze patterns of gene expression changes that may explain aggression even when common aggression-associated mutations are not present.

Our final objective was to determine the role of the immune microenvironment in thyroid cancer progression. We planned to use both computational deconvolution tools to analyze bulk RNA sequencing data as well as tissue sample staining and imaging to validate our computational findings. Previous research in other cancers has suggested that infiltrating immune cells may

support tumor growth and progression to aggressive disease. We hypothesize that infiltrating immune cells in the thyroid tumor may support metastasis, recurrence, and transformation.

In summary, our overall objective was to uncover a diverse set of genetic, gene expression, and tumor microenvironment changes that support thyroid cancer aggression. Using these findings, we also aimed to create a new molecular score based on novel biomarkers with the ability to accurately predict poor outcome risk in thyroid cancer patients. Finally, we anticipate that our findings will ultimately open up new paths for the development of novel therapies targeting aggressive thyroid cancer.

CHAPTER 2

Title of Chapter: Molecular Profiling of a Large Thyroid Lesion Patient Cohort

This chapter includes adaptation of contents from the following manuscript:

Xu, G., Loberg, M., et al. Molecular Assessment of Tumor Mutations and Microenvironment Enhances Prediction of Thyroid Cancer Outcome. *Cell Genomics.* 2023.

Contributions: Matt Loberg and I contributed equally to this paper. I performed cohort collection and curation, sequencing, omics analysis and computational deconvolution methods. Matt performed RNA sequencing analysis, tissue staining, microscopic imaging, and spatial transcriptomics methods.

2.1 Introduction

DNA and RNA sequencing have revolutionized cancer research, driving advances in cancer classification, diagnosis, risk prediction, and therapy. However, genomic understanding of thyroid cancer has lagged behind other cancers. Most advances in molecular testing of thyroid cancer have focused on malignancy detection, using common driver mutations in well-differentiated thyroid cancers such as *BRAF* V600E and *RAS* mutations.[8,9,46,47] In our research, we sought to characterize the genetic landscape of a large cohort of thyroid lesion patients. For our initial sequencing research objectives, we aimed to 1) confirm the presence of known thyroid cancer mutations and gene expression changes in our cohort, and 2) investigate associations between mutations/gene expression changes and thyroid cancer aggression.

2.2 Methods

2.2.1 Patient Cohort

Our patient cohort contains 312 formalin-fixed paraffin-embedded (FFPE) resection samples from 251 patients with thyroid nodules. Patients from Vanderbilt University Medical Center (VUMC) and the University of Washington Medical Center (UWMC) were included, and

this study was performed in compliance with VUMC and UWMC Institutional Review Boards. Diagnostic criteria are based on WHO and ATA guidelines, and each specimen's histopathology was reviewed by board-certified pathologists. Manual review of patient charts was performed to collect additional patient demographics (e.g. race), clinical histories (e.g. prior exposures to ionizing radiation), treatment courses (e.g. surgery type), tumor details (e.g. size), and outcomes (e.g. survival).

To categorize aggressive disease, patients were grouped into "indolent" or "aggressive disease" groups. The indolent category includes patients with no evidence of disease (NED), indeterminate disease, persistent disease, or recurrent disease in remission. NED was defined by undetectable thyroglobulin (Tg), undetectable anti-thyroglobulin antibodies (aThyG), and thyroid ultrasound indicating no evidence of disease. Patients without imaging follow-up were categorized as NED by laboratory testing (undetectable Tg or aThyG) alone, and patients with imaging alone (no labs) were categorized as NED if the patient had a hemithyroidectomy or no radioactive iodine. Indeterminate disease was defined by stable/detectable Tg < 1.0 ng/mL, stimulated Tg < 10 ng/mL, positive aThyG levels that were stable or decreasing, imaging without Tg labs, and/or inconclusive imaging. Persistent disease was defined by stable Tg > 1.0 ng/mL, stimulated Tg > 10 ng/mL, and/or a persistent lesion by imaging that did not increase in size over multiple years of follow-up. Recurrent disease in remission includes malignancies measurable by imaging/laboratory testing after NED designation but then had stable or decreasing tumor size/Tg following local intervention. Aggressive disease was defined by metastatic or transformed disease at presentation, distant metastasis after initial therapy completion, biopsy showing transformation to ATC; local disease recurrence without stabilization or remission following subsequent localized treatment; or increasing lesion size after initial therapy.

Progression-free survival (PFS) was calculated as the time between date of initial therapy completion and date of progression. Progression date was determined by the first date of either increasing Tg (in a thyroid stimulating hormone-suppressed patient) or increasing tumor size by

imaging. All patients that were progressive by Tg had subsequent imaging evidence of progressive disease. For patients without progression, the date of last follow-up was used to calculate PFS, and the data was appropriately censored. Patients without follow-up after therapy were omitted from analysis. Overall survival (OS) was calculated as the time between date of initial therapy completion and date of death or last follow-up (censored in the case of a living patient). For well-differentiated tumors, the date of initial therapy completion was either the date of post-operative radioactive iodine administration or the date of surgery for low-risk tumors not requiring post-operative radioactive iodine. For undifferentiated tumors, the surgery date was used as the initial therapy completion date.

2.2.2 Whole Exome Sequencing and Analysis

Nucleic acids were extracted using the COVARIS truXTRAC FFPE Total NA Kit (COVARIS, Woburn, MA) and DNA libraries were built using the NEB Ultra II DNA Library Prep Kit (NEB, Ipswich, MA) per the manufacturer's instructions. Sequencing was completed at the Vanderbilt Technologies for Advanced Genomics (VANTAGE) core facility using the IDT xGen® Exome Research Panel on an Illumina NovaSeq 6000 platform (Illumina, San Diego, CA). Raw 150 bp paired-end reads were trimmed to remove adapter sequences using Cutadapt (v2.10)[52] and the quality of the reads before and after trimming was measured by FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc).[53] Trimmed reads were aligned to the hg38 genome using Burrows-Wheeler Aligner (v0.7.17-r1188),[54] and GATK v. 4.1.8.1 was used to remove duplicate reads, perform base quality score recalibration, and discover variants.[55] Variant calling was first performed on individual samples with HaplotypeCaller in gVCF mode, all samples were jointly genotyped, and variant filtering was performed using VQSR. Variant annotation was conducted using ANNOVAR (v2018-04-16).[56] To exclude variants common in the population, variants with minor allele frequency ≥0.1% in at least one of the ExAC (Exome Aggregation Consortium),[57] 1000G (1000 Genomes Project),[57] and gnomAD (Genome Aggregation

Database)[58] databases were filtered out. *BRAF*, *RAS*, *TP53*, and *PIK3CA* mutations were assessed according to the standards and guidelines for the reporting of sequence variants in cancer by the Association for Molecular Pathology, American Society of Clinical Oncology, and the College of American Pathologists.[59] Whole exome sequencing average depth was 157X, and coverage was 91X. Mutation data was visualized by oncoplots generated using the R packages maftools[60] and ComplexHeatmap.[61]

2.2.3 Sanger Sequencing of the *TERT* promoter

TERT promoter alterations C228T and C250T were probed using Sanger sequencing. Using DNA purified as described in section 2.2.2, primers [5- TAATACGACTCACTATAGGGCA CCCGTCCTGCCCCTTCACCTT-3' (forward+T7 tail) and 5'- GTAAAACGACGGCCAGGGCTTC CCACGTGCGCAGCAGGA-3' (reverse+M13F tail)],[62] and the HotStarTaq DNA Polymerase kit (QIAGEN, Hilden, Germany), thermal cycling was performed using the following conditions: 95°C (15 minutes), 35 cycles of 94°C (30 seconds), 56°C (30 seconds), and 72°C (20 seconds), followed by 72°C (10 minutes) and 4°C hold. Sanger sequencing was then used to analyze the purified PCR products (GENEWIZ, South Plainfield, NJ).

2.2.4 Bulk RNA Sequencing and Analysis

Nucleic acids were extracted using the COVARIS truXTRAC FFPE Total NA Kit (COVARIS, Woburn, MA), and Illumina TruSeq mRNA sequencing libraries were prepared. Sequencing was performed at VANTAGE on a NovaSeq 6000 platform Raw (Illumina, San Diego, CA). Raw 150 bp paired-end reads were trimmed to remove adapter sequences using Cutadapt (v2.10).[52] Alignment to the GENCODE GRCh38.p13 genome[63] was performed using STAR (v2.7.8a),[64] and GENCODE v38 gene annotations were provided to STAR to improve the accuracy of mapping. Quality control was performed on both raw reads and adaptor-trimmed reads using FastQC,[53] and featureCounts (v2.0.2)[65] was used to count the number of mapped

reads to each gene. Genes that were significantly differentially expressed with FDR-adjusted p value < 0.05 and absolute fold change > 2.0 were detected by DESeq2 (v1.30.1)[66] and visualized using the R package EnhancedVolcano (1.18),[67] while the R package Heatmap3[68] was used for cluster analysis and visualization. Gene Ontology was performed on differentially expressed genes using the Gene Ontology Consortium resource,[69,70] and gene set enrichment analysis was performed using GSEA (v4.1.0) on the msigdb v7.1 database.[71] TIMER2.0 (http://timer.cistrome.org/),[72] a web-based deconvolution program for estimating tumor-infiltrating immune cells based on gene expression profiles across diverse cancer types, was run using THCA (Thyroid Carcinoma) as the cancer type gene signature. Deconvolution scores in TIMER 2.0 that were used included those from CIBERSORT-Abs[73], EPIC[73], and MCPCOUNTER,[74] and descriptive results were plotted using the R package ggplot2.[75] We also used the computation tool TIDE (http://tide.dfci.harvard.edu/)[76] which used gene expression data to make estimates pertaining to immune checkpoint blockade response. TIDE was run using the following settings: Cancer type = Other, Previous Immunotherapy = No. Both TIMER and TIDE score heatmaps were generated with the R package ComplexHeatmap.[77]

For fusion analysis, the STAR-Fusion (v2.7.8a) pipeline[78] was used to align and map paired-end RNA-seq reads to the human genome (GRCh38_gencode_v37) using parameters optimized to capture fusion transcripts,[79] and FusionInspector, a component of the STAR-Fusion suite, was used to validate fusion transcripts in silico. Manual review of RNA data was performed using the Integrated Genomics Viewer.[80] Two additional RET fusions were also identified by blasting soft clip reads to the human genome.

2.2.5 RNA Sequencing Score Calculation

The *BRAF-RAS* score (BRS) was calculated using 69 genes from a previously defined gene list.[81] Two genes from the original gene list were not covered in our sequencing data and were omitted from our score calculation. Bulk RNA-sequencing data was transformed into z-score

format, and *BRAF*-mutant ([B]) and *RAS*-mutant ([R]) centroids were calculated from PTCs and

FVPTCs lesions with *BRAF* V600E and *RAS* mutations (*NRAS*, *HRAS*, or *KRAS*). A vector

containing the expression of the 69 BRS genes was generated ([S]) for each sample, and then

the normalized squared Euclidean distance between [S] and [B] and [S] and [R] was calculated.

Finally, the BRS was calculated as the difference between these normalized squared Euclidean

distances.

$$\text{BRS(S)} = |[S] - [B]| - |[S] - [R]|$$

The thyroid differentiation score (TDS) was calculated from the expression of 16 genes

related to thyroid function and metabolism, as previously described.[81] To calculate TDS, the

variance stabilized gene expression data were subtracted by the median across all tumor

samples, and the TDS was calculated from the average values across the 16 genes in each

tumor.

The PI3K score was calculated using a previously described hallmark PI3K-AKT-mTOR

(PI3K) signaling gene set.[82] Looking at all samples in our cohort except MNG and HT, we log2

transformed and calculated Z-scores of RNA-expression data for each of the 105 genes in the

PI3K gene set. We then calculated the score for each sample as the sum of the Z-scores for these

105 genes.

The ERK score was calculated according to previously described methods,[81] using a set

of 48 genes previously shown to be down-regulated with MEK inhibition (set A) as well as 4 genes

previously shown to be up-regulated with MEK inhibition (set B).[83] For samples across our cohort

excluding MNG and HT, we took expression data for each gene from set A and set B, and then

applied log2 transformation before calculating Z-score. For each sample, Z-scores of set A genes

were summed, and Z-scores of set B genes were summed. The final ERK score for each sample

was calculated by subtracting the Z-score sum of set B genes from the Z-score sum of set A genes. Complete gene lists for each gene expression-based score are shown in Table 2-1.

2.2.6 Data Sharing

Code for all analyses is available at the following GitHub link (https://github.com/xgj797/ Molecular-Signature-Incorporating-Immune-Microenvironment-Enhances-Thyroid-Cancer-Outcome-Prediction). Individual-level patient clinical and sequencing data is not able to be shared due to the retrospective nature of our cohort; patients with historic samples cannot be consented, especially those with highly aggressive and rapidly lethal disease. However, aggregate-level data reported in this paper will be shared by corresponding author (Dr. Vivian Weiss, VW) of the *Cell Genomics* paper "Molecular Assessment of Tumor Mutations and Microenvironment Enhances Prediction of Thyroid Cancer Outcome" upon request.

BRS Score Genes							
Up in RAS-like							
ANKRD46	CYB561	GNA14	HGD	KATNAL2	KCNAB1	KCNIP3	LGI3
MLEC	NQO1	SFTPC	SLC4A4	SORBS2			
Up in BRAF-like							
ABTB2	AHR	ANKLE2	ANXA1	ARNTL	ASAP2	BID	CDC42EP1
COL8A2	CREB5	CTSC	CYP1B1	DTX4	DUSP5	ETHE1	EVA1A
FAM20C	FCHO1	FN1	FSTL3	GABRB2	GBP2	ITGA3	ITGB8
KCNN4	LAMB3	LLGL1	LY6E	MDFIC	MET	PDE5A	PDLIM4
PLCD3	PLEKHA6	PNPLA5	PPL	PRICKLE1	PROS1	PTPRE	NECTIN4
RASGEF1B	RUNX1	RUNX2	SDC4	SEL1L3	SFTPB	SLC35F2	SOX4
SPOCK2	STK17B	SYT12	TACSTD2	TBC1D2	TGFBR1	DCSTAMP	TMEM43
TDS Score Genes							
THRA	FOXE1	PAX8	TSHR	NKX2-1	DUOX2	DUOX1	TG
GLIS3	SLC5A8	DIO1	TPO	SLC26A4	DIO2	SLC5A5	THRB
PI3K Score Genes							
ACACA	ACTR2	ACTR3	ADCY2	AKT1	AKT1S1	AP2M1	ARF1
ARHGDIA	ARPC3	ATF1	CAB39	CAB39L	CALR	CAMK4	CDK1
CDK2	CDK4	CDKN1A	CDKN1B	CFL1	CLTC	CSNK2B	CXCR4
DAPP1	DDIT3	DUSP3	E2F1	ECSIT	EGFR	EIF4E	FASLG
FGF17	FGF22	FGF6	GNA14	GNGT1	GRB2	GRK2	GSK3B
HRAS	HSP90B1	IL2RG	IL4	IRAK4	ITPR2	LCK	MAP2K3
MAP2K6	MAP3K7	MAPK1	MAPK10	MAPK8	MAPK9	MAPKAP1	MKNK1
MKNK2	MYD88	NCK1	NFKBIB	NGF	NOD1	PAK4	PDK1
PFN1	PIK3R3	PIKFYVE	PIN1	PITX2	PLA2G12A	PLCB1	PLCG1
PPP1CA	PPP2R1B	PRKAA2	PRKAG1	PRKAR2A	PRKCB	PTEN	PTPN11
RAC1	RAF1	RALB	RIPK1	RIT1	RPS6KA1	RPS6KA3	RPTOR
SFN	SLA	SLC2A1	SMAD2	SQSTM1	STAT2	TBK1	THEM4
TIAM1	TNFRSF1A	TRAF2	TRIB3	TSC2	UBE2D3	UBE2N	VAV3
YWHAB							
ERK Score Genes							
Set A							
ARID5A	B4GALT6	BRIX1	BYSL	CCND1	CHSY1	CXCL8	DDX21
DUSP4	DUSP6	EGR1	ELOVL6	ETV1	ETV4	ETV5	FOS
FOSL1	GEMIN4	GNL3	GPR3	GTPBP4	HMGA2	IER3	LIF
MAFF	MAP2K3	MYC	NOP16	PHLDA2	PLK3	POLR1C	POLR1G
POLR3G	PPAN	PPAT	PYCR3	RRS1	SH2B3	SLC1A5	SLC4A7
SPRED2	SPRY2	SPRY4	TNC	TNFRSF12A	TSR1	WDR3	YRDC
Set B							
GTF2A1L	HYDIN	KIR3DL2	SEMA6A				

Table 2-1: BRS, TDS, PI3K, and ERK score genes

24

2.3 Results

2.3.1 Mutation Landscape of our Thyroid Lesion Patient Cohort

To begin investigating the genetic landscape of our cohort of thyroid lesions, we collected mutation and fusion data using whole-exome sequencing and bulk RNA sequencing, respectively. In our mutation and fusion data (Figure 2-1, Table 2-2), we observed patterns consistent with those described in past literature. For example, *BRAF* mutations are the most common in our cohort overall and are present in a majority of our papillary lesions, with 64% of PTCs and 33% of IFVPTCs showing these mutations. *RAS* mutations were also common and are enriched in our follicular lesions as expected; *RAS* mutations were found in 40% of NIFTPs, 42% of EFVPTCs, and 40% of FTCs. In addition, known thyroid cancer fusions were identified and followed expected trends; most of the *PPARG* fusions identified were found in FTC and PDTC samples, while *RET* fusions were associated with papillary subtypes PTC and IFVPTC. Other fusions we identified included those involving genes *NTRK3*, *BRAF*, and *TERT*. As our cohort was designed to be enriched in aggressive lesions, mutations in genes associated with aggressive thyroid cancer were common. We found that mutations in the *TERT* promoter, *TP53*, and *PIK3CA* were associated with transformed subtypes ATC and PDTC, with 31% *TERT* promoter, 27% *TP53*, and 19% *PIK3CA mutations in ATC, and* 33% *TERT* promoter, 19% *TP53*, and 12% *PIK3CA mutations in PDTC*. Additionally, *TERT* promoter, *TP53*, and *PIK3CA* were also enriched in the aggressive subset of our papillary samples. For example, for IFVPTC/PTCs with distant metastases, 41% had *TERT*p mutations, 18% had *TP53* mutations, and 12% had *PIK3CA* mutations, while for all other IFVPTC/PTCs, 28% had *TERT*p mutations, 2% had *TP53* mutations, and 7% had *PIK3CA* mutations.
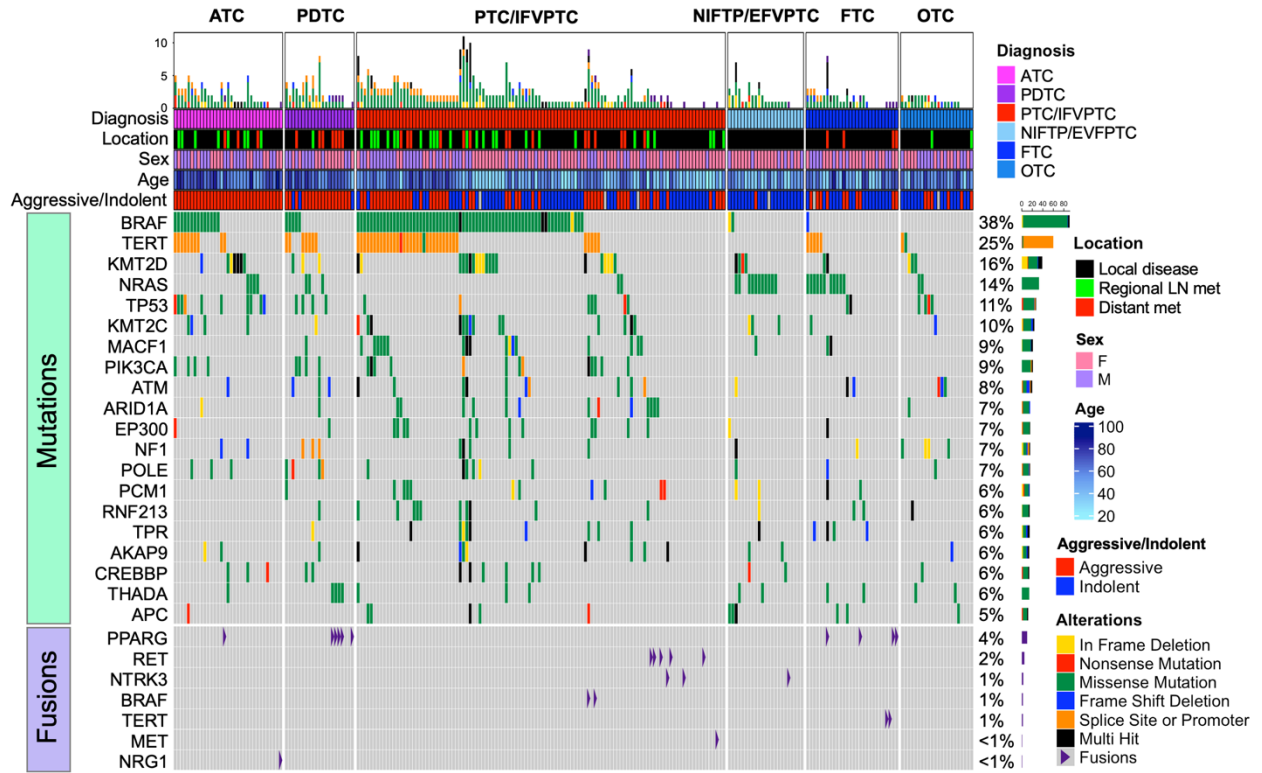
Figure 2-1: Oncoplot of select malignant thyroid lesions
Our oncoplot shows the top 20 mutated genes and top 7 genes involved in fusions in malignant samples in our cohort.

Diagnosis	BRAF mutation	RAS mutation	TERT promoter mutation	TP53 mutation	PIK3CA mutation	CTNNB1 mutation
MNG	0% (0/21)	0% (0/21)	0% (0/19)	0% (0/21)	0% (0/21)	0% (0/21)
HT	0% (0/14)	0% (0/14)	0% (0/14)	0% (0/14)	0% (0/14)	0% (0/14)
FA	0% (0/20)	15% (3/20)	0% (0/20)	0% (0/20)	0% (0/20)	0% (0/20)
OA	0% (0/7)	0% (0/7)	0% (0/7)	0% (0/7)	0% (0/7)	0% (0/7)
NIFTP	0% (0/15)	40% (6/15)	0% (0/12)	0% (0/15)	0% (0/15)	0% (0/15)
FTC	4% (1/28)	40% (11/28)	20% (5/25)	7% (2/28)	0% (0/28)	0% (0/28)
OTC	0% (0/22)	5% (1/22)	5% (1/22)	18% (4/22)	0% (0/22)	0% (0/22)
EFVPTC	17% (2/12)	42% (5/12)	0% (0/10)	0% (0/12)	0% (0/13)	0% (0/13)
IFVPTC	33% (6/18)	17% (3/18)	14% (2/14)	0% (0/18)	0% (0/18)	0% (0/18)
PTC	64% (63/98)	0% (0/98)	40% (32/80)	5% (5/98)	9% (9/98)	1% (1/98)
PDTC	24% (5/21)	14% (3/21)	33% (7/21)	19% (4/21)	19% (4/21)	5% (1/21)
ATC	41% (14/34)	12% (4/34)	31% (10/32)	27% (9/34)	12% (4/34)	0% (0/34)

Table 2-2: Quantification of key mutations organized by diagnosis
Table describing percent and counts of thyroid lesion samples with mutations in *BRAF*, *RAS*, *TERT* promoter, *TP53*, and *PIK3CA*. Lesion subtype are ordered by behavior from least to most aggressive from top to bottom. Sanger sequencing was used to detect *TERT* promoter mutations (n=276), and whole exome sequencing was used to find *BRAF* and *TP53* mutations (n=310).

2.3.2 Aggression-Associated Mutations are Associated with Worse Survival

We next looked at our mutation data in the context of thyroid cancer aggression. Using progression-free survival (PFS) analysis, we found that mutations in the *TERT* promoter, *TP53*, and *PIK3CA* were associated with significantly shorter progression-free survival time and overall survival time (Figure 2-2). We also generated oncoplots of our most commonly mutated genes in either well-differentiated papillary malignancies (PTC/IFVPTC/EFVPTC) or well-differentiated follicular malignancies (FTC/OTC) from patients with either indolent or aggressive disease (Figure 2-3). We found that well-differentiated papillary malignancies from patients with aggressive disease were more likely to have *TERT* promoter, *TP53*, and *PIK3CA* mutations, while well-differentiated follicular malignancies from patients with aggressive disease had a higher proportion of *TERT* promoter and *TP53* mutations as well. Overall, about 42% of well-differentiated tumor samples from patients with aggressive disease lacked mutations in the *TERT* promoter, *TP53*, or *PIK3CA*. In summary, while mutations in the *TERT* promoter, *TP53*, and *PIK3CA* are associated with worse survival, our data also suggest that a significant proportion of clinically aggressive thyroid cancers lack these common high-risk mutations.
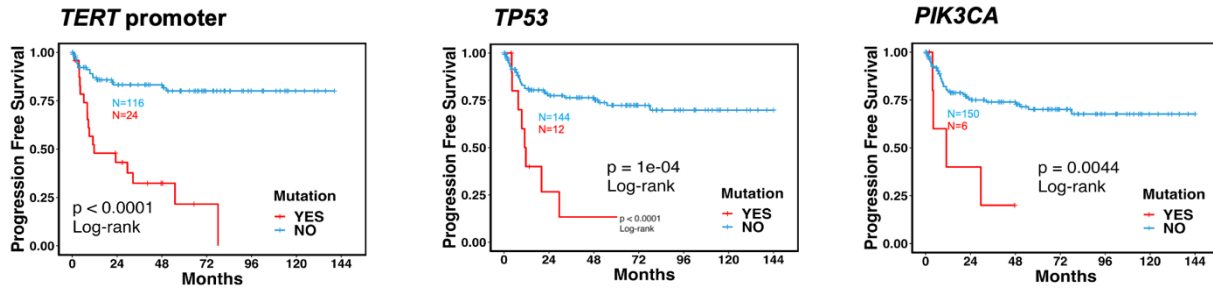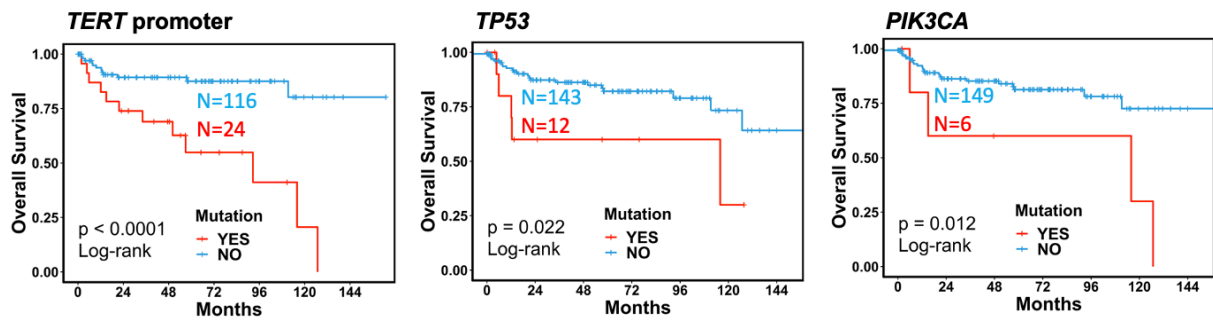
Figure 2-2: Survival curves of patients with and without mutations in the *TERT* promoter, *TP53*, and *PIK3CA*

A) Progression-free survival (PFS) of patients with and without mutations in *TERT* promoter, *TP53*, and *PIK3CA*. P values were calculated with log-rank test.

B) Overall survival (OS) of patients with and without mutations in *TERT* promoter, *TP53*, and *PIK3CA*. P values were calculated with log-rank test.
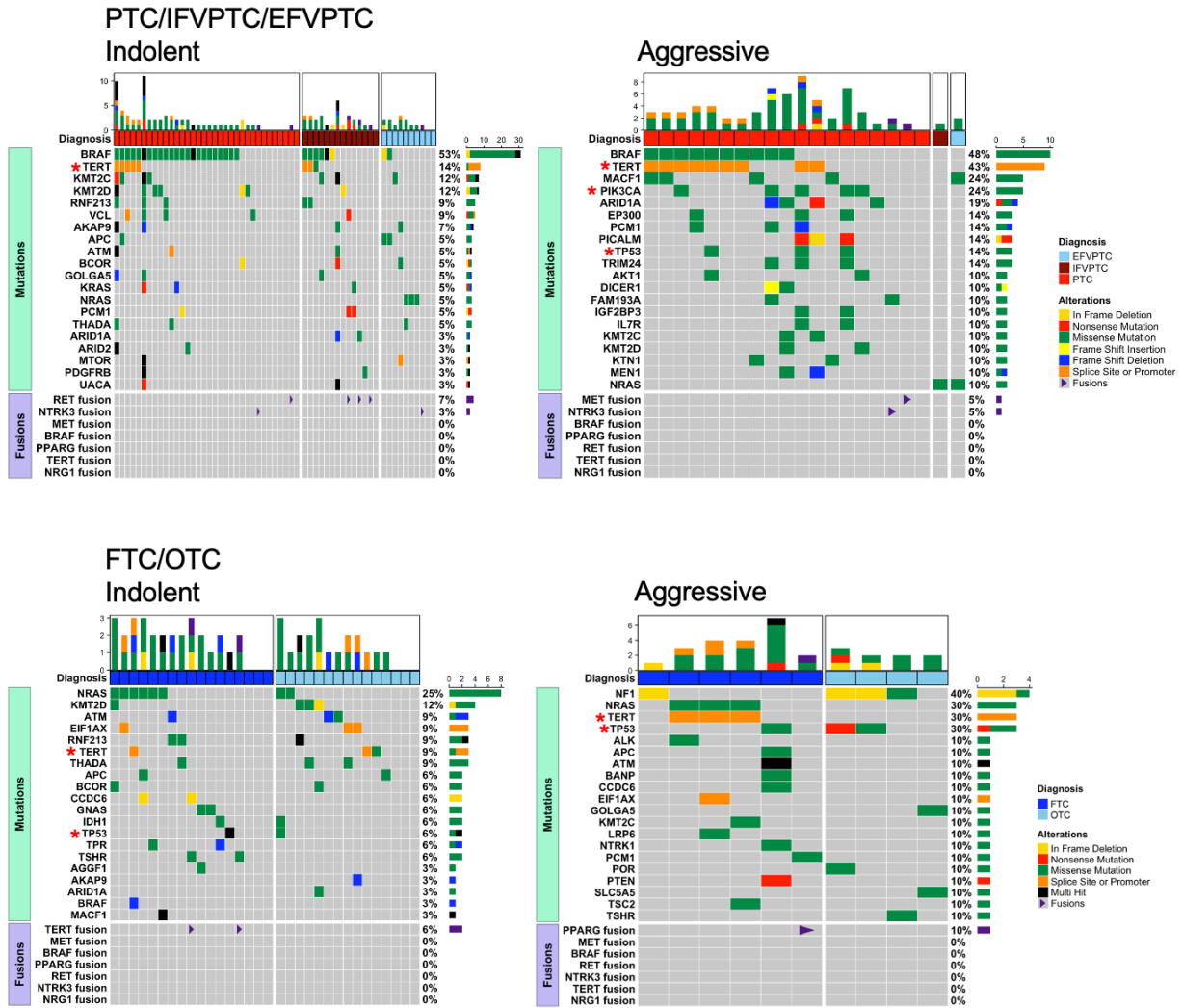
Figure 2-3: Oncoplots comparing indolent vs aggressive samples
A) Top 20 mutations in local PTC, IFVPTC, and EFVPTC lesions from patient with either indolent (left) or aggressive disease (right). *TERT*, *TP53*, and *PIK3CA* mutations are marked by red asterisks.
B) Top 20 mutations in local FTC and OTC lesions from patient with either indolent (left) or aggressive disease (right). *TERT*, *TP53*, and *PIK3CA* mutations are marked by red asterisks.

2.3.3 Thyroid Cancer Gene Expression Scores

We performed differentiational gene expression analysis of our bulk RNA sequencing data. It's known that *BRAF* and *RAS* mutations are the most common driver alterations in thyroid cancers. *BRAF* and *RAS* mutations are mutually exclusive with one another, with *BRAF* mutations being characteristic of papillary thyroid cancer (PTC), the most common type of thyroid cancer, and RAS being commonly mutated in follicular thyroid cancer subtypes.[8,9] Therefore, thyroid cancers have commonly been classified as either *BRAF*-like or *RAS*-like when analyzing gene expression.[81] In a heatmap showing differential expression of the 69 genes comprising the *BRAF-RAS* score (BRS) (Figure 2-4A), we see two general patterns of gene expression corresponding to our follicular lesions (FA, OA, NIFTP or EFVPTC, OTC, FTC), which are thought to be largely *RAS*-like, and our papillary lesions (PTC or IFVPTC) which are thought to be predominantly *BRAF*-like. PDTC and ATC are less clearly defined by these categories but overall tend to show *RAS*-like and *BRAF*-like patterns in our heatmap, respectively. Plotting our BRS scores as a boxplot organized by diagnoses (Figure 2-4B), we see that most of our ATCs and a significant proportion of our PTCs are *BRAF*-like. Previous research has found a slight association between *BRAF*-like status and thyroid cancer aggression; however, most thyroid cancers with *BRAF* mutations have good outcomes, so these signatures alone do not significantly predict aggressive disease for differentiated tumors. When looking at only well-differentiated cancer samples in our cohort, we find that *BRAF*-like status alone is not significantly associated with worse PFS (Figure 2-4C). Only after including transformed tumors PDTC and ATC is *BRAF*-like status significantly associated with shorter PFS, consistent with highly lethal ATCs being predominantly *BRAF*-like in our cohort.

As thyroid cancer progresses and becomes more aggressive, tumors become increasingly de-differentiated, losing characteristics such as their original thyroid cellular organization, functions, and markers. With this in mind, we also used gene expression analysis data to assess genes related to thyroid differentiation. Looking at 16 genes related to thyroid function and

metabolism in a gene expression heatmap (Figure 2-5A), we observed generally lower expression of a subset of these genes in PTCs/IFVPTCs, and a larger subset of these genes in our ATCs, as expected. While most ATCs have generally lower expression of *FOXE1, PAX8, TSHR, NKX2-1, DUOX2, DUOX1*, and *TG*, about half also have relatively high expression of genes *SLC5A8, DIO1, TPO, SLC26A4*, and *DIO2*. Our PDTCs do not appear to show the same extent of downregulation of these genes compared to our ATCs, which is expected as ATC is considered more de-differentiated than PDTC; however, our PDTCs also appear to have higher TDS gene expression even compared to our PTCs which are well-differentiated. Using these genes to calculate thyroid differentiation scores (TDS) and plotting them against diagnosis (Figure 2-5B), we observe a similar pattern as the one in our heatmap, with PTCs/IFVPTCs and ATCs showing overall lower TDS scores compared to other subtypes, suggesting lower differentiation.

The PI3K-AKT-mTOR (PI3K) signaling pathway is involved in regulating a wide range of processes including cell cycle progression, proliferation, adhesion, migration, survival, and differentiation.[16] This pathway is also known to be important in many cancers including thyroid cancer, especially in the transformed subtypes PDTC and ATC; some studies report one of the most commonly mutated PI3K pathway genes, *PIK3CA*, to be altered in 10-20% of these tumors.[8] Therefore, we also calculated PI3K signaling scores for our thyroid lesion cohort based on previously published methods. In our expression heatmap of genes used to calculate the PI3K score, we found patterns of gene upregulation shared between PTCs as compared to follicular subtypes. Similar to our heatmap containing TDS score genes, our PI3K gene score heatmap also shows what appear to be two unique subgroups in our ATCs (Figure 2-6A). Plotting our PI3K score against diagnosis, we find higher PI3K scores among our ATC samples compared to other those in other diagnosis (Figure 2-6B), suggesting especially strong PI3K signaling in this aggressive thyroid cancer subtype. We note that while most of our ATCs have high PI3K signaling, mutations associated with this pathway seem insufficient to explain this upregulation; for example, just 19% of ATCs have a *PIK3CA* mutation. However, we also note that copy number variations,

such as *PIK3CA* copy number gain, are also known drivers of upregulated PI3K signaling in thyroid cancer,[18] but were not studied in our cohort,.

The ERK (and MAPK) signaling pathway is a key pathway commonly altered in thyroid cancer; two of the most commonly mutated genes in thyroid cancer are *BRAF* and *RAS*, which both code for components of the ERK signaling pathway. We use previously published methods to calculate ERK signaling scores for tumors in our cohort. In a gene expression heatmap of genes used to calculate the ERK score, ATCs appear to be the most distinct and show particularly high expression of many ERK score genes (Figure 2-6A). In a boxplot with ERK score plotted against diagnosis, we find that ATC, and to a lesser extent papillary thyroid cancer subtypes PTC and IFVPTC, generally appear to show the highest ERK scores in our cohort (Figure 2-6B).

We perform principal component analysis (PCA) on our gene expression data, with colors annotating thyroid lesion diagnosis, BRS, TDS, and PI3K score (Figure 2-7). Our PCA results reaffirmed many of the patterns we observed in our previous gene expression heatmaps. Spatially, ATCs, PTCs, and follicular subtype appear to separate into distinct groups. However, PDTCs can be found intermixed with our follicular lesions, and FVPTCs are scattered both among PTCs and follicular subtypes. Looking at BRS score and TDS score annotations, our follicular subtype cluster stands out, with high BRS scores indicating RAS-like gene expression profiles, and high TDS scores indicating high differentiation, as expected. PI3K scores highlight our ATC cluster, predicting high PI3K signaling in this subtype.

A

B

BRS by Diagnosis

C

Differentiated BRS

Differentiated + Transformed BRS

34
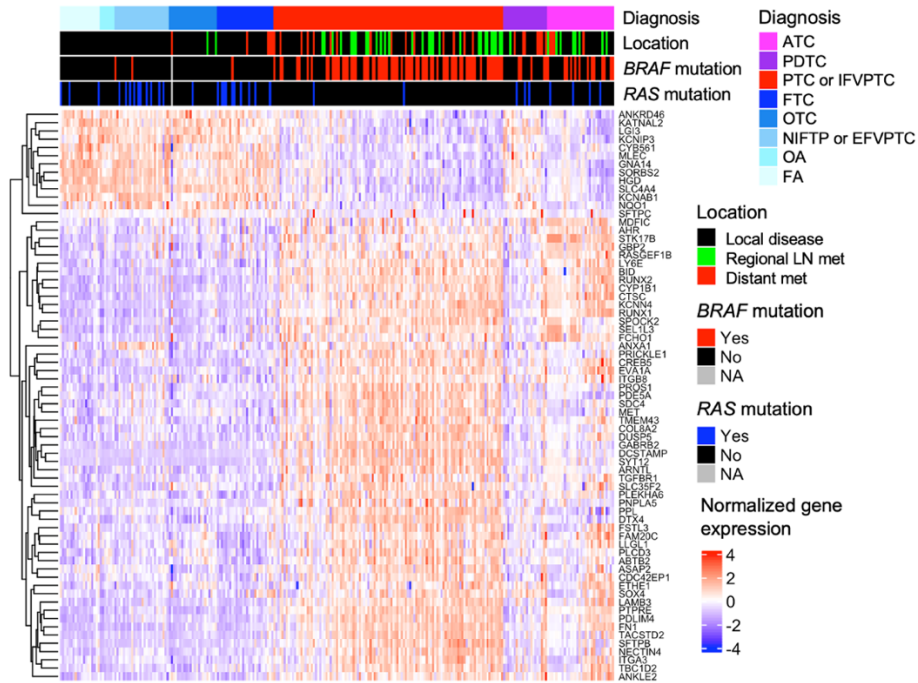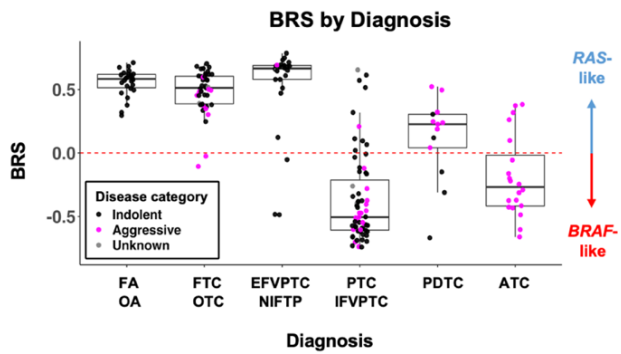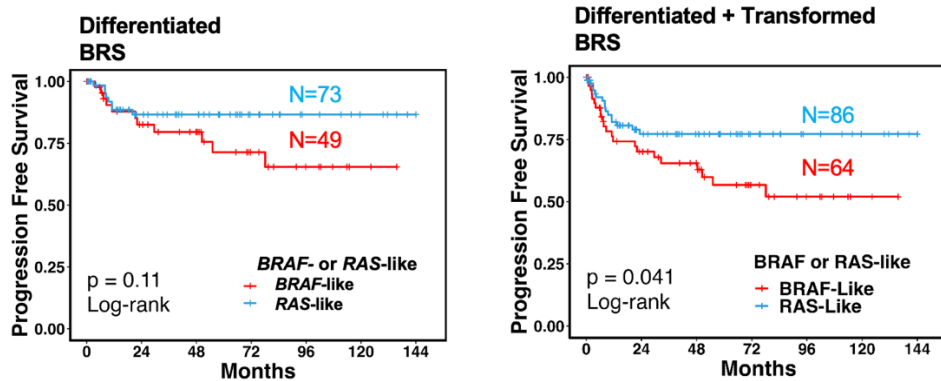
Figure 2-4: *BRAF-RAS* Score
A) Normalized gene expression heatmap showing 69 genes used to calculate *BRAF-RAS* score (BRS). Samples are ordered by diagnosis, and genes are ordered by hierarchical clustering.
B) Box plots showing BRS from local disease samples, with dot color indicating clinical behavior (pink = aggressive; black = indolent; grey = no clinical follow-up after sample collection).
C) PFS of *BRAF*-like (red) and *RAS*-like (blue) differentiated thyroid tumors (left) or both differentiated and transformed thyroid lesions (right). Local disease location only, all lesion subtypes included except MNG, OTC, FA, and OA. P values calculated using log-rank test.

A
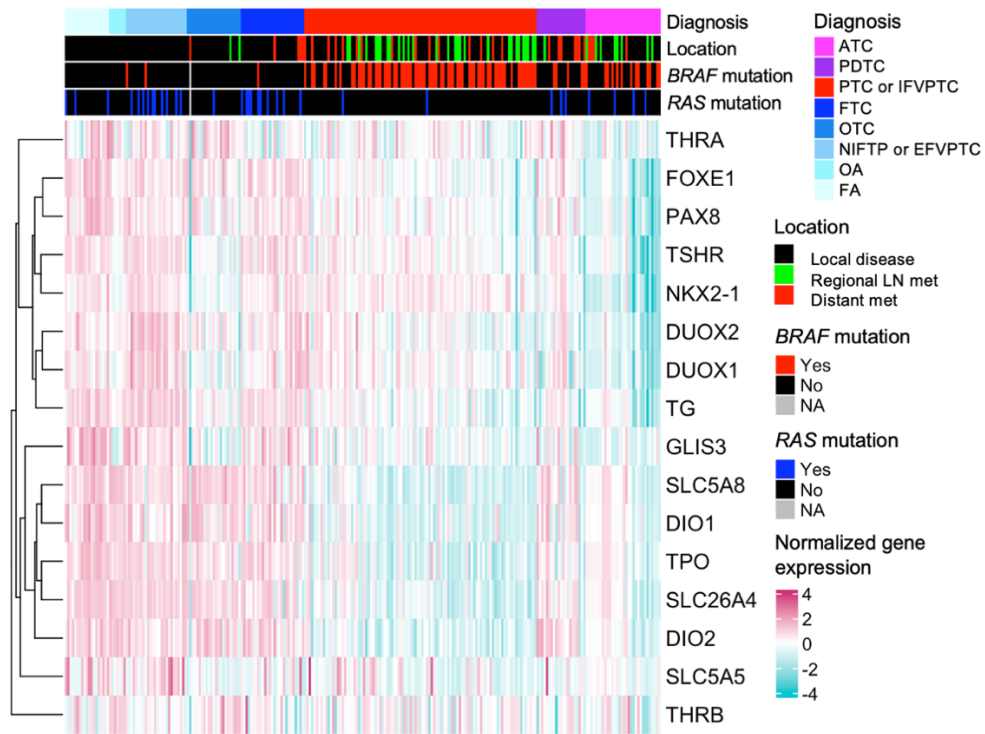
B
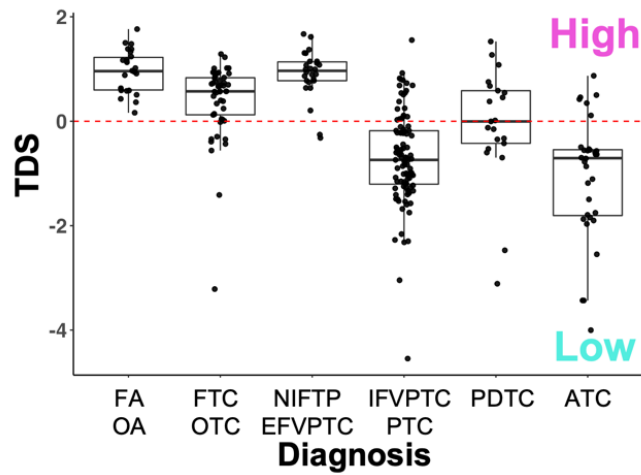
Figure 2-5: TDS Score
A) Normalized gene expression heatmap showing 16 genes used to calculate TDS. Samples are ordered by diagnosis, and genes are ordered by hierarchical clustering.
B) Box plots showing TDS from local disease samples.

A
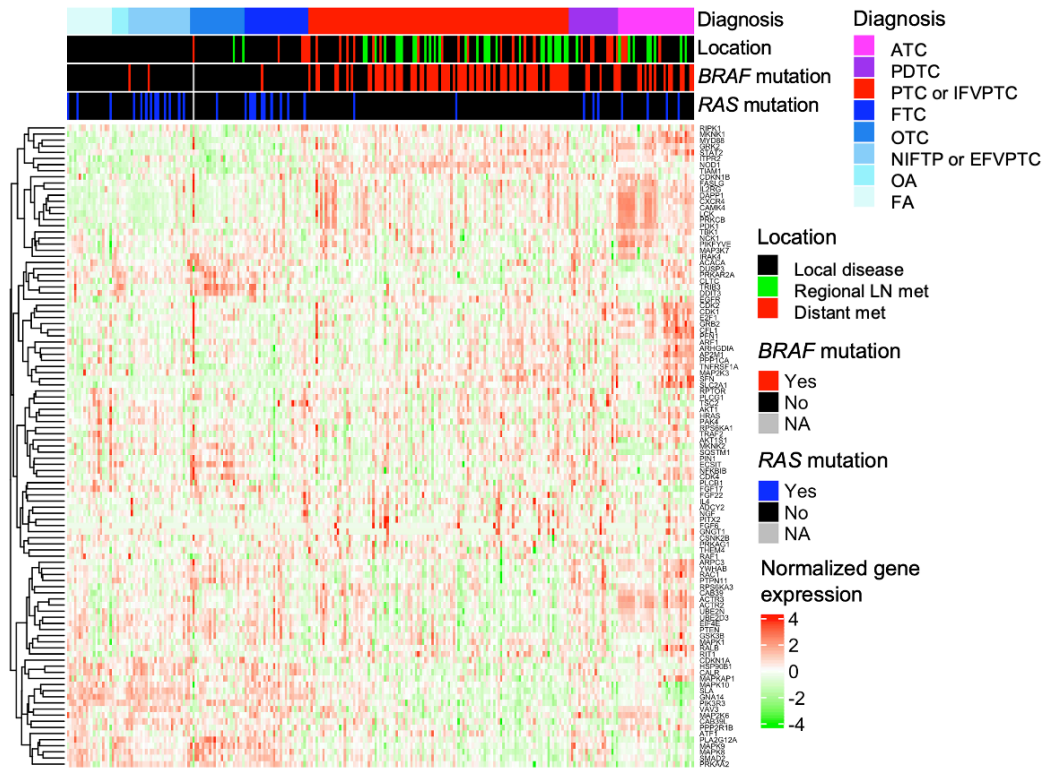
B
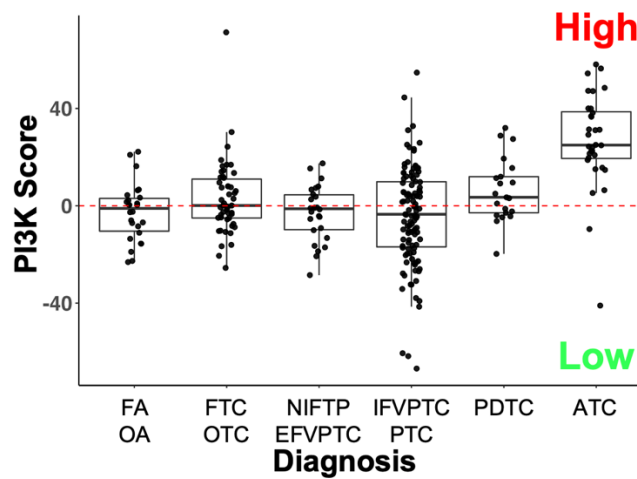
Figure 2-6: PI3K Score
A) Normalized gene expression heatmap showing 105 genes used to calculate PI3K score. Samples are ordered by diagnosis, and genes are ordered by hierarchical clustering.
B) Box plots showing PI3K score from local disease samples.

A
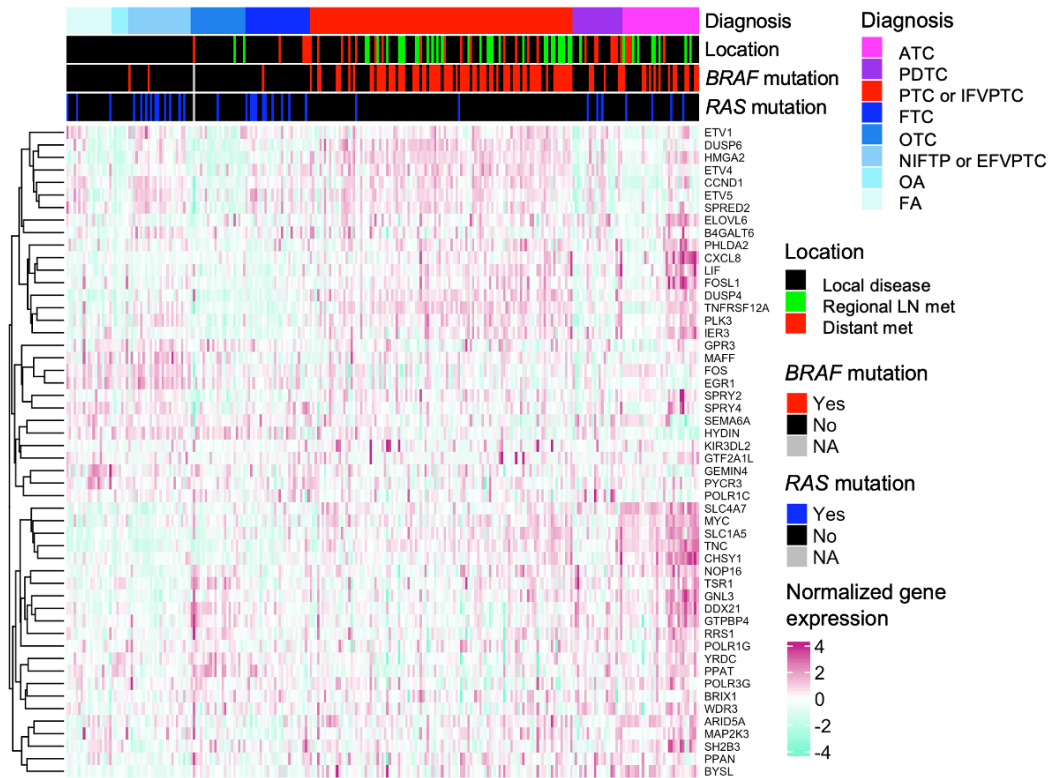
B
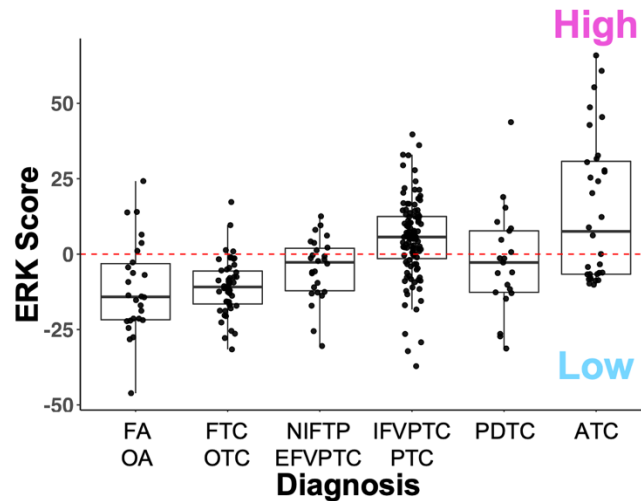
Figure 2-7: ERK Score
A) Normalized gene expression heatmap showing 48 genes used to calculate ERK score. Samples are ordered by diagnosis, and genes are ordered by hierarchical clustering.
B) Box plots showing ERK score from local disease samples.
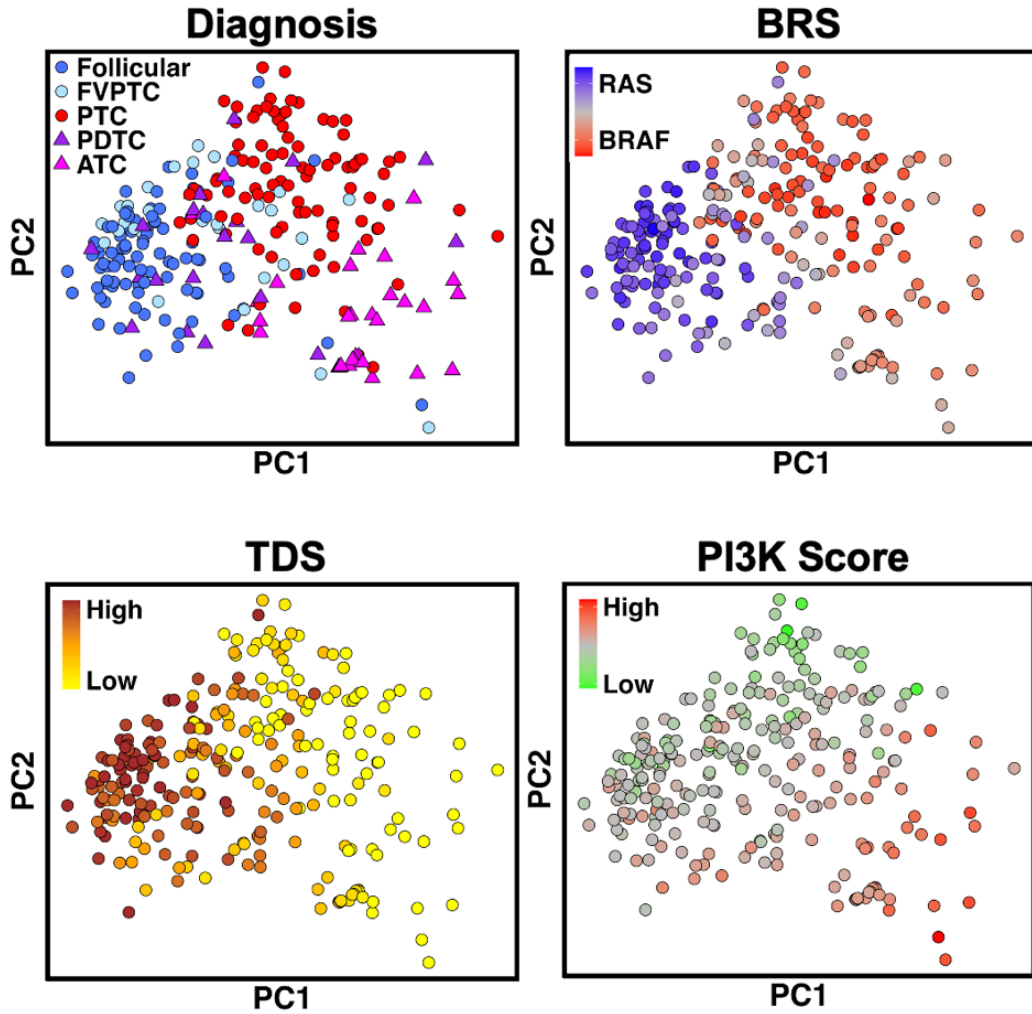
Figure 2-8: Principal component analysis
Principal component analysis (PCA) performed on bulk RNA-sequencing data from neoplastic thyroid lesions. Plots are shown colored by diagnosis, BRS, TDS, and PI3K score.

2.4 Discussion

While malignancy prediction has been the primary focus of thyroid lesion molecular profiling, recent research has made significant progress in predicting high-risk disease.[84] For example, high-risk mutations in thyroid cancer have been identified in the *TERT* promoter, *TP53*, and *PIK3CA,* often together with *BRAF* V600E.[13] However, many patients with metastatic, recurrent, or progressive thyroid cancer still lack known high-risk mutational biomarkers. In our thyroid lesion patient cohort enriched for aggressive disease, we use whole exome and bulk RNA sequencing to identify known thyroid cancer mutations, including those associated with aggressive disease. We found that *BRAF* and *RAS* mutations were associated with papillary and follicular thyroid lesions, respectively, while *TERT*p, *TP53*, and *PIK3CA* mutations were associated with decreased progression-free survival. However, our results also suggest that approximately 40% of patients with aggressive well-differentiated tumors lack these previously identified high-risk mutations. Looking at our gene expression data, we use previously published thyroid cancer scores to estimate *BRAF-RAS* gene expression, thyroid differentiation, PIK3 signaling, ERK signaling across our cohort. We observe associations between a *BRAF*-like gene expression signature, reduced thyroid differentiation marker expression, and increased PI3K and ERK signaling in our ATCs. Overall, our RNA sequencing data confirms known patterns in thyroid cancer gene expression but does not enable us to predict thyroid cancer aggression.

CHAPTER 3

Title of Chapter: A Score for Aggressive Thyroid Cancer Risk Prediction

This chapter includes adaptation of contents from the following manuscript:

Xu, G., Loberg, M., et al. Molecular Assessment of Tumor Mutations and Microenvironment Enhances Prediction of Thyroid Cancer Outcome. *Cell Genomics.* 2023.

Contributions: Matt Loberg and I contributed equally to this paper. I performed cohort collection and curation, sequencing, omics analysis and computational deconvolution methods. Matt performed RNA sequencing analysis, tissue staining, microscopic imaging, and spatial transcriptomics methods.

3.1 Introduction

The tumor microenvironment is a crucial player in the development and progression of many cancers and has great potential to inform new tools for prognostication and therapy.[85] Our previous sequencing findings suggest that known mutations and published gene expression scoring systems fall short of providing robust and comprehensive aggressive disease prediction. Our next goal was to investigate the thyroid tumor microenvironment, determine the key immune and stromal cells present, and investigate any potential for microenvironmental biomarkers to inform aggressive disease prediction. To this end, we present a gene-expression based score which we name the Molecular Aggression and Prediction (MAP) score. Our MAP score highlights markers of the tumor microenvironment and significantly predicts the development of future aggressive disease.

3.2 Methods

3.2.1 MAP Score Calculation and Gene Ontology Enrichment Analysis

Using gene expression data from our thyroid lesion patient cohort, we grouped patients by either positive BRS (*RAS*-like) or negative BRS (*BRAF*-like), and either aggressive or indolent disease. Using differential gene expression analysis, we identified 549 genes that were upregulated >4-fold in aggressive patient samples and >2-fold in BRAF-like (relative to *RAS*-like) patient samples (adjusted p value of < 0.05). For all samples in our cohort except benign subtypes MNG, FA, OA, and HT, gene expression data for each of the 549 genes was log2 transformed and Z-scores were calculated to generate our MAP scores. Enrichment analysis of the 549 MAP score genes was performed using a Panther overrepresentation test (https://pantherdb.org/).[86]

3.2.2 MAP Score Prediction Analysis

To evaluate the association between aggressive disease and predictive scores, we used logistic regression models. Penalized maximum likelihood with Jeffreys-prior penalty was used to allow for less biased and more stable estimation to account for the low number of events in some strata (R package brglm2)[87] , and area under the receiver operating characteristic curve (AUC) with corresponding 95% confidence interval (CI) were computed to evaluate the discrimination ability of a fitted model. All statistical analyses were performed using R version 4.1.2 (R Foundation, Vienna, Austria).

3.3 Results

3.3.1 MAP Score is Enriched for Extracellular Matrix, Immune, Cell Cycle, and Epithelial Differentiation Processes

As *BRAF*-like status in thyroid tumors is only slightly associated with aggressive disease, we next sought to identify genes that were upregulated in both *BRAF*-like tumors and aggressive disease. We identified 549 genes upregulated in both B*RAF*-like tumors and aggressive disease

tumors, compared to just eight genes upregulated in both RAS-like tumors and aggressive disease tumors (Table 3-1). Using the 549 genes upregulated in both B*RAF*-like tumors and aggressive disease non-metastatic tumors, we created a gene expression signature called the Molecular Aggression and Prediction (MAP) score (Figure 3-1). Comparing this score across tumors from different histologic subtypes in our cohort, we found positive MAP scores in all ATCs, the majority of PDTCs, and a portion of well-differentiated thyroid cancers, mainly from patients with aggressive disease (Figure 3-2A). We further tested our MAP score in a large external cohort of well-differentiated PTCs from the Cancer Genome Atlas (TCGA) (Figure 3-2B). In this cohort, we saw that positive MAP score correlated with aggressive PTC histologic variants, such as tall cell and diffuse sclerosing, as well as adverse pathologic features, such as extrathyroidal extension and advanced disease stage.

To better understand the biologic processes unique to aggressive MAP-positive tumors, we performed gene ontology analysis using our cohort. We found that tumors with positive MAP scores showed enrichment of biological processes including extracellular matrix, immune related processes, epithelial differentiation, and cell cycle processes. Epithelial de-differentiation and increased mitotic activity are known features associated with thyroid cancer progression and are already part of the current diagnostic criteria for many aggressive thyroid tumors.[88-90] However, biologic processes involving tumor microenvironment remodeling are less well-studied in thyroid cancer and represent an important area for future research.
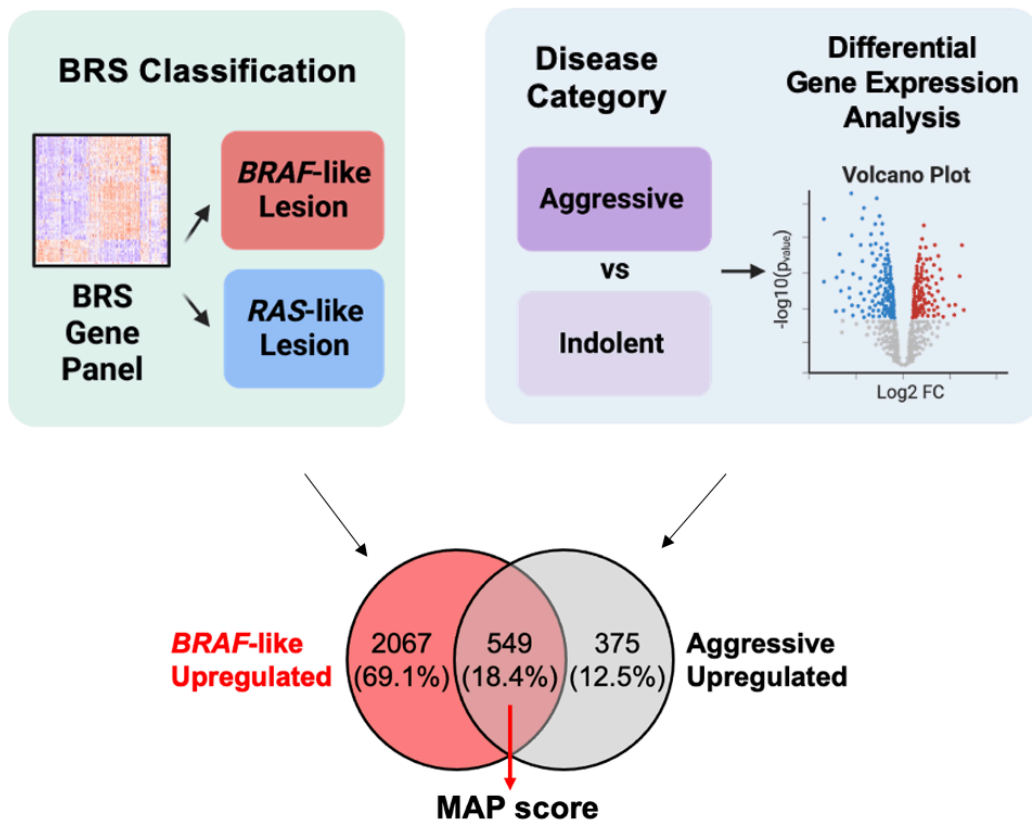
Figure 3-1: MAP score development
Diagram summarizing creation of the MAP score. Samples were categorized as either *BRAF*-like or *RAS*-like based on the expression of a previously published list of 69 genes (see section 2.3.3), and samples were categorized as aggressive or indolent based on patient disease. Differential gene expression analysis was performed to determine genes upregulated in *BRAF*-like and aggressive non-metastatic samples, and genes present in the overlap of both groups were used for the MAP score. Components of this figure were generated using BioRender.

RAS-Aggression Overlap Genes							
BRS3	EEF1A2	ISL1	NRSN1	OR4D5	OR5H2	OR6T1	OR10S1

BRAF-Aggression Overlap Genes							
ACOD1	ACP7	ACTA1	ACTC1	ACY3	ADAM19	ADAM2	ADAM29
ADAM7	ADAMTS14	ADGRF4	AFP	AGXT	AICDA	AIM2	ALDH3B2
ALPK2	ANKRD1	ANLN	ANPEP	ANXA8	ANXA8L1	APCDD1L	AQP9
ARL14EPL	ARL4C	ARNTL2	ARPP21	ARSI	ASB10	ASCL4	ASPM
ATP12A	ATP1A4	AURKA	AURKB	AVP	AZGP1	B4GALNT4	BATF
BCAS1	BEND4	BEST3	BIRC5	BPIFC	BUB1	C10orf71	C12orf75
C4orf17	C5orf46	C9	CA9	CALB2	CALHM6	CALML3	CBLN2
CCBE1	CCDC185	CCDC197	CCKAR	CCL1	CCL11	CCL13	CCL20
CCL24	CCNA2	CD19	CD2	CD300E	CD3E	CD7	CD70
CDC20	CDC25C	CDCA2	CDCA5	CDCA8	CDHR1	CDK5R2	CDX2
CEACAM3	CEACAM7	CENPA	CENPE	CENPF	CENPI	CEP55	CER1
CHAT	CHRNA1	CIB3	CKAP2L	CKM	CLC	CLCA2	CLDN14
CLEC4C	CLEC4M	CLVS2	CNGA2	COL11A1	COL1A1	COL7A1	CORO1A
COX6A2	CPA4	CPB1	CPN2	CPNE7	CR1	CR2	CREG2
CRYBA2	CSF2	CSF3	CSN1S1	CSRP3	CST8	CXCL1	CXCL13
CXCL3	CXCL5	CXCL6	CXCL8	CYP11B1	CYP27C1	CYP4F2	CYP4F22
DAZL	DDN	DEFB119	DEFB121	DEPDC1	DIAPH3	DKK1	DLGAP5
DNAJC12	DNER	DNMT3L	DRD2	DRGX	DSC3	DSCAM	DSG1
DSG3	DUSP9	DYNAP	E2F7	E2F8	EREG	ERICH3	ESPL1
EXD1	EXO1	FAM216B	FAM83A	FAM83D	FAM9A	FAM9B	FANCD2OS
FCAMR	FCRL3	FCRLA	FGF5	FOXA3	FOXD1	FOXE3	FOXG1
FOXI1	FOXR2	FPR2	FSD1	FST	FXYD3	GABRA3	GADL1
GALNTL5	GC	GDNF	GFRAL	GJA8	GJB2	GOLGA6L1	GOLGA6L22
GOLGA6L6	GOT1L1	GPR153	GPRIN1	GREM1	GSDMA	GSDMC	GTSE1
GTSF1L	H1-5	H2AC16	H2AC4	H2BC17	H3C12	H3C2	H3C3
HAS2	HCAR2	HCAR3	HDGFL1	HES2	HJURP	HMMR	HMX3
HNF4G	HOXB9	HOXC10	HOXD10	HOXD11	HOXD12	HS3ST3A1	HTN3
HTR3A	HTR7	IBSP	IFNA8	IFNL1	IGF2BP1	IL11	IL19
IL1A	IL1F10	IL2	IL21	IL24	IL2RA	IL3	IL31RA
IL36B	IL36G	IQCJ	IQGAP3	IVL	IZUMO3	KIAA1549L	KIF14
KIF15	KIF18B	KIF20A	KIF23	KIF4A	KIFC1	KLF18	KLK5
KLK6	KLK8	KRT1	KRT13	KRT14	KRT15	KRT16	KRT17
KRT31	KRT33A	KRT36	KRT4	KRT5	KRT6A	KRT6C	KRT74
KRT75	KRT79	KRTAP2-3	KRTAP3-2	KRTAP4-7	KRTAP7-1	L1CAM	LACTBL1
LAIR2	LAMA3	LBP	LCE1B	LCE2A	LCE5A	LCE6A	LGALS14
LHX1	LIM2	LORICRIN	LPAR3	LRRC30	LRRC38	LY6D	LY6L
LYPD3	LYPD6B	LYPD8	MAGEA10	MAGEA11	MAGEA3	MAGEB2	MB
MC5R	MCHR1	MDFIC2	MELK	MEPE	MKI67	MME	MMP1
MMP10	MMP11	MMP12	MMP13	MMP27	MMP3	MMP7	MSLNL
MUC7	MUCL1	MYBL2	MYBPC1	MYBPC2	MYEOV	MYF6	MYH13
MYH2	MYH7	MYH8	MYL1	MYL2	MYO18B	MYO1G	MYOD1
MYOG	MYPN	NCAN	NCAPG	NCAPH	NDC80	NEIL3	NEK2
NGB	NIPAL4	NLRP10	NLRP4	NLRP5	NPHS1	NPHS2	NRAP
NT5DC4	NTNG1	NTRK1	NXPH2	NYX	OBP2A	OOSP1	OOSP2
OPN1LW	OPRPN	OR10A3	OR10W1	OR12D2	OR14J1	OR1D2	OR4A5
OR4C11	OR4D2	OR4D9	OR4F15	OR4F17	OR4K2	OR4N2	OR52A1
OR52M1	OR52N5	OR52R1	OR5AK2	OR5BS1P	OR5H1	OR5K4	OR5P2
OR6C1	OR6K6	OR8B4	OR8G5	OR8U3	OTOP2	OTP	OTX2
P2RX5	PADI6	PAEP	PAGE2	PAGE4	PAQR9	PAX5	PBK
PDPN	PGLYRP3	PI15	PI3	PITX1	PITX2	PKP1	PLA2G2F
PLEKHS1	PLPP4	PMAIP1	POLQ	POSTN	POTEF	POU4F2	PRAME
PRAMEF1	PRAMEF14	PRAMEF18	PRAMEF2	PRAMEF20	PRAMEF9	PRDM13	PRDM8
PRDM9	PRG3	PRL	PRODH2	PRSS21	PRSS3	PRSS56	PTHLH
PTPRN	PTPRZ1	PTX3	RAB3B	RDH8	REG3A	RESP18	RFPL4AL1
RGS21	RGS7	RHOV	RNASE3	ROS1	RRM2	RTBDN	RTL3
S100A2	S100A7	S100A8	S100A9	SAA1	SAA2	SCGB3A2	SCRT2
SEMA7A	SERPINA4	SERPINA7	SERPINB13	SERPINB3	SERPINB4	SERPINB5	SERPINB7
SFN	SFTPA1	SFTPA2	SH2D1A	SHCBP1	SIGLEC14	SKA1	SLC13A2
SLC17A8	SLC1A6	SLC25A48	SLC35D3	SLCO1A2	SLITRK1	SLN	SMCO1
SMIM31	SMPX	SMYD1	SOST	SOX11	SP8	SPACA3	SPATA31D1
SPC24	SPO11	SPOCD1	SPRR2A	SPRR2D	SRD5A2	STEAP1	STRA8
STRIT1	SULT1E1	SVOP	SYT8	TAAR2	TAC3	TAF11L11	TAFA4
TARM1	TAS1R2	TBX20	TCN1	TEX13A	TEX33	TFAP2A	TFDP3
TFPI2	TGM5	TMEM158	TMEM207	TMEM40	TNFAIP6	TNIP3	TNNI2
TNNT1	TNR	TOP2A	TP63	TPX2	TRIM46	TRIML1	TRIML2
TRIP13	TROAP	TTK	TTPA	TUBA3C	TWIST1	TXNDC8	UBE2C
UCN2	UGT2B15	UGT3A1	UHRF1	UNC45B	URAD	VCAN	VRTN
WFDC9	WNT2	WNT7B	WT1	XAGE5	XDH	XIRP2	ZIC1
ZIC2	ZIC4	ZNF365	ZP4	ZPBP2			

Table 3-1: *RAS*- and *BRAF*-Aggression Overlap Genes
This table lists the eight genes that are upregulated in both *RAS*-like and aggressive non-metastatic samples, as well as the 549 genes that are upregulated in both *BRAF*-like and aggressive non-metastatic samples.

A
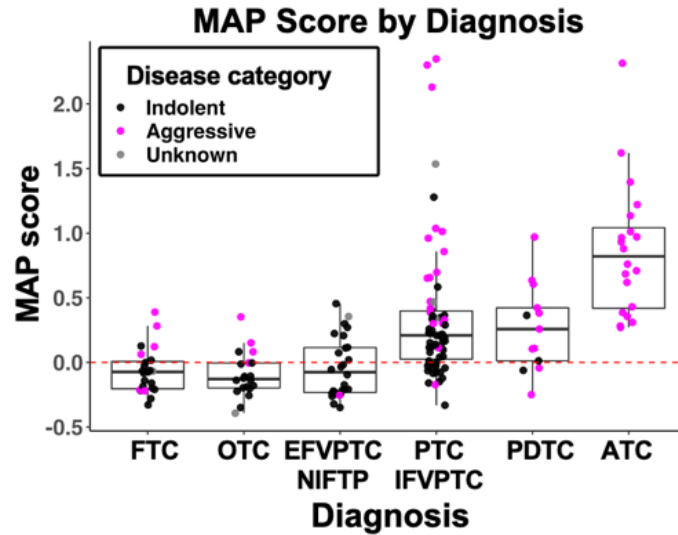
B
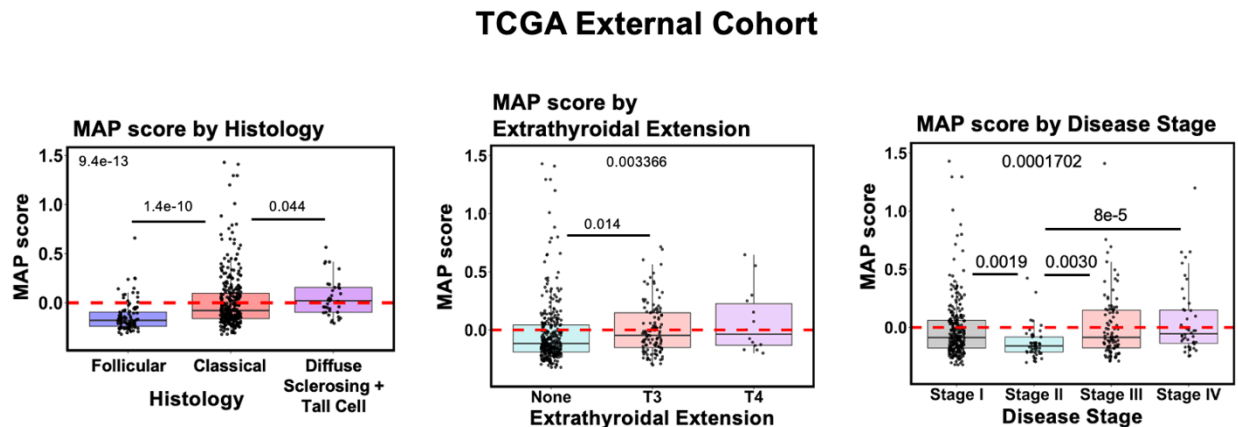
Figure 3-2: MAP scores across our cohort and an external thyroid cancer cohort

A) Box plots showing MAP score, calculated using the 549 genes that overlap between *BRAF*-like and aggressive lesions, and diagnosis. Pink dots label lesions from patients with aggressive disease.

B) Box plots using MAP scores in an external cohort of well-differentiated thyroid tumor samples from TCGA. TCGA samples are plotted by histology, extrathyroidal extension, and disease stage. Three outliers for MAP score were omitted for improved plot visualization. P values were calculated using the Kruskal-Wallis test with pairwise Wilcoxon rank-sum test and Bonferroni's correction.
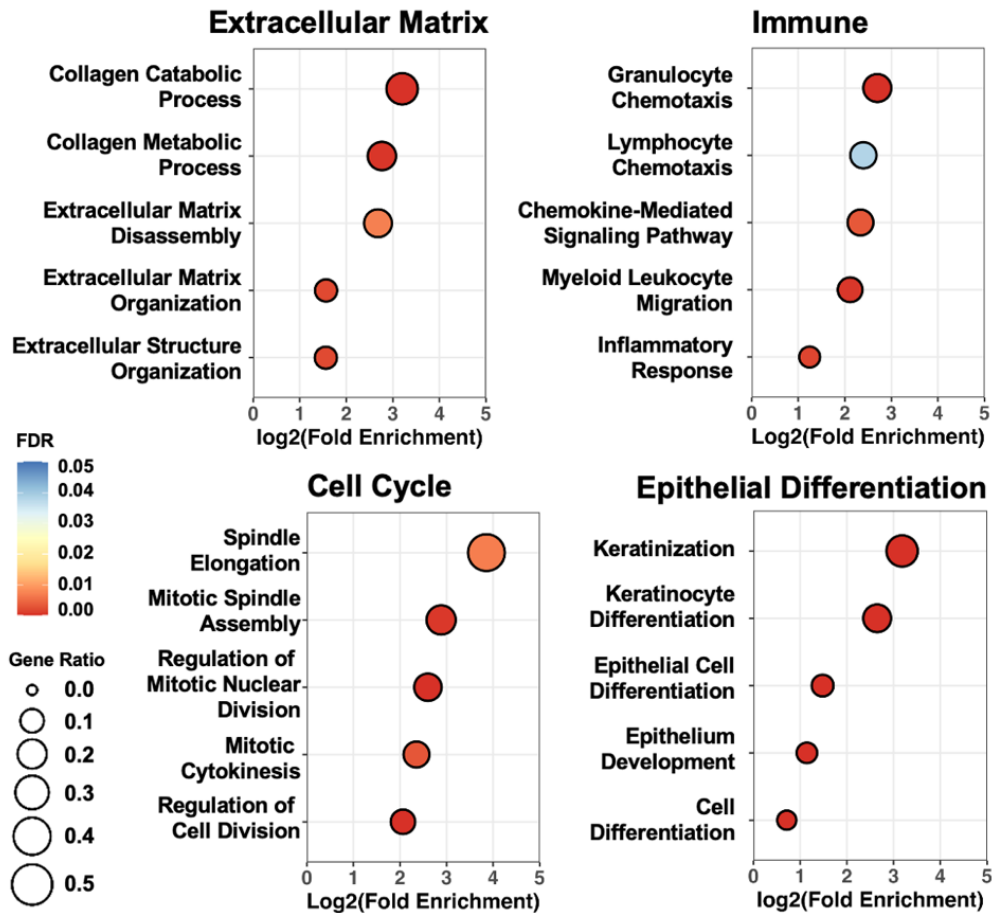
Figure 3-3: Gene ontology analysis of our thyroid cancer cohort
Gene ontology results for the 549 genes comprising the MAP score. Results show enrichment of extracellular matrix, immune, cell cycle, and epithelial differentiation processes. Statistical analysis of fold enrichment was performed using the Fisher's exact with false discovery rate correction.

3.3.2 MAP Score is Associated with Shorter Survival and Predicts Higher Risk of Aggressive Disease

We next tested whether MAP score could be used to predict aggressive disease. Survival analysis using our cohort with both well-differentiated and transformed tumors, as well as with well-differentiated tumors alone, showed that patients with positive MAP scores had significantly shorter progression-free survival and overall survival (Figure 3-4A,4B). We also performed survival analysis using the external TCGA cohort (Figure 3-4B). Despite only containing well-differentiated PTCs and not enriched in aggressive disease like our own cohort, our results for TCGA also showed significantly worse survival for positive MAP score samples.

To assess the ability of our MAP score to predict aggressive disease, we also used generalized linear models with penalized maximum likelihood estimation. In our testing, we assessed score performance using three different groups of local disease malignant samples: all malignant samples, well-differentiated malignancies, and well-differentiated malignancies that were resected prior to any evidence of disease progression, with the final group allowing for testing of future aggressive disease prediction. In our receiver-operator curves (ROCs), we compare the predictive ability of our MAP scores against and in combination with three common high-risk mutations: *TP53*, *TERTp*, and/or *PIK3CA* (Figure 3-5). Our results suggest that both our MAP score and mutation score performed similarly and were significant predictors of aggressive disease, and combining both scores provided the greatest predictive power as measured by area under the curve (AUC). We also found that MAP score provided aggressive disease prediction in samples lacking these known high-risk mutations.
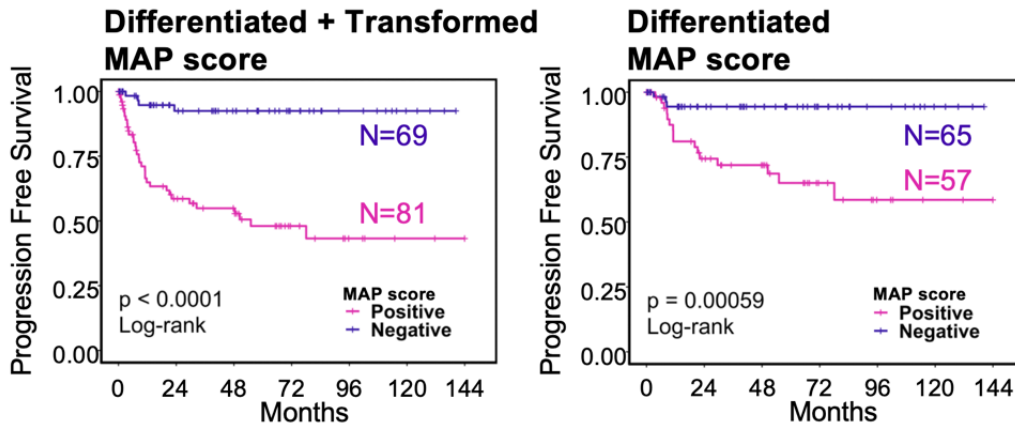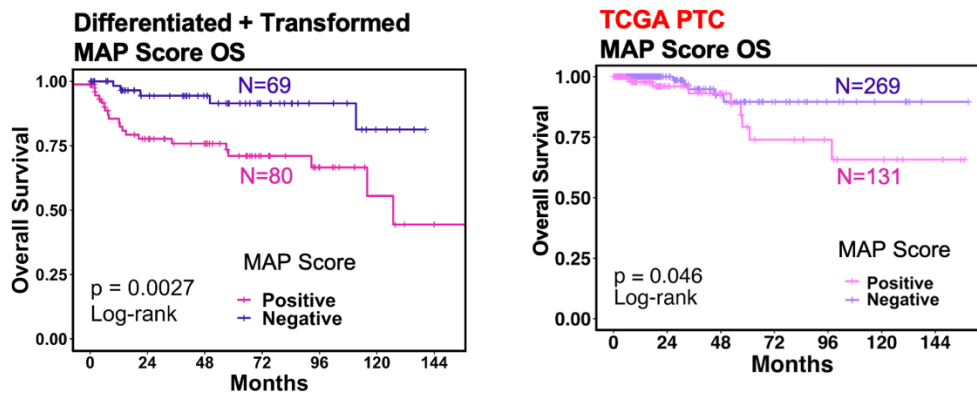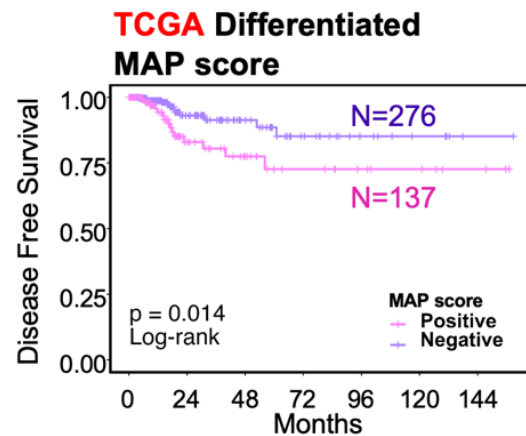
Figure 3-4: MAP score and survival
A) PFS results for patients with well-differentiated and transformed thyroid cancer (left), as well as patients with only well-differentiated thyroid cancer (right). Patients are grouped as either having positive MAP score (pink) or negative (purple) MAP score, and P values were calculated by log-rank test.
B) OS results for patients with well-differentiated and transformed thyroid cancer (left), as well as patients with only well-differentiated thyroid cancer (right). Patients are grouped as either having positive MAP score (pink) or negative (purple) MAP score, and P values were calculated by log-rank test.
C) Table with percentage of distant metastatic samples in the internal cohort from VUMC and UWMC and the external cohort from TCGA.
D) Disease-free survival in TCGA patients with well-differentiated thyroid cancer. Patients are grouped as either having positive MAP score (pink) or negative (purple) MAP score, and P values were calculated by log-rank test.

A
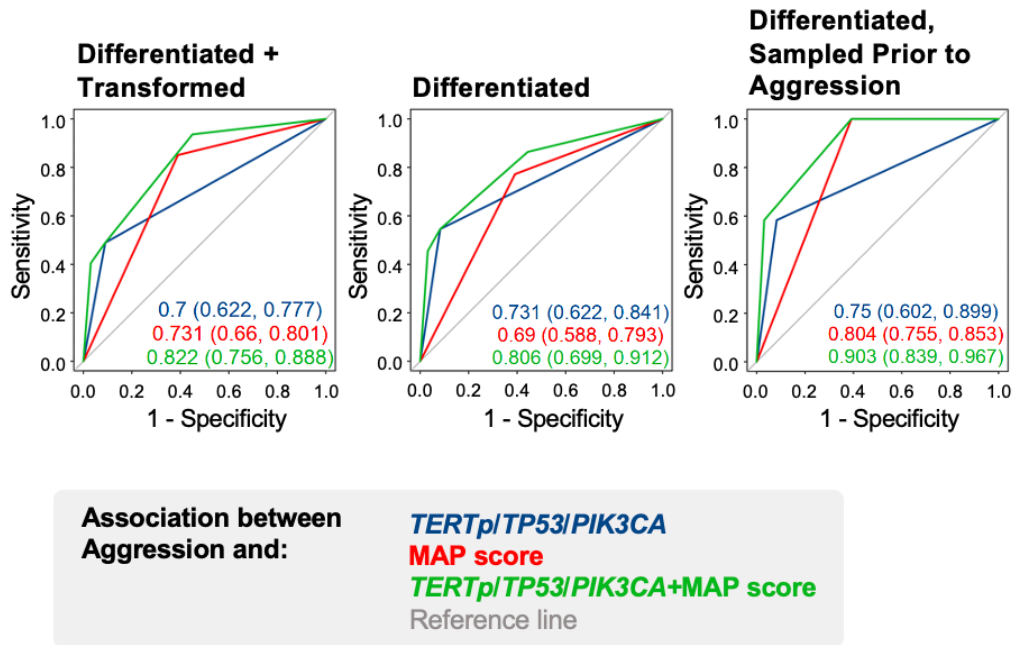
B
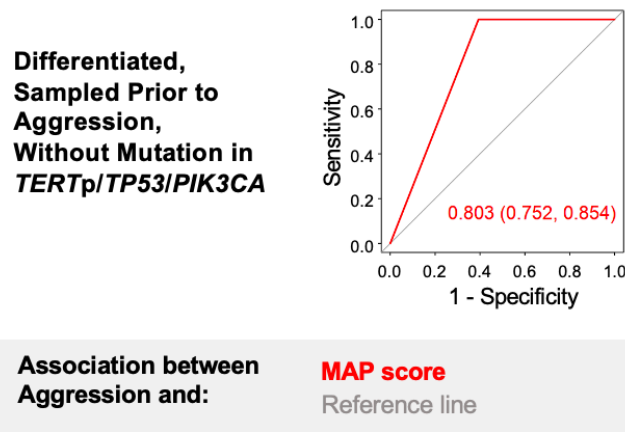
Figure 3-5: MAP score and aggression prediction
A) Receiver operating characteristic (ROC) curves showing association between aggression and TERTp/TP53/PIK3CA mutation (blue), MAP score (red), and TERTp/TP53/PIK3CA mutation + MAP score (green). ROC results include patients with both well-differentiated and transformed thyroid cancer (left), patients with only well-differentiated thyroid cancer (center), and patients with only well-differentiated thyroid cancer sampled prior to aggression (right). Metastatic tumors were excluded from analysis. Area under the curve values with 95% confidence intervals are shown.
B) Receiver operating characteristic (ROC) curve showing association between aggression and MAP score (red), for just patients with only well-differentiated thyroid cancer sampled prior to aggression excluding any samples with a mutation in *TERTp*, *TP53*, and *PIK3CA*. Metastatic tumors were excluded from analysis, and area under the curve values with 95% confidence intervals are shown.

3.4 Discussion

In summary, our MAP score highlights the tumor microenvironment as a potentially important factor in thyroid cancer aggression. Genes enriched in this score included those important for ECM processes, including matrix assembly and collagen metabolic/catabolic processes, as well as immune related processes, such as cytokine signaling, inflammatory response, and immune cell migration. Our MAP score is novel compared to current thyroid molecular tests for aggression, as most are either primarily mutation-based or have not focused on markers of the tumor microenvironment.[44,48,84] We show that this score is associated with patient survival in both our cohort and in an external cohort of well-differentiated thyroid cancer patients (TCGA). We additionally show that this score, in combination with known high-risk mutations, improves prediction of thyroid cancer aggression. The MAP score may also provide outcome prediction for patients lacking known high-risk mutations; however, due to limitations in sample number, we were unable obtain statistically significant results for this test. In addition, we acknowledge that our findings used sequencing data from cells collected by surgical resection, so additional research is needed to show that such tests retain the ability to predict aggressive disease using samples taken by less invasive biopsy techniques such as FNA. However, if these limitations can be addressed, we propose that this score or a similar score evaluating markers from the tumor microenvironment could someday be applied clinically to profile tumors, predict future aggressive disease, and inform potential avenues for therapy. In summary, our findings highlight the importance of non-mutational biomarkers for revealing thyroid cancer behavior and suggest a need for more tumor microenvironment research in order to advance classification and aggression prediction for thyroid tumors.

CHAPTER 4:

Title of Chapter: Characterization of the Thyroid Tumor Microenvironment

This chapter includes adaptation of contents from the following manuscript:

Xu, G., Loberg, M., et al. Molecular Assessment of Tumor Mutations and Microenvironment Enhances Prediction of Thyroid Cancer Outcome. *Cell Genomics.* 2023.

Contributions: Matt Loberg and I contributed equally to this paper. I performed cohort collection and curation, sequencing, omics analysis and computational deconvolution methods. Matt performed RNA sequencing analysis, tissue staining, microscopic imaging, and spatial transcriptomics methods.

4.1 Introduction

The tumor microenvironment, including the extracellular matrix (ECM), signaling molecules, and non-tumor cells such as immune cells and stromal cells, is thought to play a critical role in disease progression. While there has been limited research on the microenvironment of thyroid cancer, a recent study reported a subgroup of *BRAF*-like lesions enriched in cancer-associated fibroblasts (CAFs) that may have more aggressive behavior.[91] Another study investigating thyroid tumors driven by *BRAF* V600E mutation and *PTEN* loss found that fibroblasts may remodel collagen in the tumor microenvironment to promote progression.[92] Macrophages and other infiltrating immune cells have also been implicated in supporting the aggressive behavior of thyroid cancer.[38,39] Given new and expanding research on the tumor microenvironment, it is not surprising that clinical trials are now evaluating checkpoint inhibitor therapy for the most aggressive thyroid cancer subtype, anaplastic thyroid cancer.[93-98] To further examine the immune and stromal microenvironment of thyroid cancer, we next evaluate the immune composition of thyroid tumors in our cohort, as well as the potential for response to immunotherapy.

4.2 Methods

4.2.1 TIMER 2.0

To estimate levels of infiltrating immune and stromal cells, we applied bulk RNA sequencing data to the online computational deconvolution tool TIMER2.0 (http://timer.cistrome.org/) (Figure 4-1), using THCA (Thyroid Carcinoma) as the cancer type gene signature. TIMER 2.0 immune deconvolution scores used in our study include those from CIBERSORT-Abs[73], EPIC[99], and MCPCOUNTER[74], and descriptive results were plotted using the R package ggplot2.[75] TIMER scores were plotted in heatmaps using R package ComplexHeatmap.[61]
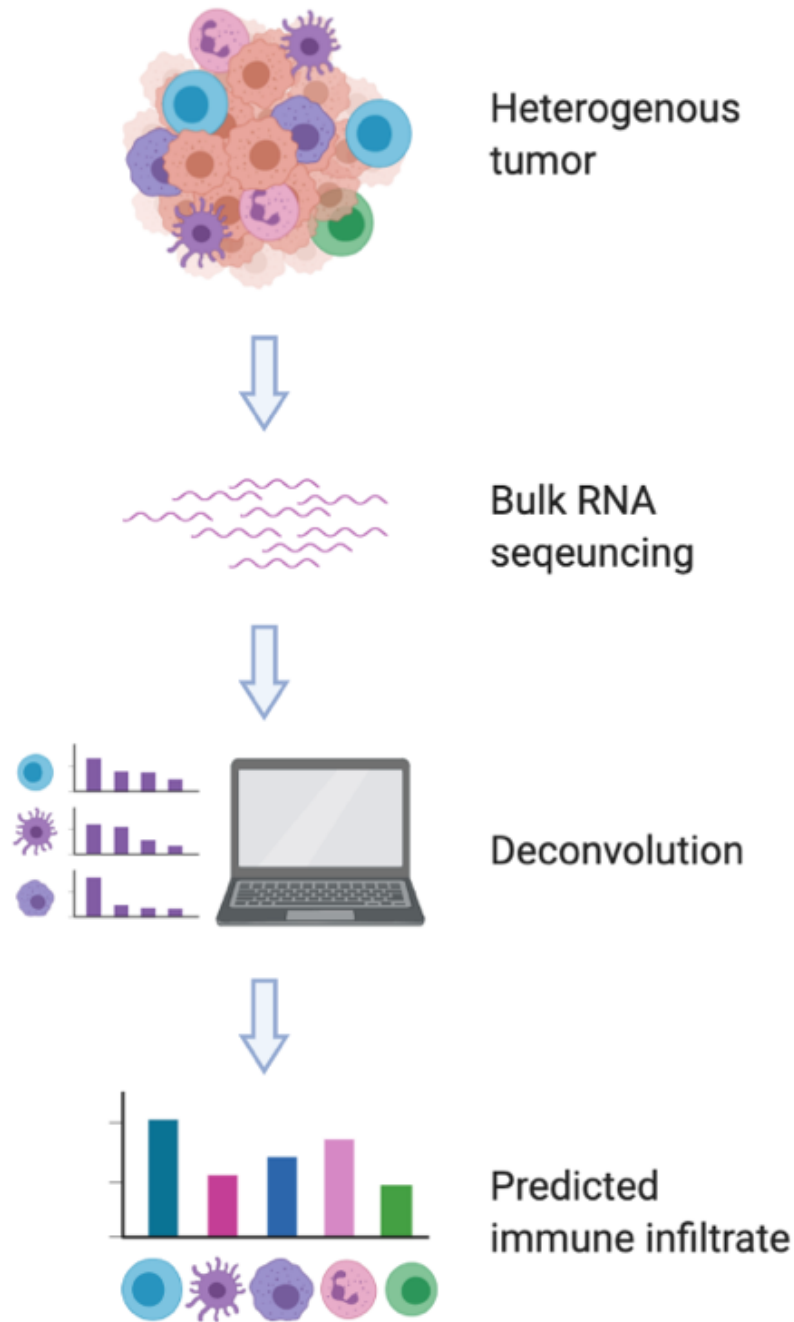
Figure 4-1: Overview of computational deconvolution process
This diagram summarizes the steps taken to generate our computation deconvolution data. First, heterologous tumor samples are processed to undergo bulk RNA sequencing. Differential gene expression data from bulk RNA sequencing is used as the input by computational deconvolution programs. These programs use known gene expression profiles to predict levels of each immune cell type in the tumor. This figure was generated using BioRender.

4.2.2 TIDE

To estimate immune checkpoint response using gene expression data, we used the computational tool TIDE (http://tide.dfci.harvard.edu/),[76] with the following settings: Cancer type = Other, Previous Immunotherapy = No. TIMER scores were plotted in heatmaps using R package ComplexHeatmap.[77]

4.2.3 Tissue Staining

To collect multiplex immunofluorescence (IF) data, five μm ATC tissue sections were cut from 33 formalin-fixed paraffin-embedded (FFPE) tissue blocks and stored at -20°C. Tissue sections were thawed at room temperature overnight and heated at 60°C for 1 hour. Tissue sections were deparaffinized with xylene (2x15 minutes), ethanol (100% 2x5 minutes, 95% 1x5 minutes), and water (5 minutes), and then washed with PBS. Antigen retrieval was performed by heating slides for 45 minutes in sodium citrate buffer (pH 6.0) in a rice cooker, and then cooling at room temperature for 30 minutes. Tissues were washed with PBS before blocking for 2 hours with 10% goat serum in PBS (blocking buffer). Primary antibodies (Abcam ab207178 recombinant rabbit monoclonal anti-fibroblast activation protein alpha (FAP) IgG, clone EPR20021, 1:100; Invitrogen MA5-16868 rat monoclonal anti-MRC1 IgG2a, clone MR5D3, 1:25) were diluted in blocking buffer before incubating on tissue sections at 4°C for 16 h (Abcam, Cambridge, UK; Thermo Fisher, Waltham, MA). Tissue sections were washed using 0.05% Tween 20 in PBS. Secondary antibodies (Invitrogen A-21245 polyclonal goat anti-rabbit IgG alexa fluor 647 1:150; Abcam ab6953 polyclonal goat anti-rat IgG Cy3 1:150) as well as conjugated primary antibodies (eBioscience 53-9003-82 mouse monoclonal anti-pan cytokeratin IgG1 AF488, clone AE1/AE3, 1:100) were diluted in blocking buffer containing Hoechst 33342 nuclear stain (1:1000) and incubated on tissue sections at 37°C for 1 h (Abcam, Cambridge, UK; Thermo Fisher, Waltham, MA). 12 representative 20X and 12 representative 60X images were captured for each tissue section using a Nikon Spinning Disc confocal microscope.

To analyze multiplex immunofluorescence (IF) data, representative images were blindly scored by a practicing pathologist (Dr. Vivian Weiss, VW). Each ATC tissue section was scored for intensity (0-3) and frequency (0-3) of FAP staining of non-malignant cells, and an overall FAP staining score was calculated as the product of the intensity and frequency scores (0-9). FAP staining scores of 0-1 were categorized as low, and FAP staining scores of greater than 1 were categorized as high. Non-malignant MRC1 stained cells were counted on 12 20X images for each tissue section. An average of less than 1 MRC1+ cell per 20X field was categorized as low, and an average greater than 1 MRC1+ cell per 20X field was categorized as high.

4.2.4 Spatial Transcriptomics

To perform spatial transcriptomics, the Visium FFPE platform was used to generate spatial transcriptomics data (10x Genomics, Pleasonton, CA). 8 FFPE block thyroid carcinomas with ATC histology were selected for analysis. Following pathologist review (VW), 5 μm sections up to 6 mm x 6 mm in size were cut onto a Visium Gene Expression Slide (Visium Spatial Gene Expression Slide Kit, PN-1000188) to capture the transition zone between PTC and ATC. After sectioning, the slide was incubated at 42°C and then stored in a desiccator.

Using manufacturer's protocols (Visium FFPE 10X Genomics), samples were deparaffinized, stained (hematoxylin and eosin), and scanned at 20X. Due to a supply chain shortage of eosin, 3 of 8 ATCs were stained with hematoxylin only. The Visium Human Transcriptome Probe Set v1.0 was hybridized to samples overnight at 50°C, and following RNA digestion and tissue permeabilization, sequencing libraries were prepared according to manufacturer's protocols. Using the NovaSeq 6000 platform (Illumina, San Diego, CA), sequencing was performed at a depth of >40,000 reads per spot and >150 million reads per sample.

To analyze spatial transcriptomics data, Visium sequencing data was pre-processed with Space Ranger 2.0.0 (10X Genomics), and analysis of Space Ranger outputs was performed with

Seurat 4.0.[100] Seurat 4.0 was used to perform normalization, dimensionality reduction, and clustering. Dimensionality was determined by elbow plot, and for clustering, resolution was set to 0.2. PTC and ATC histology designation was conducted by pathologist review (VW). Deconvolution of immune cell frequencies within individual capture areas was performed using the R package SpaCET,[101] and to determine capture area malignant cell fraction, the SpaCET PANCAN setting was used. ATC classification was based on current standard-of-care clinical practice as defined by the WHO and ATA guidelines and was reviewed by a practicing pathologist (VW). Individual Visium samples were classified as high MAP score or moderate MAP score based on the MAP score of the associated bulk RNA sequencing sample.

4.3 Results

4.3.1 Computational Deconvolution of the Tumor Microenvironment

Using our cohort's bulk RNA sequencing data, we performed differential gene expression comparing samples with positive vs negative MAP scores (to separate samples based on predicted outcome) and found that MAP-positive samples were enriched in inflammatory genes (Figure 4-2A). These genes included markers of immune cell infiltrates such as CAFs, M2 macrophages, and neutrophils (Figure 4-2B).

To further investigate the thyroid tumor microenvironment, we used the computational deconvolution tool TIMER, which uses bulk RNA sequencing data to make estimations of infiltrating immune cells as well as a few other non-tumor cells.[72] Using TIMER, we confirmed that CAFs, M2 macrophages, and neutrophils are significantly higher in MAP-positive tumors (Figure 4-3A). Looking at a wider range of TIMER scores in a heatmap (Figure 4-3B), we observe overall higher infiltration of immune cells in our PTC and ATC samples compared to our FTC and PDTC samples.

In addition, we applied TIMER to a large external cohort of well-differentiated PTC samples in the TCGA (Figure 4-4). While the TCGA cohort only contains PTCs, and no ATCs, we observed trends comparable to those we saw in our own cohort, with enrichment of CAFs, neutrophils, and M2-macrophages in MAP-positive tumors, and enrichment of CD8+ T cells in MAP-negative tumors.
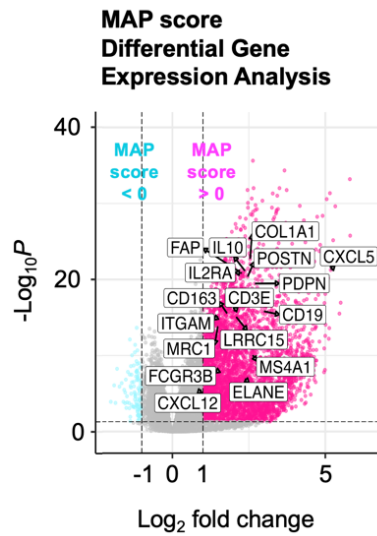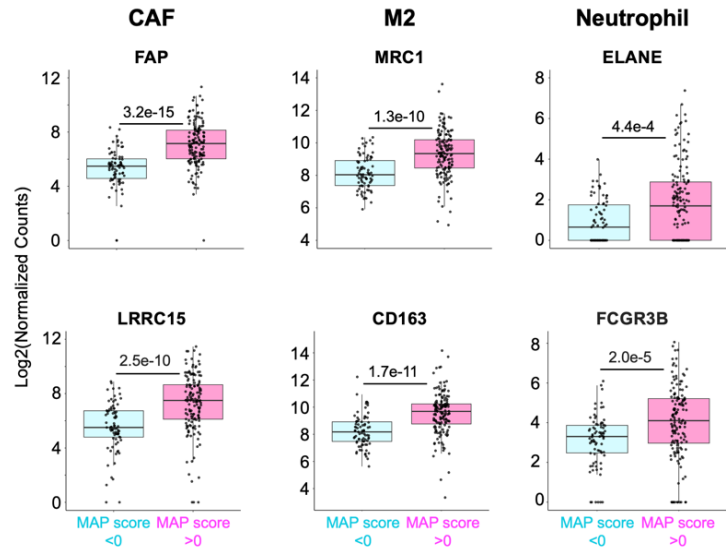
Figure 4-2: MAP score is associated with cancer associated fibroblast (CAF) and immune infiltrate gene expression

A) Volcano plot showing differentially expressed genes between malignant localized thyroid lesions with either positive (pink) or negative (blue) MAP score. Dotted lines indicate fold-change > 2 and adjusted p value < 0.05. Samples with Hashimoto Thyroiditis were excluded from this analysis. Markers of extracellular matrix, CAFs, and key immune cell populations are labeled.

B) Box plots showing log-transformed gene expression of select cell markers in malignant thyroid lesions with either negative MAP score (light blue) or positive MAP score (dark pink). Hashimoto Thyroiditis samples were excluded from this analysis. Results include CAF markers *FAP* and *LRRC15*, M2 macrophage polarization markers *MRC1* and *CD163*, and neutrophil markers *ELANE* and *FCGR3B*. P values were calculated using the Wilcoxon rank-sum test.
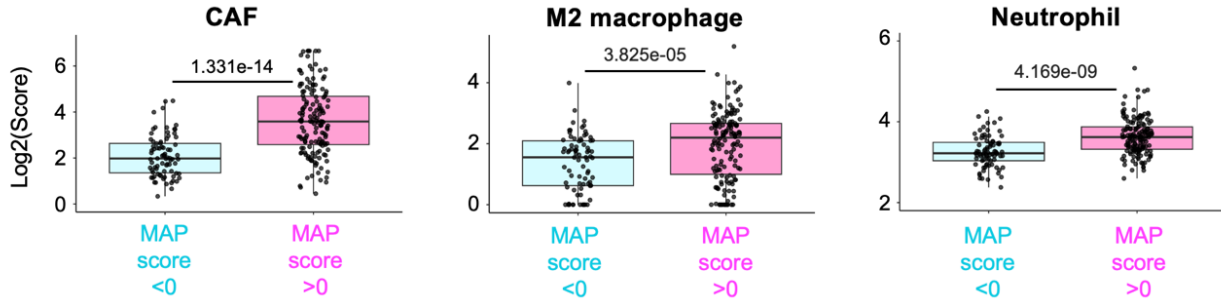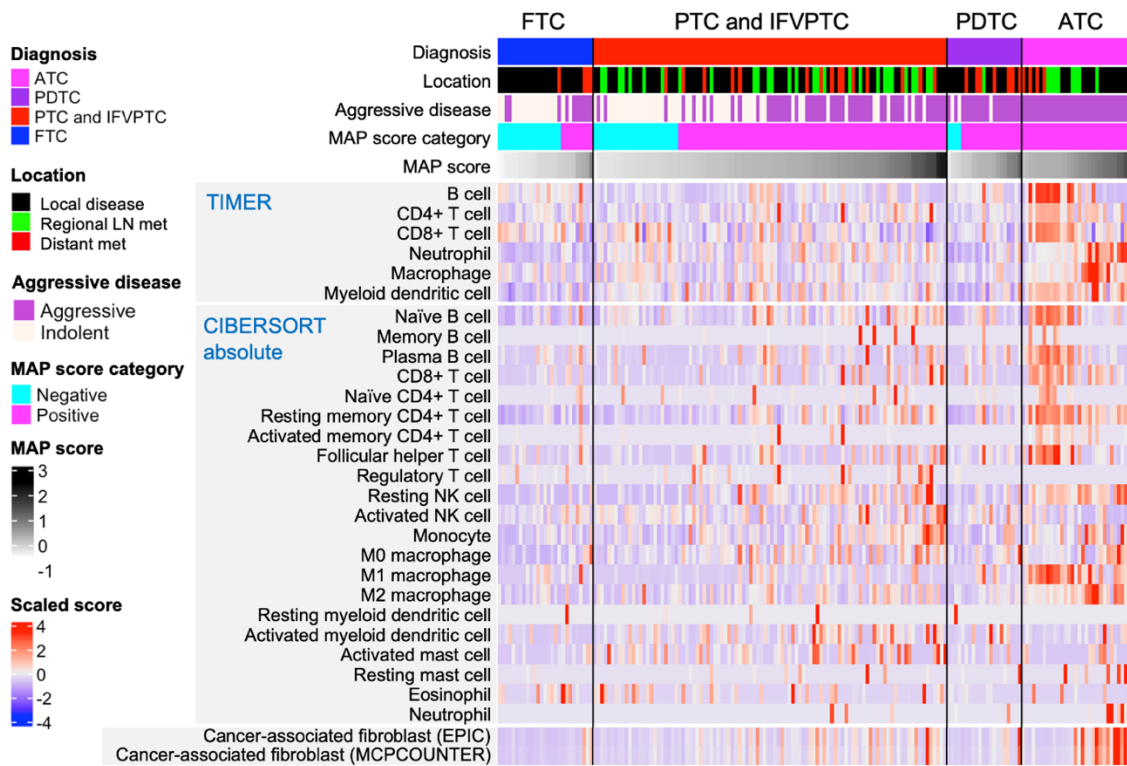
Figure 4-3: TIMER scores in our internal thyroid cancer cohort
A) Box plots showing log-transformed EPIC CAF score, CIBERSORT absolute value M2 macrophage score, and TIMER neutrophils score, for malignant thyroid lesions categorized as either negative MAP score (light blue) or positive MAP score (dark pink). Hashimoto Thyroiditis samples were excluded. P values were calculated with Wilcoxon rank-sum test.
B) Heatmap of TIMER deconvolution results. Diagnosis, tissue location, aggressive disease, MAP score category, and MAP score annotations are displayed on the top of the heatmap, followed by sample location, TIMER and absolute value CIBERSORT (CIBERSORT-Abs) immune deconvolution scores, and CAF scores estimated by EPIC and MCPCOUNTER. Samples are sorted by increasing MAP score from left to right within each diagnosis. PTC samples from patients with Hashimoto Thyroiditis (HT) were excluded.
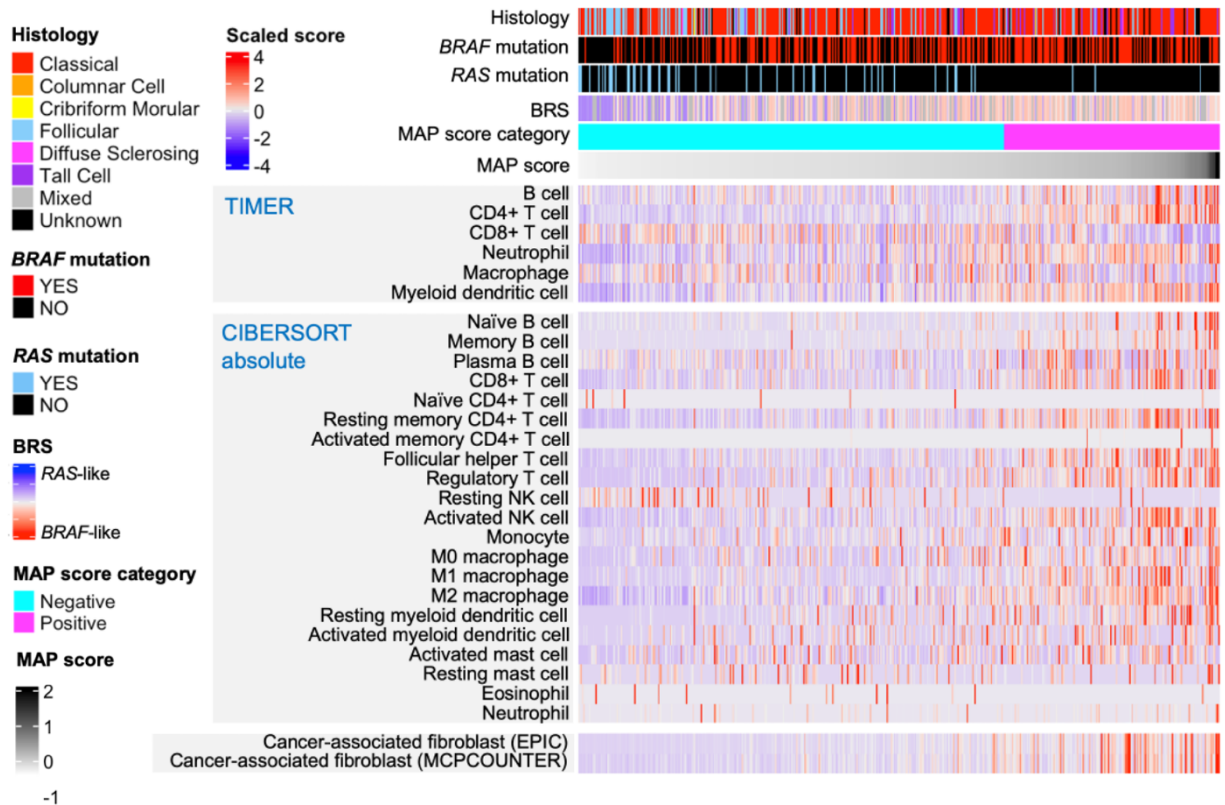
Figure 4-4: TIMER scores in the TCGA thyroid cancer cohort
Heatmap of select deconvolution results calculated using external cohort of well-differentiated thyroid cancers from TCGA. Histology, BRAF mutation, RAS mutation, BRS, MAP score category, and MAP score annotations are displayed on the top of the heatmap, followed by, TIMER and absolute value CIBERSORT (CIBERSORT-Abs) immune deconvolution scores, and CAF scores estimated by EPIC and MCPCOUNTER. Samples are sorted by increasing MAP score from left to right.

As transformed tumors such as PDTC and ATC have recently been tested in immunotherapy clinical trials, we focused our investigation of our TIMER findings in these specific subtypes. In a heatmap of TIMER results for just PDTC and ATC, we find three striking patterns of immune cell infiltrate: immune dessert, lymphocyte-rich, and CAF-rich (Figure 4-4A). We find that PDTCs generally show low scores for immune infiltration, regardless of MAP score. These results suggest "immune desert" microenvironments and agree with prior research in PDTCs.[102,103] ATCs, in contrast, are predicted to contain abundant stomal and immune cells, displaying either a lymphocyte/M1 macrophage-rich or CAF/M2 macrophage-rich infiltrate that appears to correlate with MAP score. CAF-rich microenvironments were more common in, but not limited to, ATCs sampled from a location in the thyroid and surrounding soft tissues. We also collected representative H&E staining images for these groups using samples from our cohort (Figure 4-4B). While our representative PDTC appears immune cold, our representative lymphocyte-rich ATC shows lymphocytes as dark purple dots, while our representative CAF-rich ATC shows light pink areas indicating abundant stoma with CAFs and macrophages scattered throughout. To test the association between MAP score and TIMER scores in ATCs, we categorized ATCs as either having a moderate MAP score or a high MAP score using a 50th percentile cutoff (Figure 4-4C) and found a significant association with lymphocyte-rich versus CAF-rich microenvironments, respectively (Figure 4-4D).
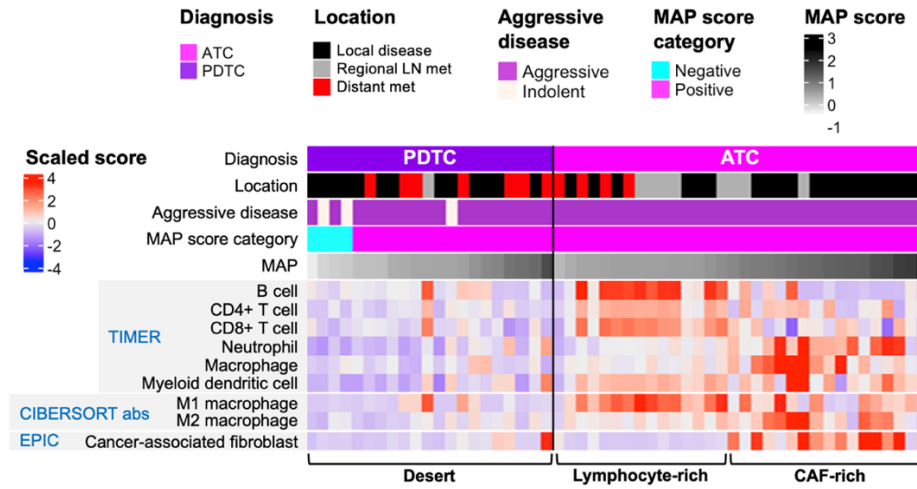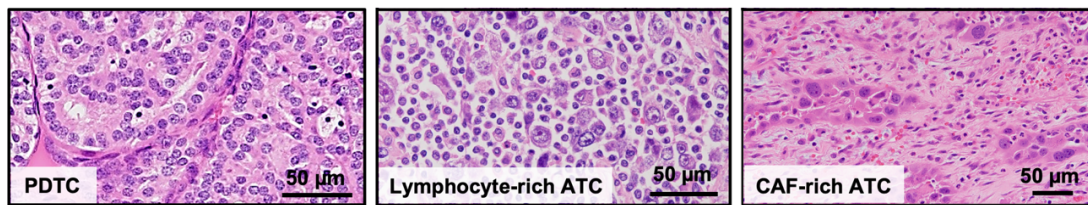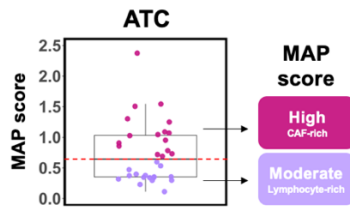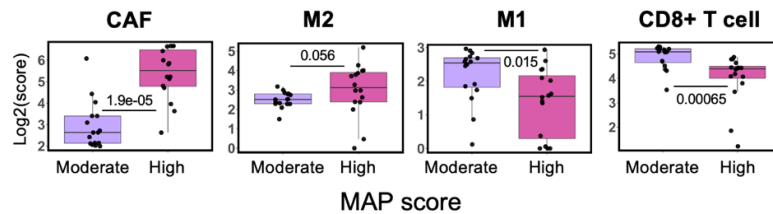
Figure 4-5: TIMER scores in transformed tumors

A) Heatmap of select deconvolution scores for transformed subtypes PDTC and ATC. Diagnosis, tissue location, aggressive disease, MAP score category, and MAP score annotations are shown at the top of the heatmap, followed by TIMER scores, M1/M2 absolute value CIBERSORT immune deconvolution scores, and EPIC CAF scores. Samples are sorted by increasing MAP score from left to right within each diagnosis.

B) Representative H&E histology of PDTC, lymphocyte-rich ATC, and CAF-rich ATC.

C) Diagram showing division of ATC tumors into moderate and high MAP score groups using a 50th percentile MAP score cutoff.

D) Box plots of EPIC CAF score, CIBERSORT absolute M2 macrophage score, CIBERSORT absolute M1 macrophage score, and TIMER CD8+ T cell score, comparing moderate and high MAP score ATCs. All scores are shown on a log2 scale, and p values were calculated using Wilcoxon rank-sum test.

4.3.2 MAP Score is Associated with Predicted Immunotherapy Response in ATCs

We used the prediction algorithm TIDE to estimate response to immune checkpoint blockade (ICB) immunotherapy and T cell inclusion in our ATCs. Based on our TIDE score results, our ATCs with moderate MAP score are predicted to respond to immunotherapy, while our ATCs with high MAP score are predicted to be less responsive (Figure 4-6A). Our TIDE exclusion results show an association between moderate MAP score ATCs with predicted T cell inclusion and high MAP score ATCs with predicted T cell exclusion (Figure 4-6B). To validate our T cell exclusion score results, we additionally performed immunohistochemical staining of CD3 in ATC whole tumor sections (Figure 4-6C), followed by blinded scoring of ATCs with either CD3 exclusion or CD3 inclusion, and found a strong correlation of CD3 score with our TIDE exclusion score (Figure 4-6D).
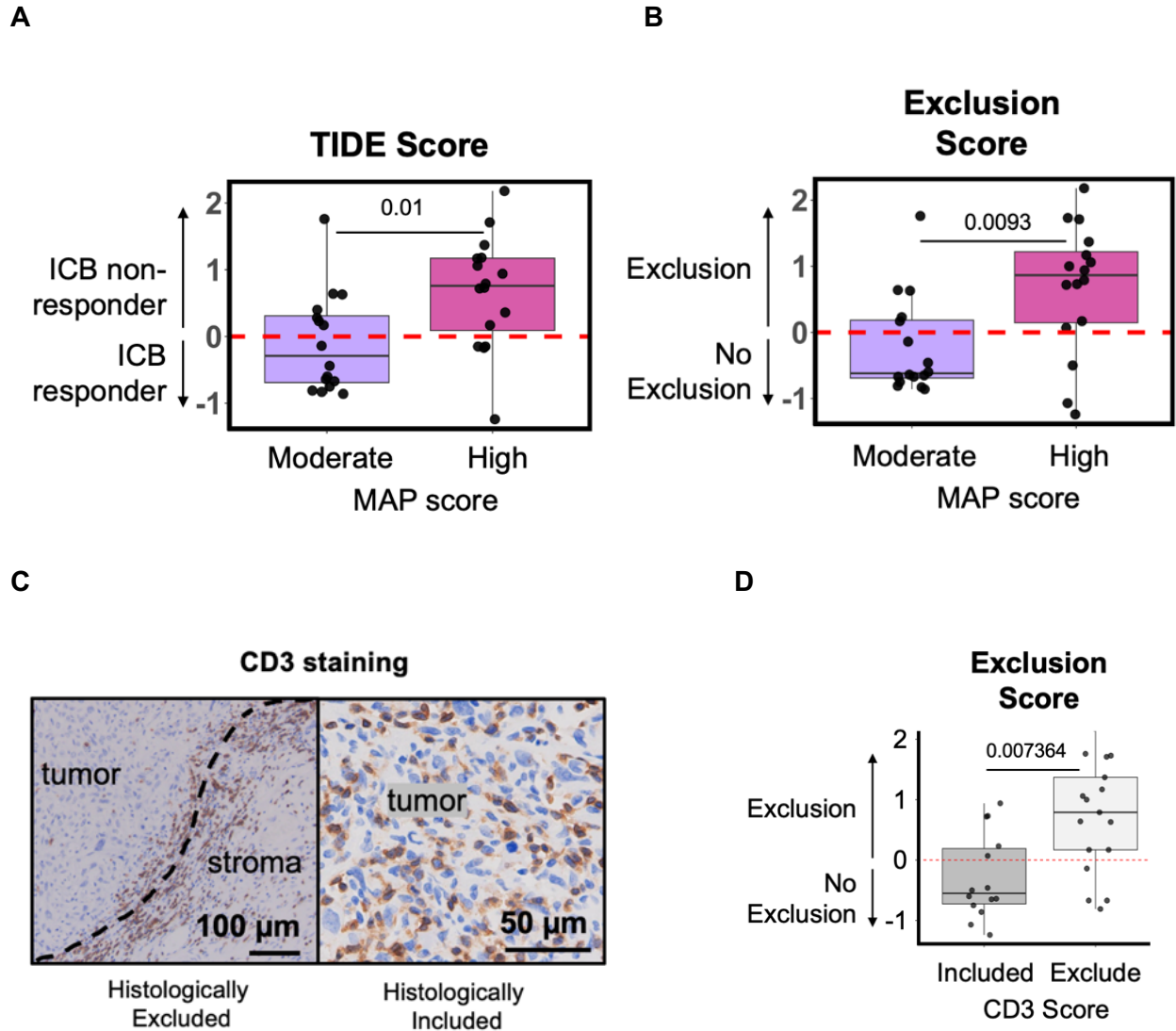
Figure 4-6: TIDE results and validation by CD3 staining

A) TIDE score in moderate and high MAP score ATCs. P value was calculated with Wilcoxon rank-sum test.

B) TIDE exclusion score in moderate and high MAP score ATCs. P value was calculated with Wilcoxon rank-sum test.

C) CD3 staining in representative ATC that is histologically excluded (left) and representative ATC that is histologically included (right).

D) TIDE exclusion score and CD3 staining score (included vs excluded) in ATCs. P value was calculated with Wilcoxon rank-sum test.

4.3.3 Additional Validation of Computational Deconvolution Results

In collaboration with Matthew Loberg, we used a number of additional approaches to confirm our bulk sequencing infiltrate findings in our ATCs. We performed spatial transcriptomics on eight ATC samples from our cohort, using the algorithm, SpaCET, to spatially deconvolute the immune cell populations present within individual spatial capture areas. We found robust CAF and macrophage infiltrate in all eight ATC specimens (Figure 4-7A). We further confirmed our findings through blinded pathologist review, with scoring of H&E and immunofluorescence staining for FAP (CAF marker) and MRC1 (M2 macrophage marker) in all ATCs of our cohort (Figure 4-7B). We found that ATCs with higher levels of CAFs, neutrophils, and M2 macrophages also had higher MAP scores.

To further investigate the association of CAFs and M2 macrophages in ATC, we also performed multiplex immunofluorescence for the CAF marker FAP and the M2 macrophage marker MRC1 in our ATC samples. We found a strong correlation between the abundance of MRC1-positive (MRC1+) macrophages and FAP-positive (FAP+) CAFs, and in our imaging, we see that FAP+ CAFs abutted tumor cells, while MRC1+ macrophages predominantly localized within the tumor stroma adjacent to fibroblasts (Figure 4-8A). We next quantitatively analyze co-localization of CAFs and M2 macrophages by looking at the correlation between predicted CAF and M2 macrophage frequency for individual capture areas in our spatial transcriptomics data (Figure 4-8B). We found that ATC samples had significant spatial correlation between CAFs and M2 macrophages. However, we also found that the magnitude of the correlation was more pronounced in tumors with higher MAP scores, suggesting greater CAF/M2 macrophage co-localization within high MAP tumors. We also used spatial transcriptomics and immunohistochemical staining to confirm our findings of infiltrating lymphocytes in moderate-MAP ATCs (Figure 4-8C). Spatial transcriptomics of ATC samples identified increased abundance of lymphoid populations in samples with moderate versus high MAP scores.
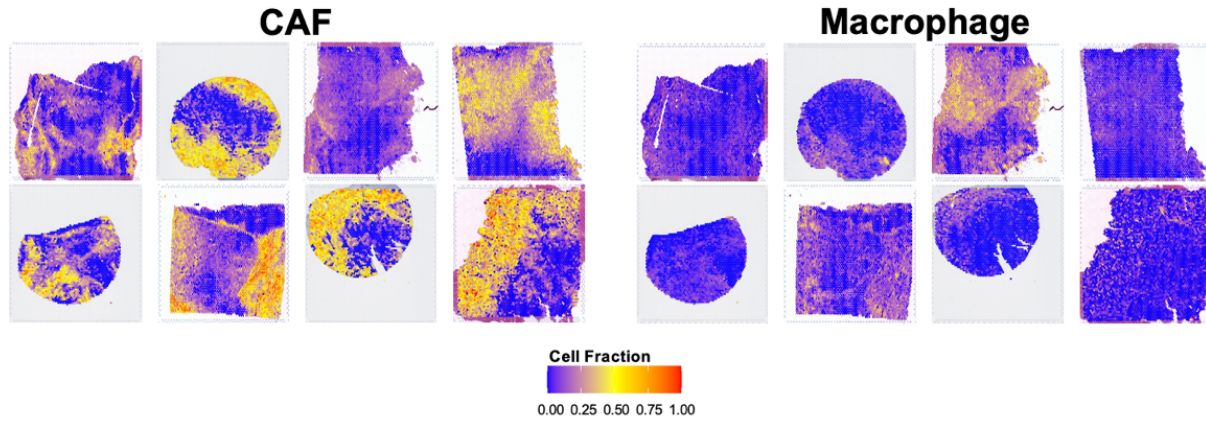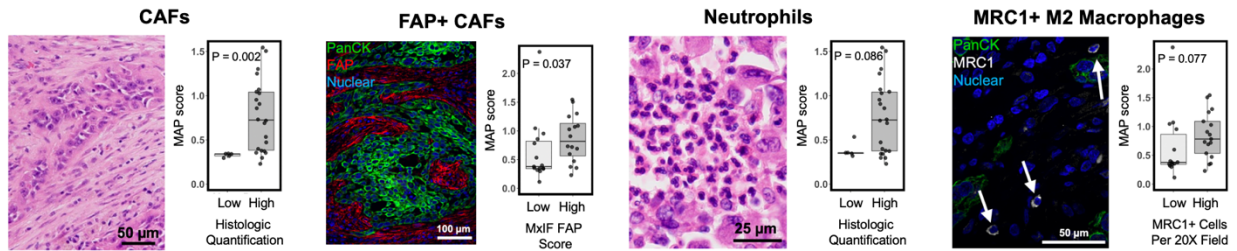
Figure 4-7: Validation of CAF and macrophage infiltrate in ATCs
A) SpaCET spatial deconvolution results for 8 ATC samples, showing estimated spatial capture area cell fractions for CAF and macrophage.
B) Box plots of MAP scores in ATCs, split into groups with either low or high histologic quantification of CAFs, FAP+ CAFs, neutrophils, and MRC1+ macrophages. Representative histology of each specific cell type is shown to the left of each box plot. P values were calculated using the Wilcoxon rank-sum test.

A

PanCK MRC1 FAP Nuclear 20 μm

Adjusted R^2: 0.26
P-value: 0.0014

Log2(MRC1⁺ Cells/20X Field)
FAP Staining Score

B

CAF/macrophage Co-localization

Adjusted R^2: 0.45
P-value: 0.040

50th percentile MAP score

M2/Fibroblast Co-Localization Rho
MAP score

Moderate MAP score

Spearman Rho = 0.088
P = 7.86e-09

Cell fraction (Macrophage M2)
Cell fraction (CAF)

High MAP Score

Spearman Rho = 0.51
P = 1.53e-248

Cell fraction (Macrophage M2)
Cell fraction (CAF)

Both CAF M2 Other

C

Spatial Transcriptomic Lymphoid Fraction

Lymphoid Fraction
1.00
0.50
0.00

Moderate MAP score

High MAP score

0.036

Average Lymphoid Deconvolution
Moderate High
MAP score

Figure 4-8: Validation of macrophage and CAF co-localization

A) Representative multiplex IF image of ATC with MRC1+ macrophages (indicated by white arrows) and adjacent FAP+ fibroblasts. Green = pan-cytokeratin, white = MRC1, red = FAP, blue = nuclear. Quantification of staining on the right shows the relationship between FAP staining and MRC1 staining. R^2 and p values are generated from a linear model with FAP staining score as the independent variable.

B) Linear model showing M2 macrophage and fibroblast co-localization from SpaCET deconvolution of eight ATCs as a dependent variable of MAP score (left), with representative spatial capture area M2 macrophage and fibroblast non-parametric rho correlation plots of moderate and high MAP score ATCs (right).

C) Representative images of lymphocyte deconvolution from spatial transcriptomics data of moderate and high MAP score tumors (left), and comparison of average lymphoid spatial capture fraction between moderate and high MAP score tumors for eight ATC samples (right). P value was calculated with Wilcoxon rank-sum test.

4.4 Discussion

Previous literature has suggested that CAFs, neutrophils, and macrophages may interact closely with tumor cells and support anti-tumor immunity.[104] Our finding of robust infiltration of CAFs, neutrophils, and M2 macrophages in high MAP score ATC suggests that these cells may play important roles in thyroid cancer progression. We speculate that these interactions may inform the development of novel immunotherapies targeting aggressive CAF-rich thyroid tumors. Furthermore, we found that our MAP score is associated with predicted immunotherapy response in ATCs. Recent clinical trials have showed some moderate response of ATC to ICB therapy.[93-98] Immune profiling using a molecular signature like our MAP score could someday help identify ICB-responsive ATC patients and direct non-ICB-responsive ATC patients to other more effective therapies.

CHAPTER 5

Title of Chapter: Papillary Thyroid Microcarcinoma with Distant Metastasis

This chapter includes adaptation of contents from the following manuscript:

Hu, R., et al. Incidental pulmonary metastases revealing subcentimeter papillary thyroid carcinoma. *AACE Clinical Case Reports*. 2020.

Contributions: I performed DNA sequencing analysis and methods writing.

5.1 Introduction

Papillary thyroid microcarcinomas (PMCs) are papillary thyroid cancers that are under 10 mm across their largest measurable diameter. PMCs are mostly indolent and have excellent prognosis; for example, autopsy data shows that 5% of the population has an undiagnosed PMC. In previous decades, PMCs were treated with a fairly aggressive approach where all tumors were recommended for removal. However, current practices are much more conservative. According to both the 2015 American Thyroid Association (ATA) guidelines and the 2017 Thyroid Imaging Reporting and Data System (TI-RADS), PMCs are not typically recommended for biopsy unless there is extrathyroidal extension or nodal metastasis.[3,105] However, in rare cases, some PMC patients may develop distant metastases. In this project, we perform molecular testing on PMC samples from two patients who also developed distant metastatic disease. In doing so, we highlight the need for improved markers for distinguishing low-risk and high-risk PMCs.

5.2 Methods

5.2.1 Patients

Patient 1 was a 70-year-old female with a history of partial thyroidectomy for benign disease who presented with incidental lung metastasis on computed tomography (CT) scan. Wedge biopsy of the left lung revealed PTC, and thyroid ultrasound revealed two subcentimeter nodules. Pathology following completion thyroidectomy revealed follicular variant PTC. Only the lung wedge resection was sequenced for this patient due to the small size of the intrathyroidal tumor. Patient 2 was a 29-year-old female who presented with incidental lung metastasis on CT scan; a core needle biopsy later confirmed metastatic cancer by thyroid transcription factor-1 and thyroglobulin expression. Neck ultrasound revealed two subcentimeter nodules, and total thyroidectomy revealed a follicular variant PTC.

5.2.2 DNA Isolation and Whole Exome Sequencing

DNA was isolated from formalin-fixed paraffin embedded (FFPE) tissue using the Covaris DNA FFPE isolation kit. Library building was performed using the NEB DNA Ultra II kit, and sequencing was performed on a NovaSeq 6000 system using the IDT research exome panel. FastQC was used to check raw sequencing data quality,[53] and DNA reads were then aligned to the hg19 reference human genome using Burrows-Wheeler alignment before sorting and indexing by SAMtools.[106] GATK tools including Picard and HaplotypeCaller were used for variant calling.[55,107] Variants were filtered by strand, base quality, mapping quality, and end distance biases or adjacent gaps before annotation by ANNOtate VARiation.[108] Variants were then filtered on 119 known cancer genes (Table 5-1). We additional filtered for variants with an exome allele frequency <0.001 to exclude variants common in the population.[57] Variants were categorized according to published standards.[59]

AGGF1	CREB3L2	GOLGA5	LRP6	PTEN	TERT
AGK	CTNNB1	GORASP2	LTK	PTH	TFG
AKAP13	DICER1	GSK3B	MACF1	RAF1	THADA
AKAP9	EIF1AX	GTF2IRD1	MEN1	RBPMS	TP53
AKT1	EML4	HOOK3	MET	RET	TPM3
ALK	EP300	HRAS	MKRN1	RNF213	TPR
APC	ERBB4	IDH1	NCOA4	RNF43	TRA2A
AXIN1	ERC1	IDH2	NF2	ROS1	TRIM24
BANP	ETV6	IGF2BP3	NRAS	SLC26A11	TRIM27
BCL2L11	EZH1	IRF2BP2	NTRK1	SLC5A5	TRIM33
BRAF	EZR	KIAA1217	NTRK3	SND1	TRIM61
C7orf10	FAM114A2	KIAA1598	OFD1	SPECC1L	TSC2
CALCA	FAM193A	KIF5B	PAX8	SQSTM1	TSHR
CCDC149	FARSB	KLK1	PCM1	SS18	UACA
CCDC30	FBXW7	KRAS	PGK1	SSBP2	VCL
CCDC6	FGFR2	KRT20	PICALM	STK11	VHL
CCNY	FKBP15	KRT7	PIK3CA	STRN	WARS
CHEK2	GFPT1	KTN1	POR	SYN2	ZBTB8A
CHGA	GLIS3	LGR5	PPARG	TBL1XR1	ZC3HAV1
CITED1	GNAS	LOC389473	PRKAR1A	TCF7L2	--

Table 5-1: List of 119 known cancer genes used for mutation filtering

5.3 Results

Sequencing of patient 1's lung metastasis revealed a missense mutation *TSHR*, c.1721C>G, which codes for the protein alteration p.Thr574Ser. This mutation is a variant of unknown clinical significance. Additional mutations in genes including *GNAS, ALK, NCOA4*, and *CHGA* were identified, and are not predicted to be damaging, potentially representing passenger or germline alterations.

Sequencing of patient 2's primary intrathyroid tumor revealed a NRAS mutation, c.182A>G, coding for the protein alteration p.Gln61Arg. This alteration is a well-described pathogenic mutation. Sequencing also identified a novel GNAS missense mutation, c.46C>T (p.Arg16Cys), representing a potential passenger mutation, as well as alterations in *MACF1* and *RNF213,* which may represent germline alterations.

In addition, we applied our MAP score (described in Chapter 3) to the sequencing data from both patients. We found that patient 1 had a MAP score of 0.38, while patient 2 had a MAP score of -0.328958. While only patient 1 had a positive MAP score which predicts aggressive disease, this patient's sample also lacked any detectable pathogenic mutations during our study. This suggests the potential utility of our score for predicting aggressive disease in patients without aggression-associated mutations.

A complete table of mutation findings in shown in Table 5-2. Mutations in *TERT* promoter, *TP53*, and *PIK3CA*, typically associated with aggressive thyroid cancer, were not identified in samples from either patient.

Patient	Gene	Chromosome	Alteration type	Coding DNA alteration	Protein alteration	Allele frequency	Reference gene
1	ALK	2	Missense	c.932G>A	p.Arg311His	55%	NM_004304.4
	NCOA4	10	Missense	c.1457A>G	p.Gln486Arg	52%	NM_001145260.1
	TSHR	14	Missense	c.1721C>G	p.Thr574Ser	45%	NM_000369
	CHGA	14	In-frame deletion	c.1095_1097 delGGA	p.Glu365del	11%	NM_001275.3
	GNAS	20	In-frame insertion	c.1188_1196 dupTGCAGCCCC	p.Ala398 _Ala400dup	61%	NM_001077490.1, NM_08042
2	MACF1	1	Missense	c.12404C>T	p.Ala4135Val	52%	NM_012090
	NRAS	1	Missense	c.182A>G	p.Gln61Arg	13%	NM_002524.4
	RNF213	17	Frameshift deletion	c.11658delG	p.Lys3887 ArgfsTer49	47%	NM_001256071.1
	GNAS	20	Missense	c.46C>T	p.Arg16Cys	30%	NM_001077490.1, NM_080425.2

Table 5-2: Mutations identified in patient samples

5.4 Discussion

In this project, we analyzed two patients with PMC who also developed metastatic disease in the lungs. These patients would not have met current biopsy criteria, which would have potentially led these patients to have missed or delayed diagnosis of distant metastasis. In addition, even if the patients had undergone biopsy and sequencing, the mutations detected in the patient samples did not include any previously described mutations associated with aggressive thyroid cancer, such as *BRAF*, *TP53*, or *TERT* promoter mutations. These two patients highlight a need for improved methods to distinguish truly low-risk PMCs from PMCs at risk for metastasis. Our findings utilizing our MAP score for aggression disease prediction further revealed a potential utility of our score for predicting aggressive disease in patients without aggression-associated mutations; however, our ability to assess risk prediction power in PMCs is limited by the number of samples in this study. Additional research to expand our understanding of the molecular drivers and tumor microenvironment of thyroid cancer will be critical for improving our ability to anticipate aggressive tumor behavior.

CHAPTER 6

Title of Chapter: Thyroid Cancer and the Canonical Wnt Pathway

6.1 Introduction

Increased Wnt signaling is classically associated with aggressive behavior across cancers, including thyroid cancer. For example, anaplastic thyroid cancer (ATC), the most aggressive form of thyroid cancer, has consistently been shown to have Wnt/β-catenin (canonical) signaling activation in previous studies.[29-31] Treatment options are extremely limited for patients with ATC; there are currently no targeted therapies available for *BRAF*-wildtype ATCs, and combination dabrafenib-trametinib therapy (approved for ~25% of *BRAF* V600E mutant ATCs) has shown limited efficacy to date.[109] The Wnt signaling pathway may offer new potential targets for therapy for such aggressive thyroid cancers. In our study, we use our large cohort of patient thyroid lesions to study upregulation and potential drivers of Wnt signaling in ATCs. We analyze DNA and RNA sequencing data and show that Wnt ligands rather than Wnt pathway mutations appear to drive most Wnt signaling upregulation in ATCs. In addition, we find an association between Wnt signaling activation and cancer-associated fibroblasts (CAFs), as well as reduced progression-free survival (PFS) time with increased Wnt signaling. We additionally investigate the expression of factors that might function downstream by Wnt signaling to influence cancer cell adhesion, differentiation, and aggression.

6.2 Methods

6.2.1 Patient Cohort Sequencing and Data Analysis

Patient cohort, sequencing methods, mutation analysis, gene expression analysis, and survival analysis methods used for this project were the same as those described in section 2.2.

6.2.2 Transcription Factor Binding Prediction

To perform transcription factor binding prediction, we used JASPAR (https://jaspar.genereg.net/), an open-access database containing manually curated, non-redundant transcription factor (TF) binding profiles.[110] To perform our analysis in JASPAR, we first located and loaded the transcription factor matrix profile of interest, TCF7. Next, for each gene we tested, the 1500 base pairs upstream of the gene's promoter were copied from the UCSC Genome Browser (Human Assembly: GRCh38/hg38) (https://genome.ucsc.edu/index. html)[111] and input into the JASPAR scan function. After running the scan, the sequence with the top JASPAR relative score was selected for each gene. For controls, known Wnt target genes *CCND1, MYC, LGR5, AXIN2, LEF1, JUN, VEGFA*, and non-Wnt target genes *TG, LIME1, ARIH2, COPS7B, and GALNT18* were also tested.

6.2.3 qPCR Timecourse

ATC cell line THJ-21T was passaged and counted before plating ~374,000 cells in 2 mL of RPMI + 10% FBS media per well in a 6-well plate, before placing in an incubator to allow cells to attach overnight. The next day, three aliquots of RPMI + 10% FBS media were warmed to 37°C and used to prepare the following treatments: 1 μM Chiron, 50 ng/mL Wnt3a, and no treatment control. Old media was removed from the cells by aspiration and treatments were added slowly down the side of each well. At timepoints 4, 8, and 12 hours after starting the treatment, RNA was isolated using the Qiagen RNeasy Mini Kit (Product #: 74106) with the Qiagen RNase-free DNase Set (Product #: 79254), and the ThermoFisher Qubit RNA High Sensitivity (HS) Kit (Product #: Q32852) was used to quantify RNA concentration before storage at -80°C. After RNA had been isolated from each timepoint, cDNA was made using the BioRad iScript cDNA Synthesis Kit (catalog #: 1708891) according to manufacturer's instructions. qPCR was performed in a clear-bottomed 96-well plate with Itaq™ Universal SYBR (catalog #: 1725125), diluted cDNA, and the

appropriate forward and revere primers. Primers used included those for control housekeeping genes PGK1, with forward primer 5'-GAACAAGGTTAAAGCCGAGCC-3' and reverse primer 5'-GTGGCAGATTGACTCCTACCA-3, as well as GAPDH, with forward primer 5'-AGCCACATCGCTCAGACAC-3' and reverse primer 5'-GCCCAATACGACCAAATCC-3'. To measure AXIN2 as a proxy for Wnt signaling, we used forward primer 5'-TGGCCTCGTCTTTGAATCCC-3' and reverse primer 5'-TGAACGGCTTTCATGTCGGA-3'. For VIM, we used forward primer 5'-GAGAACTTTGCCGTTGAAGC-3' and reverse primer 5'-GCTTCCTGTAGGTGGCAATC-3'. For CD44, we used forward primer 5'-TCTGCAAGGCCTTTAATAGCACG and reverse primer 5'-GTTCGCAGCACAGATGGAATTG-3'.

6.3 Results

6.3.1 Wnt Pathway Upregulation is Ligand Driven in Aggressive Thyroid Cancer

Wnt signaling is a well-known driver of aggressive thyroid carcinoma. To begin identifying Wnt pathway mutations and gene expression changes in thyroid cancer, we performed whole-exome and bulk RNA sequencing on 312 formalin-fixed paraffin-embedded (FFPE) resection samples from 251 patients with a diverse range of thyroid nodules from two tertiary care centers. This cohort included 34 ATC tumors. Despite the known role of Wnt signaling, DNA sequencing analyses identified very few Wnt pathway associated mutations in our thyroid cohort, with just 6% of ATCs having mutations in the Wnt pathway genes tested (Figure 6-1A,1B). Despite the lack of mutations, IHC staining of β-catenin in ATC patient samples confirmed upregulated Wnt signaling. Intriguingly, nuclear β-catenin staining was identified in both the tumor cells and stroma, suggesting that Wnt signaling may be ligand-driven rather than mutation-driven (Figure 6-1C). To further investigate the role of Wnt signaling in ATC and other thyroid cancers, we performed differential gene expression analyses using our cohort. We found that gene expression of 11 Wnt

ligands was significantly higher in ATCs compared to either benign thyroid lesions or well-differentiated thyroid cancer (WDTC) subtypes (Figure 6-2). We conclude from our findings that increased Wnt signaling in ATCs is commonly ligand-driven rather than mutation-driven.

A
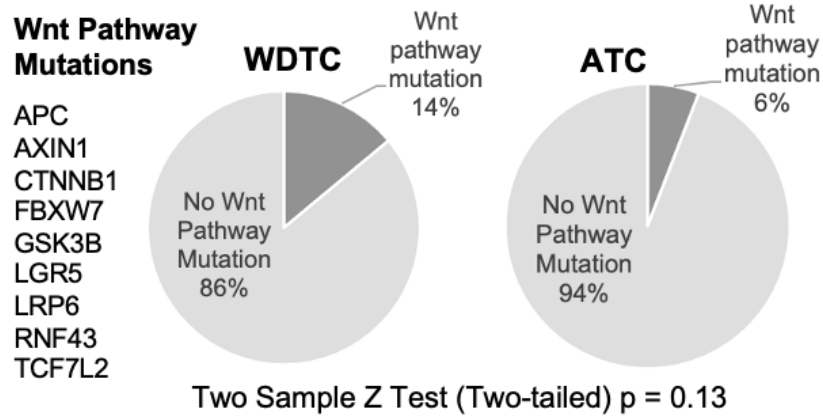
B

Diagnosis	APC	AXIN1	CTNNB1	FBXW7	GSK3B	LGR5	LRP6	RNF43	TCF7L2
NIFTP	6.7% (1/15)	0% (0/15)	0% (0/15)	0% (0/15)	0% (0/15)	0% (0/15)	6.7% (1/15)	0% (0/15)	0% (0/15)
HC	4.5% (1/22)	0% (0/22)	0% (0/22)	0% (0/22)	0% (0/22)	0% (0/22)	0% (0/22)	4.5% (1/22)	0% (0/22)
FTC	7.1% (2/28)	0% (0/28)	0% (0/28)	0% (0/28)	0% (0/28)	0% (0/28)	7.1% (2/28)	0% (0/28)	0% (0/28)
EFVPTC	15.4% (2/13)	0% (0/13)	0% (0/13)	0% (0/13)	0% (0/13)	0% (0/13)	0% (0/13)	0% (0/13)	0% (0/13)
IFVPTC	0% (0/18)	0% (0/18)	0% (0/18)	5.6% (1/18)	0% (0/18)	0% (0/18)	5.6% (1/18)	0% (0/18)	0% (0/18)
PTC	5.7% (5/88)	2.3% (2/88)	2.3% (2/88)	1.1% (1/88)	1.1% (1/88)	3.4% (3/88)	5.7% (5/88)	1.1% (1/88)	2.3% (2/88)
PDTC	0% (0/21)	0% (0/21)	9.5% (2/21)	0% (0/21)	0% (0/21)	4.8% (1/21)	0% (0/21)	0% (0/21)	0% (0/21)
ATC	2.9% (1/34)	0% (0/34)	0% (0/34)	0% (0/34)	0% (0/34)	0% (0/34)	2.9% (1/34)	0% (0/34)	0% (0/34)

C
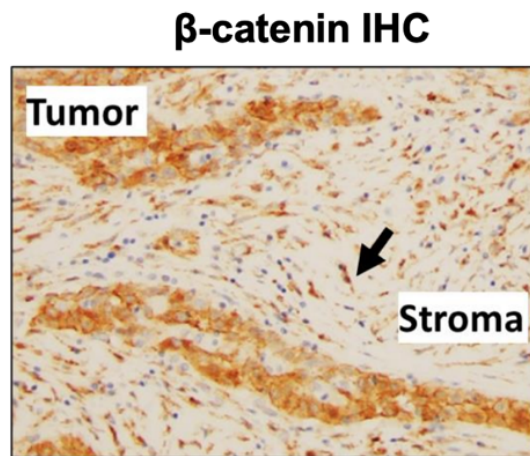
Figure 6-1: Wnt pathway upregulation in ATCs is not commonly mutation driven
A) Wnt pathway mutations in WDTC and ATC samples.
B) Table of Wnt pathway mutations organized by diagnosis
C) β-catenin IHC of representative ATC patient sample.
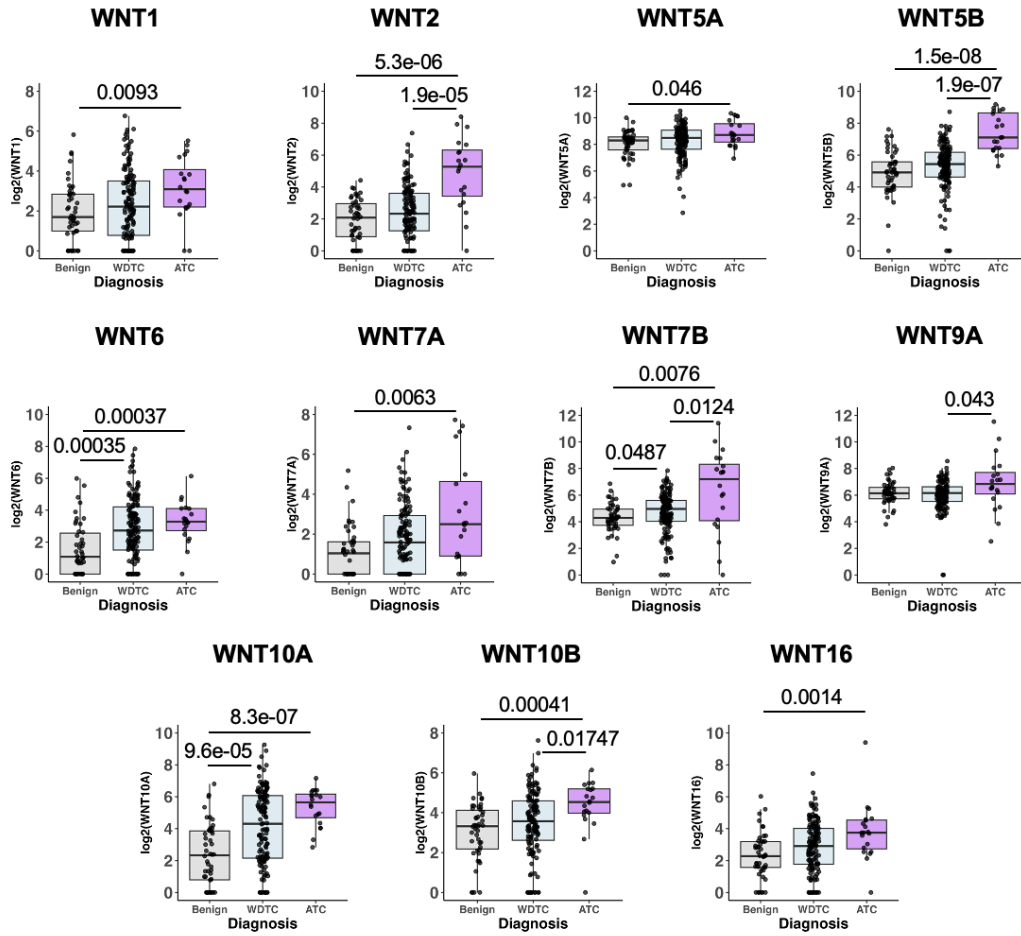
Figure 6-2: Wnt ligand drives Wnt pathway upregulation in ATCs
Wnt ligand expression in benign, WDTC, and ATC samples.

6.3.2 Wnt Pathway Signaling is Associated with High CAF and Macrophage Score

We next explored the tumor microenvironment of Wnt-driven thyroid cancers in our patient cohort. Our previous work identified unique tumor microenvironments in MAP high and MAP moderate ATCs, CAF-rich and lymphocyte-rich, respectively. We now utilized gene ontology (GO) analysis to look at enrichment of Wnt pathway gene sets between ATCs with either high or low CAF score (Figure 6-3A). Our results showed that CAF-high ATCs were enriched in six different Wnt pathway gene sets, suggesting that ATCs with high CAF scores also have increased Wnt pathway signaling. Our previous work also found that CAF-rich ATCs were associated with higher macrophage scores. Therefore, we also used GO analysis to look at enrichment of Wnt pathway gene sets between ATCs with either high or low macrophage score (Figure 6-3B). Our results showed that ATCs with high macrophage scores were also enriched in these Wnt pathway gene sets.
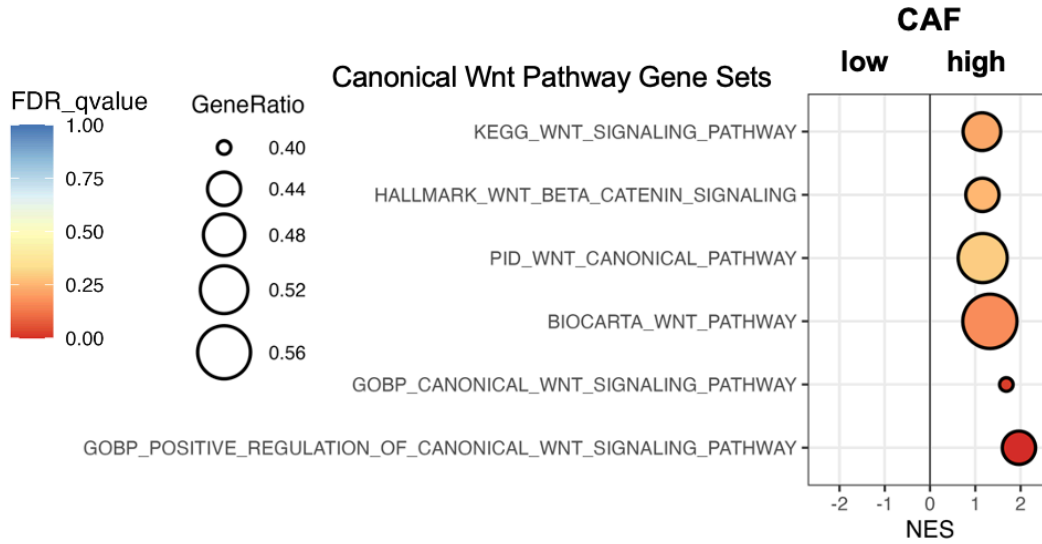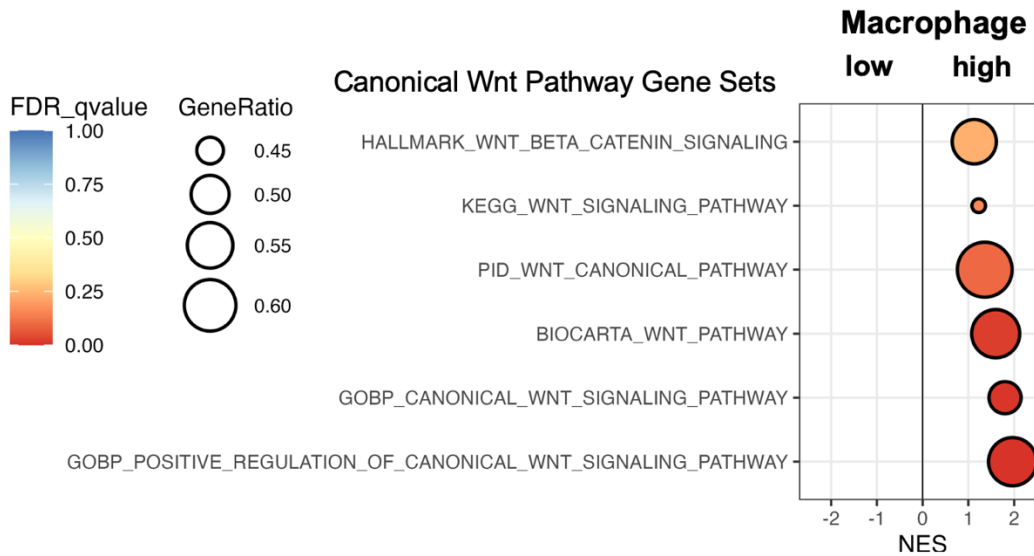
Figure 6-3: Wnt pathway upregulation is associated with increased CAF and macrophage score in ATCs
A) Gene ontology (GO) analysis of ATCs with high or low CAF (EPIC) score.
B) Gene ontology (GO) analysis of ATCs with high or low macrophage (TIMER) score.

6.3.3 Wnt Pathway Signaling is Associated with Worse Progression-free Survival

Next, we tested the association between increased Wnt signaling and thyroid cancer patient survival. Using the "Hallmark Wnt Beta Catenin Signaling" score, we divided patients into two groups with either positive or negative scores based on the patient's local malignant lesion (Figure 6-4A). We found that patients with positive Wnt scores had significantly shorter progression-free survival (PFS) compared to patients with negative Wnt scores, suggesting that Wnt signaling is associated with more aggressive thyroid cancer. We also compared PFS between ATCs with high vs low expression of the gene *WNT2*, using a 50th percentile cutoff. *WNT2* is a Wnt ligand known to be involved in canonical Wnt signaling and also reported to be produced by CAFs to promote increased aggression in colorectal cancer.[112] We found that high expression of *WNT2* was significantly associated with shorter PFS (Figure 6-4B). To further study the clinical relevance of *WNT2*, we looked at thyroid differentiation in our ATCs using a thyroid differentiation score (TDS), based on the expression of 16 thyroid function genes (Figure 6-4C). We found that ATC samples with lower TDS, and therefore less differentiation, had higher *WNT2* gene expression. In addition, we looked at local thyroid samples from patients with either associated metastasis or no associated metastases (Figure 6-4D). We found that patients with associated metastasis had significantly higher *WNT2* expression. In a gene expression heatmap containing only our ATC samples (Figure 6-5), we see that ATCs with high expression of *WNT2* appear to be correlated with high cancer-associated fibroblast (CAF) score. These findings support the idea that CAFs could be producing Wnt2 ligand in tumors with upregulated canonical Wnt signaling.
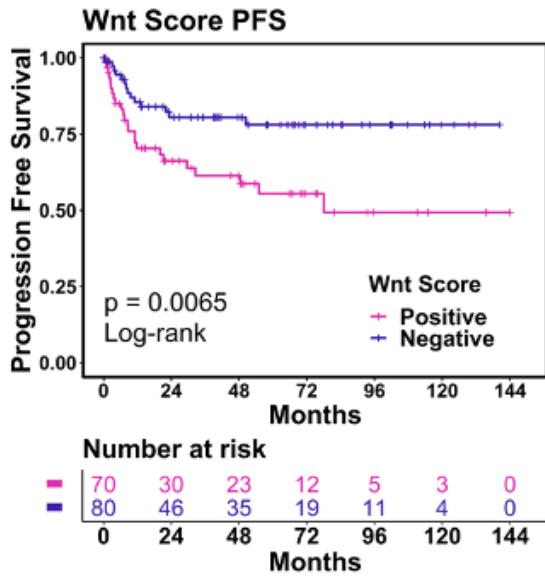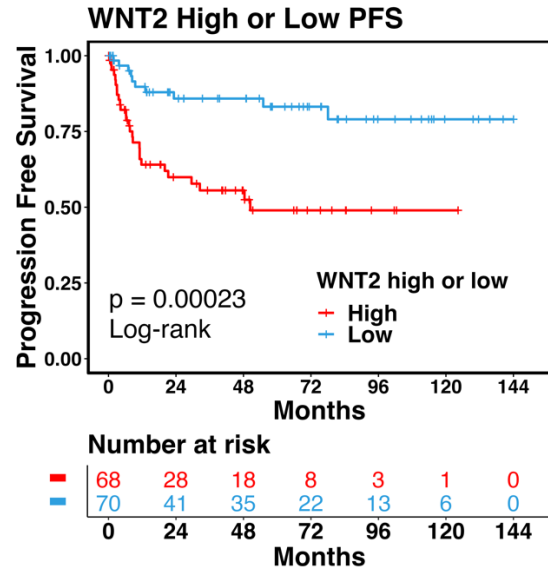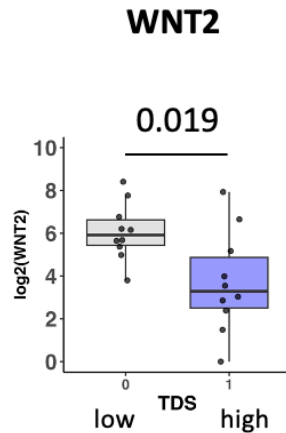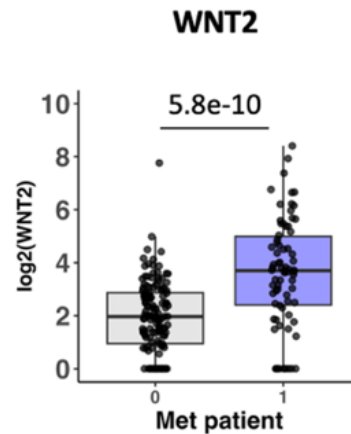
Figure 6-4: Wnt pathway upregulation and *WNT2* expression are associated with decreased progression-free survival (PFS), lower TDS, and metastasis in ATC patients
A) PFS curve for thyroid cancer patients with positive vs negative Wnt scores, using "Hallmark Wnt Beta Catenin Signaling" score calculated from each patient's local malignant lesion.
B) PFS curve for thyroid cancer patients with high vs low *WNT2* gene expression, using "Hallmark Wnt Beta Catenin Signaling" score calculated from each patient's local malignant lesion.
C) Gene expression for *WNT2* in ATCs with either low or high TDS score. P values were calculated with Wilcoxon rank-sum test.
D) Gene expression for *WNT2* for all local thyroid tumors with either no associated metastatic disease (0) or associated metastatic disease (1). P values were calculated with Wilcoxon rank-sum test.
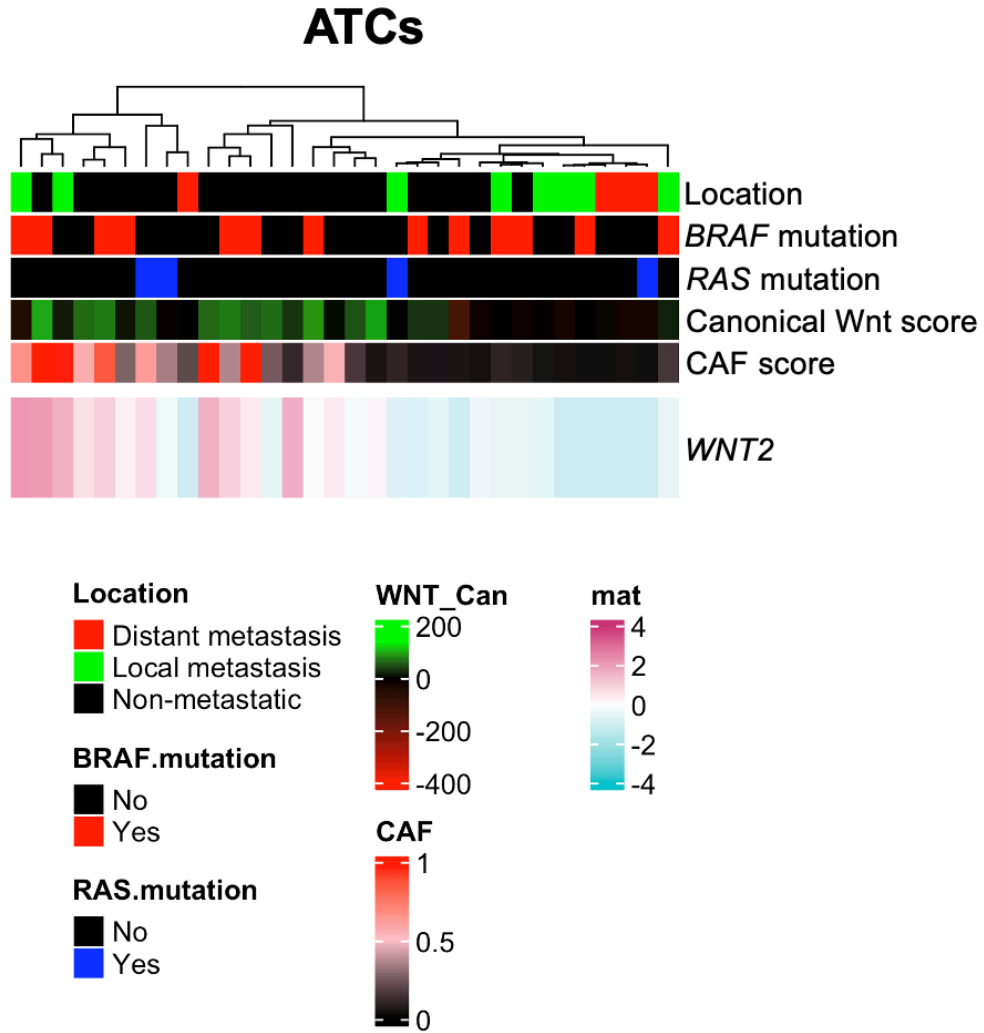
Figure 6-5: *WNT2* **gene expression heatmap**
Normalized gene expression heatmap of ATC samples showing *WNT2* expression. Annotations include sample location, BRAF mutation status, RAS mutation status, canonical Wnt signaling score, and cancer-associated fibroblast (CAF) score (EPIC).

6.3.4 Wnt Pathway Signaling is Associated with Cadherin and Laminin Gene Expression

While we have showed that Wnt signaling may be associated with survival of patients with thyroid cancer, the exact mechanistic role that Wnt plays in thyroid cancer behavior is unknown. In previous literature, researchers have suggested that canonical Wnt signaling may interact with laminin and cadherin to control cancer cell adhesion and migration.[113] Laminin is the main non-collagenous glycoprotein found in the basement membrane and has been shown to play important roles in tumor invasion and metastasis.[114] Cadherins are well-known epithelial cell-cell adhesion molecules that can stabilize cellular β-catenin levels and potentiate β-catenin signaling.[115] Additionally, expression of *CDH3*, which codes for P-cadherin, has been associated with metastasis and poor prognosis.[116,117] Therefore, we investigated cadherin and laminin gene expression in our thyroid cancer cohort. We use the online tool JASPAR to make predictions of TCF7 transcription factor binding sites. Using Jaspar, we generate relative scores for the sequences up to 1500 base pairs (1.5kb) upstream of the *CDH3, LAMA3*, *LAMB3*, and *LAMC2* promoter and compare them to those of other known Wnt target genes (Figure 6-6A). We find that JASPAR scores for the sequences up to 1.5kb upstream of these genes are comparable to those of some known Wnt target genes, in particular *LAMA3*, which suggests the possibility of these genes being Wnt pathway targets. In addition, in a gene expression heatmap containing local thyroid lesion samples (Figure 6-6B), we find that higher expression of *CDH3, LAMA3*, *LAMB3*, and *LAMC2* may trend with increased Wnt signaling score, particularly in our PTCs and ATCs. In a volcano plot comparing lesions with negative and positive Wnt signaling score (Figure 6-6C), we see that samples with higher Wnt signaling score appear to be enriched in *CDH3, LAMA3*, *LAMB3*, and *LAMC2* gene expression. Overall, these findings show that these cadherin and laminin genes are potential targets of the Wnt pathway and appear to be upregulated in thyroid tumors with increased Wnt signaling,

A
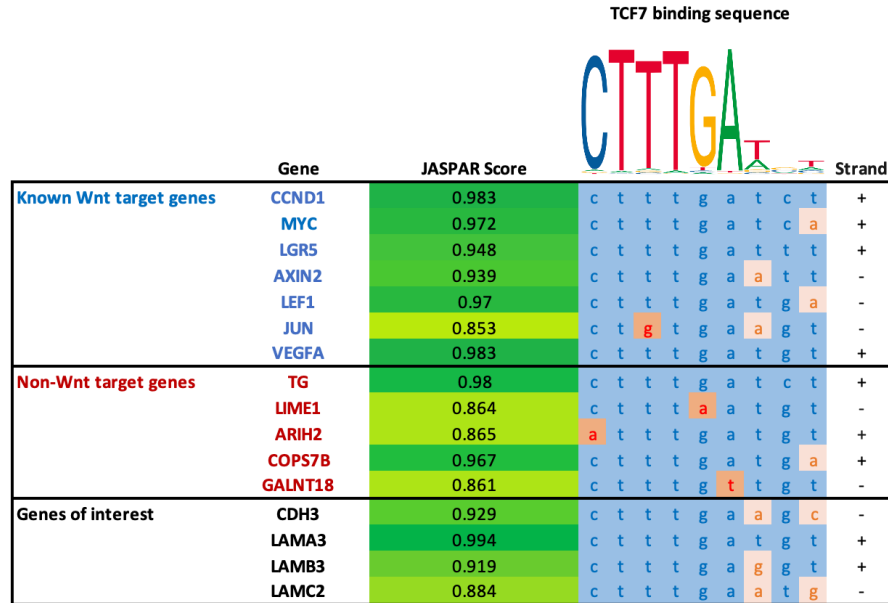
B
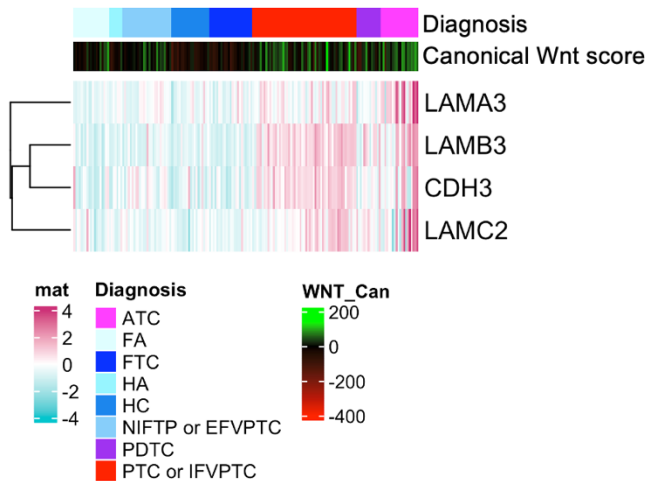
C
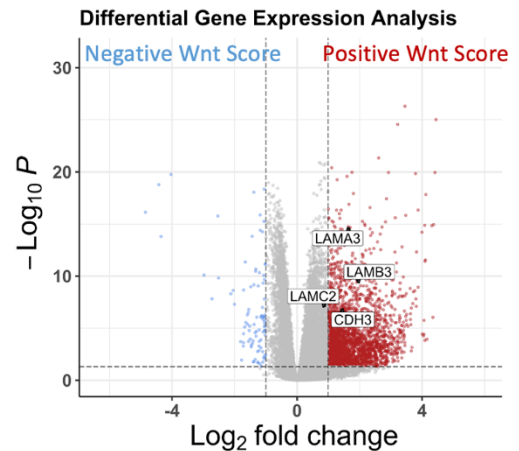
Figure 6-6: Wnt signaling may target expression of *CDH3*, *LAMA3*, *LAMB3*, and *LAMC2* in ATCs

A) JASPAR results showing top match by JASPAR score for the TCF7 binding sequence in DNA within 1.5kb upstream of the promoter of *CDH3*, *LAMA3*, *LAMB3*, and *LAMC2*, compared to known Wnt target and non-Wnt target genes.

B) Normalized gene expression heatmap of local thyroid lesion samples showing *CDH3*, *LAMA3*, *LAMB3*, and *LAMC2* expression. Annotations include diagnosis and canonical Wnt signaling score.

C) Volcano plot showing differential gene expression of *CDH3*, *LAMA3*, *LAMB3*, and *LAMC2* expression between thyroid lesion samples with either negative or positive Wnt score.

6.3.5 Wnt Pathway Signaling is Associated with VIM and CD44 Gene Expression

Previous research has also shown the genes *VIM* and *CD44* may be upregulated in cancers with increased Wnt signaling.[118-120] *VIM* codes for vimentin, a widely expressed intermediate filament (IF) protein that is important for maintaining cell integrity in the cytoskeleton, and is also associated with EMT and cancer progression.[119,121] *CD44* is a cell-surface glycoprotein involved in cell–cell interactions, cell adhesion, and migration, and is also a known cancer stem cell marker for many solid tumors including colorectal cancer.[118] We proposed that Wnt signaling upregulates expression of the genes *VIM* and *CD44* in thyroid cancer to cancer cell de-differentiation and invasion. Again, we use the online tool JASPAR to make predictions of TCF7 transcription factor binding sites. We used JASPAR to generate relative scores for the sequences (up to 1000bp) upstream of the *VIM* and *CD44* promoter and compare them to those of other known Wnt target genes (Figure 6-7A). We find that JASPAR scores for the 1.5kb region upstream of *VIM* and *CD44* are comparable to those of some known Wnt target genes. In a gene expression heatmap containing only our ATC samples (Figure 7B), we see that ATCs with high *VIM* and *CD44* gene appears correlated with high canonical Wnt signaling score, further supporting our prediction that these genes are Wnt pathway targets.

To further support our VIM and CD44 findings, we performed additional experiments measuring gene expression of VIM and CD44 in ATC cell lines with Wnt pathway activation. To test gene expression, we performed qPCR timecourse experiments in ATC cell line THJ-21T treated with Wnt3a or CHIR99021 (Chiron) to activate Wnt signaling. We confirmed Wnt pathway activation in Wnt3a or Chiron treated THJ-21T cells using qPCR of AXIN1 gene expression. We then measured expression of *VIM* and *CD44* by qPCR at timepoints 4, 8, and 12 hours after Wnt3a or Chiron treatment (Figure 6-8A,8B). We found that Wnt3a and Chiron treatment showed increased expression of *VIM* and *CD44*, peaking at 8hrs with Chiron treatment, and peaking at 12 hours with Wnt3a treatment. These findings further support our hypothesis that *VIM* and *CD44* are downstream targets upregulated by canonical Wnt signaling.

A

TCF7 binding sequence

	Gene	JASPAR Score	Sequence	Strand
Known Wnt target genes	CCND1	0.983	c t t t g a t c t	+
	MYC	0.972	c t t t g a t c a	+
	LGR5	0.948	c t t t g a t t t	+
	AXIN2	0.939	c t t t g a a t t	-
	LEF1	0.97	c t t t g a t g a	-
	JUN	0.853	c t g t g a a g t	-
	VEGFA	0.983	c t t t g a t g t	+
Non-Wnt target genes	TG	0.98	c t t t g a t c t	+
	LIME1	0.864	c t t t a a t g t	-
	ARIH2	0.865	a t t t g a t g t	+
	COPS7B	0.967	c t t t g a t g a	+
	GALNT18	0.861	c t t t g t t g t	-
Genes of interest	CD44	0.894	c t t t g a a t g	-
	VIM	0.895	c t t t t a t c t	-

B

ATCs
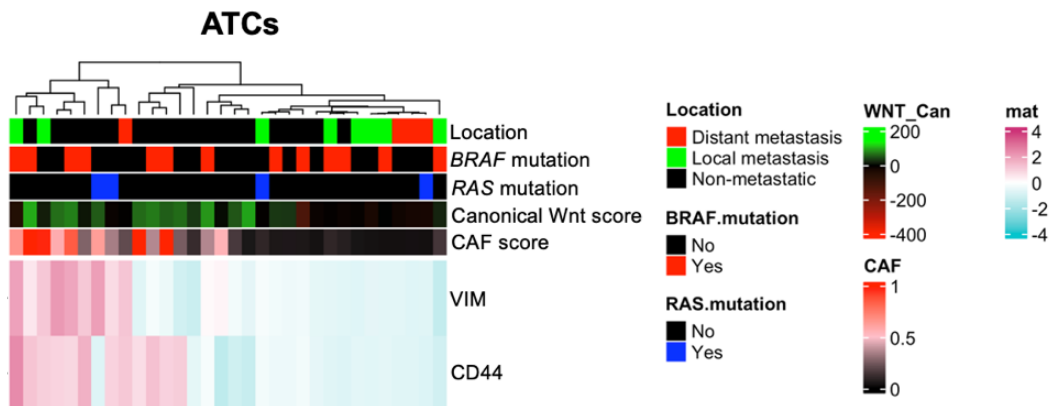
Figure 6-7: CD44 and VIM are potential Wnt signaling targets

A) JASPAR results showing top match by JASPAR score for the TCF7 binding sequence in DNA within 1.5kb upstream of the promoter of *CD44* and *VIM*, compared to known Wnt target and non-Wnt target genes.

B) Normalized gene expression heatmap of ATC samples showing *CD44* and *VIM* expression. Annotations include sample location, BRAF mutation status, RAS mutation status, canonical Wnt signaling score, and cancer-associated fibroblast (CAF) score (EPIC).

A
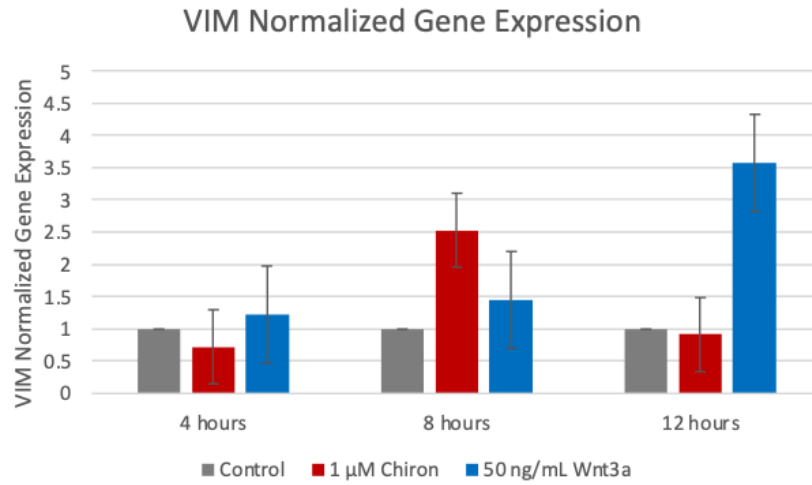
B
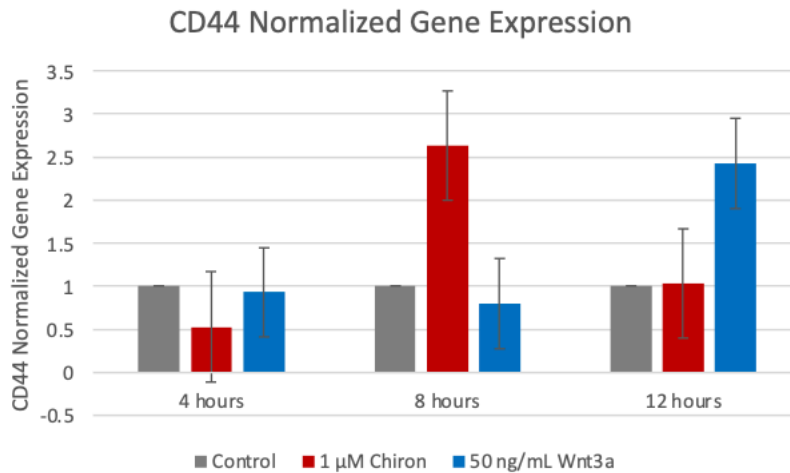
Figure 6-8: Upregulation of CD44 and VIM with Wnt activation in vitro
qPCR timecourse results showing normalized gene expression of (A) *VIM* and (B) *CD44*, with Wnt3a treatment, Chiron treatment, or no treatment control. Three biological replicates were performed. Error bars indicate standard error of the mean (SEM).

6.3.6 VIM and CD44 Gene Expression are Associated with Aggressive Disease

As previous research has suggested that *VIM* and *CD44* have been associated with EMT, stemness, migration, and cancer progression,[118,119,121] we next investigated whether increased *VIM* and *CD44* are associated with increased dedifferentiation and aggression in ATCs. We compared thyroid differentiation in our ATCs as estimated by our previously described thyroid differentiation score (TDS) (Figure 6-9A), and found that ATC samples with lower TDS, and therefore less differentiation, had higher *VIM* gene expression. To further study to clinical relevance of *VIM* and *CD44*, we looked at local thyroid samples from patients with either associated metastasis or no associated metastases (Figure 6-9B). We found that patients with associated metastasis had significantly higher *VIM* and *CD44* expression. Our findings suggest that *VIM* and *CD44* expression is associated with increased thyroid cancer aggression.
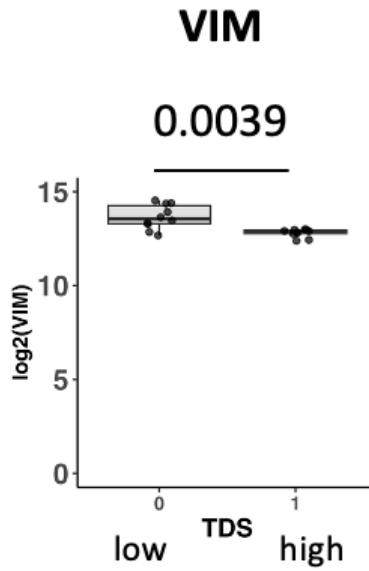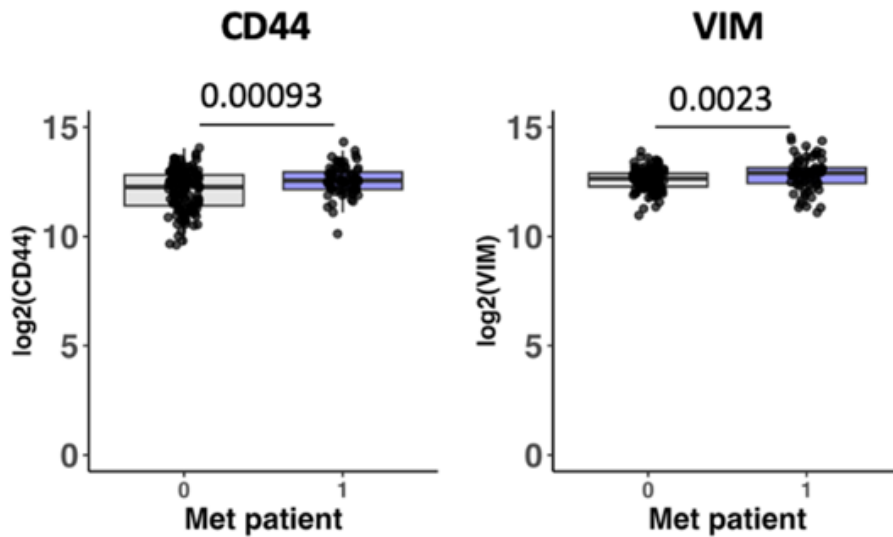
Figure 6-9: Gene expression of *VIM* and *CD44* are associated with aggressive thyroid cancer

A) Gene expression for *VIM* in ATCs with either low or high TDS score. P values were calculated with Wilcoxon rank-sum test.

B) Gene expression for *CD44* and *VIM* for all local thyroid tumors with either no associated metastatic disease (0) or associated metastatic disease (1). P values were calculated with Wilcoxon rank-sum test.

6.4 Discussion

Our findings suggest that Wnt signaling in ATCs may be primarily Wnt ligand-driven rather than mutation-driven, and that this signaling may play a role in ATC cell de-differentiation and invasion. *CDH3*, *LAMA3*, *LAMB3*, and *LAMC2* are all potential candidates involved downstream of Wnt signaling that may affect cell invasion. Furthermore, we found evidence that *CD44* and *VIM* may also be activated by Wnt signaling, and that increased *CD44* and *VIM* gene expression are associated with metastasis and loss of thyroid differentiation markers. Additional experiments are needed to determine whether these genes are involved in Wnt-mediated cell adhesion or other changes related to aggressive cancer behavior. For example, experiments could test the effect of knocking down any of the previously mentioned genes in ATC cell lines in the presence of a Wnt signaling activator, and then looking at the effect of the knockdown on cell adhesion, migration, or gene expression.

CHAPTER 7

Title of Chapter: Canonical Wnt Pathway and E3 ubiquitin ligases

This chapter includes adaptation of contents from the following manuscript:

Kassel, S., et al. USP47 deubiquitylates Groucho/TLE to Promote Wnt-β-catenin Signaling. *Science Signaling*. 2023.

Contributions: I performed statistical analysis and methods writing.

7.1 Introduction

Canonical Wnt signaling is fundamental in normal development and adult cell homeostasis, while its dysregulation is a key factor in many cancers. E3 ubiquitin ligases are important regulators of activities such as homeostasis, the cell cycle, and DNA repair, making them key players across a range of cancer types.[122,123] In addition, E3 ubiquitin ligases are involved in critical roles to the canonical Wnt pathway; for example, the E3 ubiquitin ligase β-Trcp is required for ubiquitinating β-catenin, marking it for proteasomal degradation.[124] However, other Wnt signaling factors may also be targets of E3 ubiquitin ligase activity. For example, previous research by the Lee Lab identified the E3 ubiquitin ligase, X-linked inhibitor of apoptosis (XIAP), as a positive regulator of Wnt signaling by supporting Gro/TLE ubiquitylation. During Wnt signaling, XIAP is recruited to TCF/LEF to promote ubiquitylation of Gro/TLE that is bound to TCF/LEF, decreasing Gro/TLE affinity for TCF/LEF and allow β-catenin to effectively compete for TCF/LEF binding.[125,126] In the paper published by Kassel et. al., we show both in cells and in vitro that the deubiquitylase ubiquitin-specific protease 47 (USP47) reverses the action of XIAP by deubiquitylating Gro/TLE, enhancing the ability of β-catenin to cycle off and on TCF.

TRIP12 is another E3 ubiquitin ligase that may play an important role in several cancers. TRIP12 is classically considered a tumor repressor, with a number of studies reporting anti-tumor

roles.[127,128] Mutations in the TRIP12 gene are relatively uncommon compared to other well-known cancer drivers; for example, data on 27,235 patients from the cBioPortal database shows that the *TRIP12* gene is altered in just 1.84% of patients.[129] Gene expression data from the Gepia2 database shows that different types of cancer may be associated with either downregulation or upregulation of TRIP12 gene expression.[129] Recent research has begun to look at the potential role TRIP12 may play in the Wnt pathway. In this chapter, I focus on analyses comparing patient survival and *USP47* and *TRIP12* gene expression in an external cohorts of pancreatic adenocarcinoma patients, as well as gene expression analysis using data from our internal thyroid lesion patient cohort.

7.2 Methods

To perform survival analysis using cBioPortal, we used data from a publicly available pancreatic cancer cohort, Pancreatic Adenocarcinoma TCGA.[130] Patients were grouped by *USP47* mRNA expression into those with high expression and those with medium to low expression, before comparing probability of overall survival between each group. For survival analysis with *TRIP12,* we used Z-scores of *TRIP12* mRNA expression (log 2 RSEM-UQ), and grouped patients by a median expression cutoff before comparing probability of overall survival between each group.

For our thyroid lesion patient cohort, all sequencing and gene expression analysis used for this project were the same as those described in section 2.2.

7.3 Results

Studies have suggested that USP47 is increased in several types of cancer and may be associated with aggressive cancer behaviors, such as metastasis in breast cancer.[131] To further investigate the role of USP47 in cancer aggression, we performed survival analysis using an external publicly available database from The Cancer Genome Atlas and the online tool

cBioPortal.[132,133] We found that for patients with pancreatic cancer, high amounts of USP47 were significantly associated with shorter patient survival time (Figure 7-1).

Our analysis for TRIP12 was performed similarly to our analysis of USP47, again using The Cancer Genome Atlas and the online tool cBioPortal. We categorized pancreatic cancer patients by high or low *TRIP12* gene expression using a 50th percentile cutoff (Figure 7-2) and found that the group of patients with high TRIP12 gene expression had significantly shorter survival time.

We next performed preliminary analysis to look at the relevance of *USP47* and *TRIP12* in thyroid cancer. Our cBioPortal analysis of the TCGA thyroid cancer cohort did not suggest an association between *USP47* or *TRIP12* and patient survival (data not shown), so we decided to perform preliminary analysis of these genes in our own thyroid lesion cohort. Our gene expression heatmap containing thyroid lesion samples local to the thyroid (Figure 7-3A) did not indicate a clear relationship between canonical Wnt signaling and *USP47* or *TRIP12* expression. We also show a volcano plot of differential expression of genes in either aggressive or indolent local thyroid lesions (Figure 7-3B). Our results suggest that *USP47* or *TRIP12* expression is not associated with aggressive thyroid lesions, unlike *Wnt2* expression.
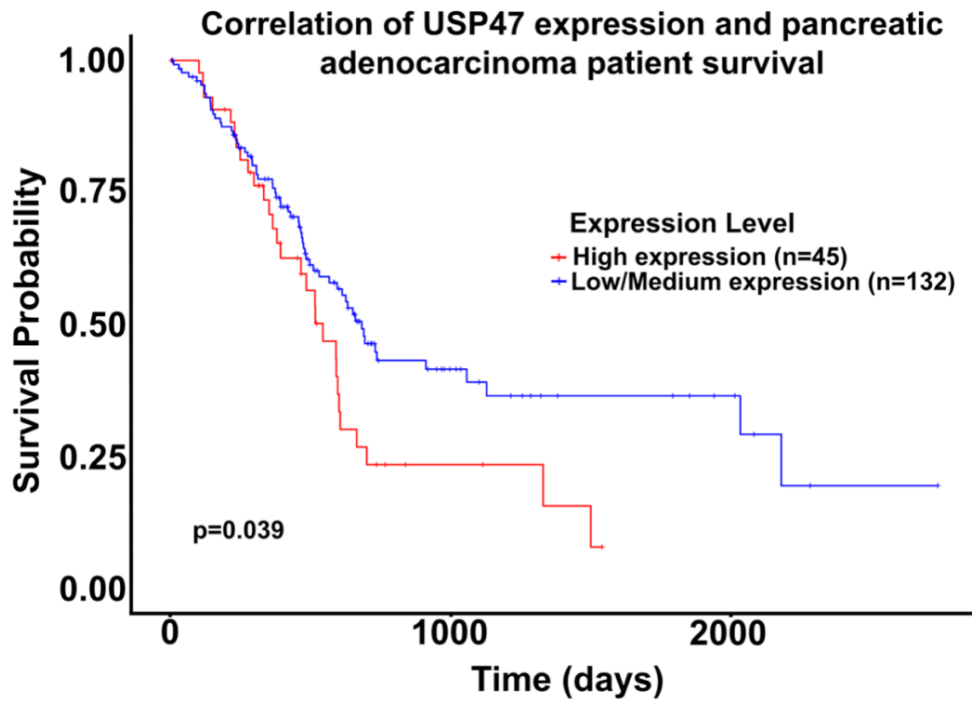
Figure 7-1: *USP47* gene expression and survival in pancreatic adenocarcinoma patients
Survival curve comparing TCGA pancreatic adenocarcinoma patients with high vs medium or low *USP47* mRNA expression. Significance was calculated by the log-rank test (p-value = 0.039).

Overall Survival Over Time (Months)
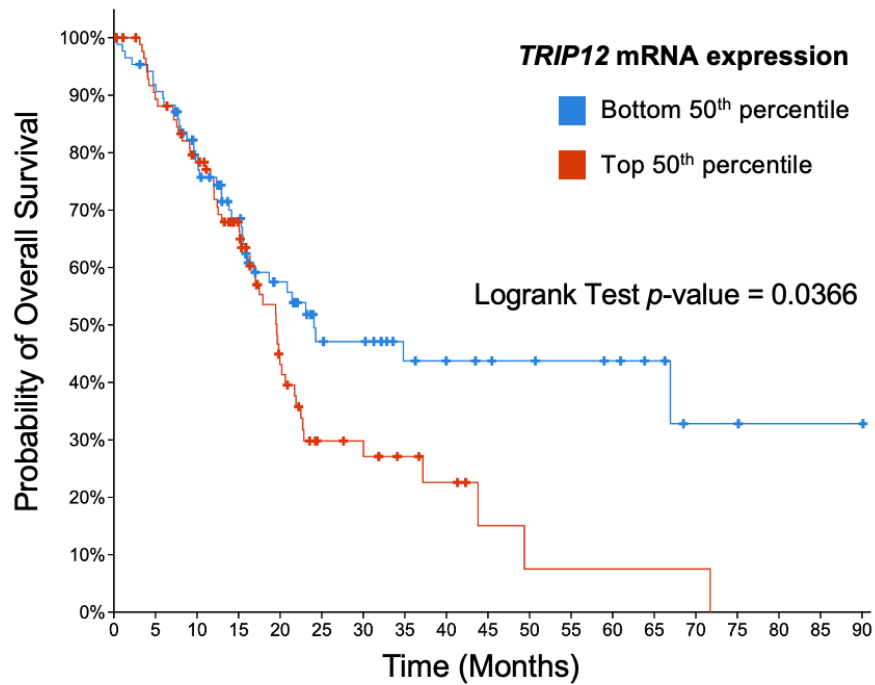
TRIP12 mRNA expression	Number of Cases, Total	Number of Events	Median Months Overall (95% CI)
Bottom 50th percentile	88	39	24.07 (16.80 - NA)
Top 50th percentile	89	53	19.59 (17.00 – 21.90)

Figure 7-2: *TRIP12* **gene expression and survival in pancreatic adenocarcinoma patients** Survival curve comparing TCGA pancreatic adenocarcinoma patients with top vs bottom 50th percentile *TRIP12* mRNA expression. Table at the bottom summarizes counts of number of cases, events, and median months overall for each group. Significance was calculated by the log-rank test.
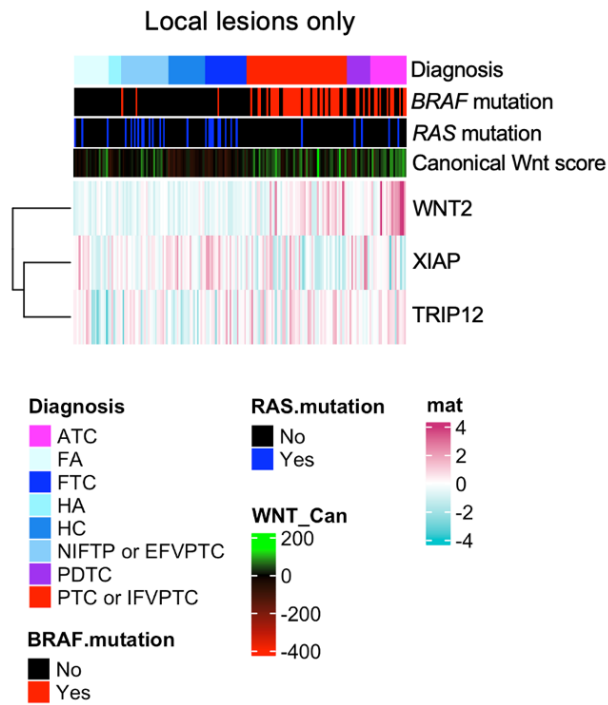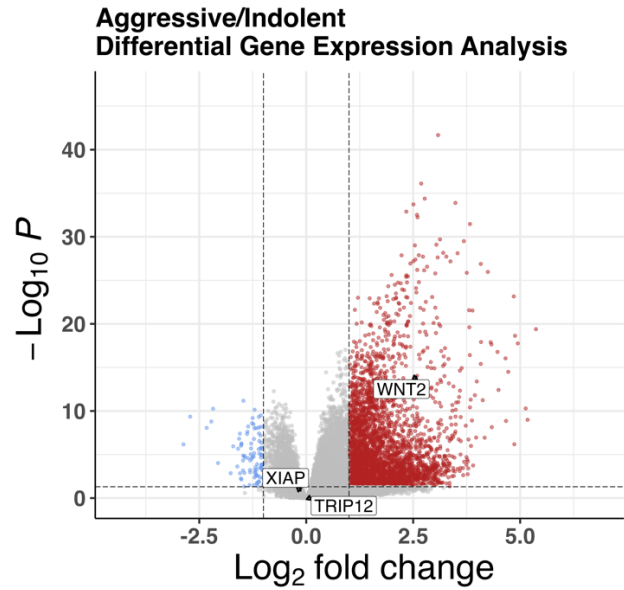
Figure 7-3: *XIAP* and *TRIP12* gene expression in local thyroid lesions
A) Normalized gene expression heatmap of local thyroid lesions showing *WNT2, XIAP, and TRIP12* gene expression. Annotations include diagnosis, BRAF mutation status, RAS mutation status, canonical Wnt signaling score, and cancer-associated fibroblast (CAF) score (EPIC).
B) Volcano plot comparing differential expression of genes enriched in aggressive (right) vs indolent (left) local thyroid lesions. *WNT2, XIAP, and TRIP12* are labeled. Dotted lines indicate fold-change > 2 and adjusted p value < 0.05.

7.4 Discussion

In summary, our findings suggest that both *USP47* and *TRIP12* are significantly associated with patient survival in pancreatic cancer but may not be involved to the same degree in thyroid cancer. We note that USP47, in addition to its role of deubiquitylating Gro/TLE to enhance β-catenin cycling off and on TCF, has previously been shown to support Wnt signaling through other interactions such as deubiquitylating β-catenin.[134] In addition, USP46 has also been reported to participate in DNA damage repair, cell adhesion, and epithelial-mesenchymal transition.[135] Therefore, our finding of reduced survival time with increased *TRIP12* expression has several potential mechanisms that could support increased cancer aggression. We also note that previous studies have also shown that pancreatic cancer is associated with recurrent inactivating mutations in RNF43, a ubiquitin ligase that targets Wnt receptors.[136] Future studies may include further investigation a potential role of increased USP47 expression in potentiating Wnt signaling in pancreatic cancers also containing RNF43 mutations. Regarding TRIP12, the relationship between *TRIP12* gene expression and overall survival in pancreatic cancer patients is interesting, but its relationship with Wnt signaling is still unclear. While TRIP12 has known roles in regulating major biological processes and known alterations reported in cancers, there are currently few cancer studies in previous literature investigating its relationship with the Wnt pathway. Future research is needed to resolve any possible interactions between TRIP12 and Wnt signaling.

CHAPTER 8

Title of Chapter: Discussion and Future Directions

8.1 Implications

Despite recent advances in cancer molecular profiling, thyroid cancer pathogenesis remains poorly understood compared to similarly prevalent cancers. My thesis research aims to improve our understanding of the molecular and microenvironmental markers of thyroid cancer, particularly aggressive thyroid cancers such as anaplastic thyroid cancer. Using DNA and RNA sequencing on a large patient cohort, we describe the genetic and transcriptional landscape of a wide range of thyroid lesion subtypes. We found that commonly described primary and secondary mutations known in thyroid cancer were present but were often insufficient to explain thyroid cancer aggression and increased Wnt signaling.

Several molecular diagnostic scores for thyroid cancer have been developed for malignancy prediction,[44,45] but tools for predicting aggressive thyroid cancer remain limited. With the need in mind, we created our molecular aggression and prediction (MAP) score, which demonstrates the potential to improve prognostication of aggressive thyroid cancer, even in patients lacking mutations commonly associated with aggressive disease. Our MAP score highlights the tumor microenvironment, including stromal cells as well as components of the immune infiltrate, providing new insight into potential drivers of aggression (Figure 8). In particular, our MAP score demonstrates the ability to classify the highly aggressive thyroid cancer subtype ATC, revealing two unique patterns of immune and stromal cell infiltration that are associated with differential response to checkpoint blockade immunotherapy.

In addition, our findings on canonical Wnt signaling in thyroid cancer, particularly anaplastic thyroid cancer, may provide further opportunities for biomarker discovery and therapeutic strategies. One of our novel findings is that Wnt signaling is upregulated through

increased Wnt ligand expression without widespread mutations in Wnt pathway genes as reported in previous studies of ATC. We found that thyroid lesions with higher Wnt signaling score also had expression of select laminin and cadherin genes, suggesting a potential link between these adhesion molecules and Wnt signaling. In addition, we found that ATCs with higher predicted Wnt signaling score also had higher predicted CAF infiltrate, as well as increased expression of *CD44* and *VIM*, which could point towards a role for Wnt signaling in modifying cancer cell adhesion and invasion in ATC. While more research is needed to determine the mechanisms at play, our findings could help open the path towards new therapeutic strategies targeting upregulated Wnt signaling in aggressive thyroid cancers like ATC.
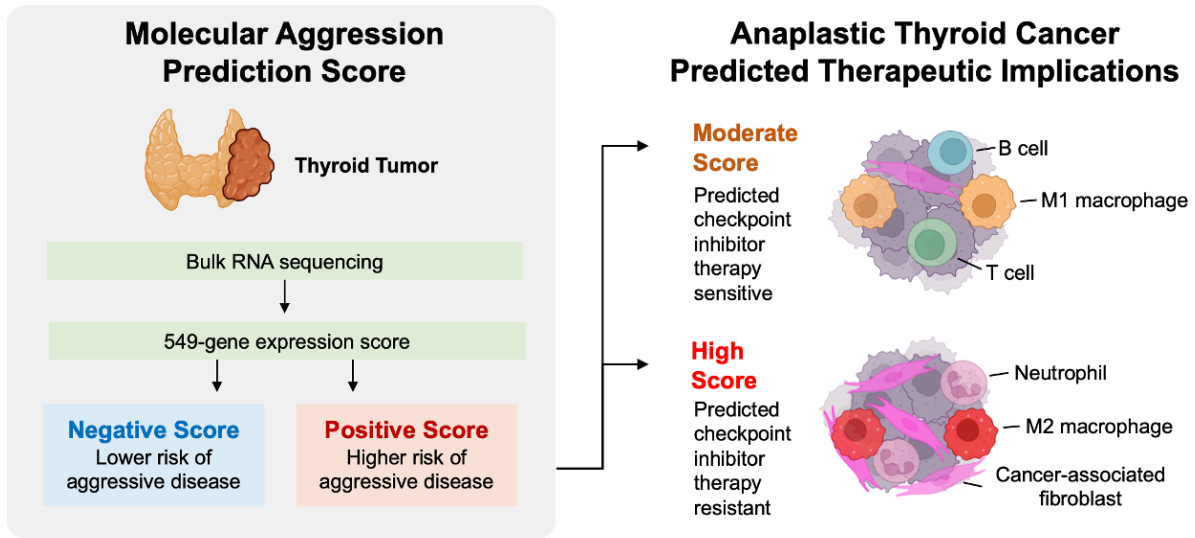
Figure 8: Diagram Summarizing MAP Score Findings
Bulk RNA sequencing was used to generate a 549-gene expression score incorporating tumor microenvironment markers. Score differentiates between low and high risk of aggressive disease in thyroid lesions. Score also categorizes ATCs into two groups with unique tumor-infiltrating cell makeup and predicted checkpoint inhibitor therapy response.

8.2 Limitations

8.2.1 Sequencing and Computational Limitations

We note a number of limitations relating to our sequencing and computational methods. Our study was not oriented specifically for mutation discovery, so it is possible that some mutations that could be used predict aggressive disease remain undetected in our cohort. The depth of our targeted whole-exome sequencing could have limited our ability to detect subclonal mutations or mutations in samples with lower tumor purity. While we worked to sequence samples with a high percent tumor, some samples still contained a lower percentage of cancer cells, such as those that were stromal-rich. Detection of subclonal mutations in lower percent tumor samples would be inherently limited. In addition, our analysis did not include copy number alterations, which have been implicated in thyroid cancer aggression in recent studies.[137,138] We also note that patient samples used for sequencing were formalin-fixed paraffin-embedded (FFPE) tissues samples, which may be subject to reduced DNA quality. In addition, bulk RNA sequencing and spatial transcriptomics are limited in their ability to subclassify small immune cell populations and CAF subsets.

Furthermore, we note that all previously published gene expression-based scores, including computational deconvolution scores, may be limited in their ability to estimate true cellular content. While we provide non-computational methods to support our computational results, additional research using external cohorts can help further validate our findings. Additionally, we note that our study of papillary thyroid microcarcinomas (PTMs) with distant metastasis is limited by our small sample size of only two patients. This limitation is not unexpected given the relative rarity of PTMs with distant metastasis. However, future research involving more PTM samples with distant metastasis is needed to improve our understanding of these unusual cases.

8.2.2 Resection Sample Limitations

Molecular tools for thyroid cancer prediction typically use a minimally-invasive technique called fine needle aspiration (FNA) to collect biopsy material. While our MAP score performed well using FFPE resection samples, the potential clinical utility of this gene score is unclear without additional research to determine whether our MAP score can also perform well using FNA samples. FNA may enrich for specific cell types, such as tumor cells or certain immune cell types, while other cell types such as CAFs may be less amenable to isolation by FNA. If not accounted for, such variations would negatively impact the accuracy of our current MAP score. To begin investigating these questions, our lab has performed flow cytometry of thyroid cancer aspirate biopsy material and confirmed the robust identification of neutrophils, macrophages, and T cells.

8.2.3 Retrospective Study Limitations

All patient data used for my thesis research was collected retrospectively. Therefore, we faced some limitations in our ability to complete collection of certain clinical data variables for some patients, such as those with incomplete records of disease earlier in life, or those lost to follow-up. In addition, the real-world clinical utility of our MAP score for predicting thyroid cancer aggression and response to immunotherapy could not be tested. Based on our available data, none of the patients in our cohort were recorded to have been treated with immunotherapy, so ICB response could only be estimated computationally.

8.2.4 Cost Effectiveness in Healthcare

Our findings of gene expression and stromal markers of aggressive disease in thyroid cancer may open up potential avenues for new predictive tools and therapies. However, implementation of new practices involving sequencing of large panels of genes may be both costly and impractical, at least in the near future. Additional research is needed to demonstrate the real-world utility of these tests over existing methods; as the majority of thyroid cancer lesions are

indolent, the need for improved prediction of future thyroid cancer aggression may not be enough to outweigh the increased cost of adopting novel testing methods. Sudden application of such tests to the vast number of new indolent thyroid lesion cases every year could potentially overwhelm healthcare systems and is unlikely to be compatible with existing sequencing resources. However, as advancements in personalized medicine continue to appear and high-throughput sequencing continues to fall in cost, we foresee a future where widespread use of such tests becomes more feasible.

8.3 Future Directions

8.3.1 Investigate the Relationship between Wnt Signaling and the Tumor Microenvironment

Our research revealed distinct patterns of infiltrating immune cells and stromal cells in the microenvironment of thyroid cancers with differing levels of aggression. In addition, we found that Wnt signaling appeared to be associated with CAF levels, patient survival, and aggressive characteristics such as dedifferentiation and metastasis. Future research is needed to more closely investigate the mechanisms linking Wnt signaling to the tumor microenvironment and thyroid cancer aggression. Other research in our lab is currently investigating a possible role of CAFs secreting WNT2 ligand to recruit monocytes, drive polarization favoring M2 macrophages, and altering PD-L1 expression. As only adults were included in the studies in this paper, future studies may also investigate the role of tumor infiltrating cells in pediatric thyroid cancer cases, which have not been previously profiled for the tumor microenvironment.

8.3.2 Fine-needle Aspiration (FNA) and Tumor Microenvironment Profiling

As we explained previously, FNA may enrich for tumor cells and certain immune cell types, while other cell types such as CAFs may be harder to collect. Our MAP score depends on cells

in the tumor microenvironment, so large biopsy-based studies will be needed to assess the differential cell enrichment in aspirate biopsies and determine what modifications are required for the MAP score to continue providing prognostic information in FNA.

8.3.3 Prospective Study of Immunotherapy Response and MAP Score

To more thoroughly assess whether the tumor's stromal infiltrate can be used as an accurate predictor of ICB response, future studies are needed with more thyroid cancer patients receiving ICB therapy. We anticipate such studies becoming more feasible as immunotherapy becomes more widely used to treat thyroid cancer.

8.4 Concluding Remarks

In conclusion, our findings emphasize the importance of the thyroid tumor microenvironment as a source of biomarkers for informing improvements in therapy and prediction of aggressive disease. Our MAP score incorporating stromal genes demonstrates the potential to improve thyroid cancer risk-stratification and predict immunotherapy response, and we envision future testing platforms that take advantage of both mutational and stromal microenvironment data for better outcome and ICB response prediction. Continued research on the thyroid cancer microenvironment also has the potential to identify novel targets for therapy, especially for ATC, the most aggressive form of thyroid cancer which has virtually no cure. Future research of the stromal microenvironment has the potential to deepen our understanding of cancer biology and redefine tumor classification, and molecular tests incorporating stromal genes have the potential to inform treatment across a wide range of stromal-rich cancers.

References

1. Rahib, L., Smith, B.D., Aizenberg, R., Rosenzweig, A.B., Fleshman, J.M., and Matrisian, L.M. (2014). Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer Res *74*, 2913-2921. 10.1158/0008-5472.CAN-14-0155.

2. Siegel, R., Ward, E., Brawley, O., and Jemal, A. (2011). Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA Cancer J Clin *61*, 212-236. 10.3322/caac.20121.

3. Haugen, B.R., Alexander, E.K., Bible, K.C., Doherty, G.M., Mandel, S.J., Nikiforov, Y.E., Pacini, F., Randolph, G.W., Sawka, A.M., Schlumberger, M., et al. (2016). 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid *26*, 1-133. 10.1089/thy.2015.0020.

4. Filetti, S., Durante, C., Hartl, D., Leboulleux, S., Locati, L.D., Newbold, K., Papotti, M.G., Berruti, A., and clinicalguidelines@esmo.org, E.G.C.E.a. (2019). Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Ann Oncol *30*, 1856-1883. 10.1093/annonc/mdz400.

5. Duggan, M.A., Anderson, W.F., Altekruse, S., Penberthy, L., and Sherman, M.E. (2016). The Surveillance, Epidemiology, and End Results (SEER) Program and Pathology: Toward Strengthening the Critical Relationship. Am J Surg Pathol *40*, e94-e102. 10.1097/PAS.0000000000000749.

6. Hwangbo, Y., Kim, J.M., Park, Y.J., Lee, E.K., Lee, Y.J., Park, D.J., Choi, Y.S., Lee, K.D., Sohn, S.Y., Kim, S.W., et al. (2017). Long-Term Recurrence of Small Papillary Thyroid Cancer and Its Risk Factors in a Korean Multicenter Study. J Clin Endocrinol Metab *102*, 625-633. 10.1210/jc.2016-2287.

7. Nagaiah, G., Hossain, A., Mooney, C.J., Parmentier, J., and Remick, S.C. (2011). Anaplastic thyroid cancer: a review of epidemiology, pathogenesis, and treatment. J Oncol *2011*, 542358. 10.1155/2011/542358.

8. Nikiforov, Y.E., and Nikiforova, M.N. (2011). Molecular genetics and diagnosis of thyroid cancer. Nat Rev Endocrinol *7*, 569-580. 10.1038/nrendo.2011.142.

9. Cohen, Y., Xing, M., Mambo, E., Guo, Z., Wu, G., Trink, B., Beller, U., Westra, W.H., Ladenson, P.W., and Sidransky, D. (2003). BRAF mutation in papillary thyroid carcinoma. J Natl Cancer Inst *95*, 625-627. 10.1093/jnci/95.8.625.

10. Xing, M. (2005). BRAF mutation in thyroid cancer. Endocr Relat Cancer *12*, 245-262. 10.1677/erc.1.0978.

11. Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W., et al. (2002). Mutations of the BRAF gene in human cancer. Nature *417*, 949-954. 10.1038/nature00766.

12. Boos, L.A., Dettmer, M., Schmitt, A., Rudolph, T., Steinert, H., Moch, H., Sobrinho-Simões, M., Komminoth, P., and Perren, A. (2013). Diagnostic and prognostic implications of the PAX8-PPARγ translocation in thyroid carcinomas-a TMA-based study of 226 cases. Histopathology *63*, 234-241. 10.1111/his.12150.

13. Yip, L., Gooding, W.E., Nikitski, A., Wald, A.I., Carty, S.E., Karslioglu-French, E., Seethala, R.R., Zandberg, D.P., Ferris, R.L., Nikiforova, M.N., and Nikiforov, Y.E. (2021). Risk assessment for distant metastasis in differentiated thyroid cancer using molecular profiling: A matched case-control study. Cancer *127*, 1779-1787. 10.1002/cncr.33421.

14. Chen, J. (2016). The Cell-Cycle Arrest and Apoptotic Functions of p53 in Tumor Initiation and Progression. Cold Spring Harb Perspect Med *6*, a026104. 10.1101/cshperspect.a026104.
15. Cho, Y., Gorina, S., Jeffrey, P.D., and Pavletich, N.P. (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. Science *265*, 346-355. 10.1126/science.8023157.
16. Engelman, J.A., Luo, J., and Cantley, L.C. (2006). The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. Nat Rev Genet *7*, 606-619. 10.1038/nrg1879.
17. García-Rostán, G., Costa, A.M., Pereira-Castro, I., Salvatore, G., Hernandez, R., Hermsem, M.J., Herrero, A., Fusco, A., Cameselle-Teijeiro, J., and Santoro, M. (2005). Mutation of the PIK3CA gene in anaplastic thyroid cancer. Cancer Res *65*, 10199-10207. 10.1158/0008-5472.CAN-04-4259.
18. Hou, P., Liu, D., Shan, Y., Hu, S., Studeman, K., Condouris, S., Wang, Y., Trink, A., El-Naggar, A.K., Tallini, G., et al. (2007). Genetic alterations and their relationship in the phosphatidylinositol 3-kinase/Akt pathway in thyroid cancer. Clin Cancer Res *13*, 1161-1170. 10.1158/1078-0432.CCR-06-1125.
19. Liu, R., and Xing, M. (2016). TERT promoter mutations in thyroid cancer. Endocr Relat Cancer *23*, R143-155. 10.1530/ERC-15-0533.
20. Ceja-Rangel, H.A., Sánchez-Suárez, P., Castellanos-Juárez, E., Peñaroja-Flores, R., Arenas-Aranda, D.J., Gariglio, P., and Benítez-Bribiesca, L. (2016). Shorter telomeres and high telomerase activity correlate with a highly aggressive phenotype in breast cancer cell lines. Tumour Biol *37*, 11917-11926. 10.1007/s13277-016-5045-7.
21. Wang, Y., Meeker, A.K., Kowalski, J., Tsai, H.L., Somervell, H., Heaphy, C., Sangenario, L.E., Prasad, N., Westra, W.H., Zeiger, M.A., and Umbricht, C.B. (2011). Telomere length is related to alternative splice patterns of telomerase in thyroid tumors. Am J Pathol *179*, 1415-1424. 10.1016/j.ajpath.2011.05.056.
22. Low, K.C., and Tergaonkar, V. (2013). Telomerase: central regulator of all of the hallmarks of cancer. Trends Biochem Sci *38*, 426-434. 10.1016/j.tibs.2013.07.001.
23. StatPearls. (2023). In. NBK536943.
24. Landa, I., Ganly, I., Chan, T.A., Mitsutake, N., Matsuse, M., Ibrahimpasic, T., Ghossein, R.A., and Fagin, J.A. (2013). Frequent somatic TERT promoter mutations in thyroid cancer: higher prevalence in advanced forms of the disease. J Clin Endocrinol Metab *98*, E1562-1566. 10.1210/jc.2013-2383.
25. Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell *61*, 759-767. 10.1016/0092-8674(90)90186-i.
26. Hoang-Vu, C., Dralle, H., Scheumann, G., Maenhaut, C., Horn, R., von zur Mühlen, A., and Brabant, G. (1992). Gene expression of differentiation- and dedifferentiation markers in normal and malignant human thyroid tissues. Exp Clin Endocrinol *100*, 51-56. 10.1055/s-0029-1211176.
27. S., I. Histology   PathologyOutlines.com.
28. Kwong, L.N., and Dove, W.F. (2009). APC and its modifiers in colon cancer. Adv Exp Med Biol *656*, 85-106. 10.1007/978-1-4419-1145-2_8.
29. Kurihara, T., Ikeda, S., Ishizaki, Y., Fujimori, M., Tokumoto, N., Hirata, Y., Ozaki, S., Okajima, M., Sugino, K., and Asahara, T. (2004). Immunohistochemical and sequencing analyses of the Wnt signaling components in Japanese anaplastic thyroid cancers. Thyroid *14*, 1020-1029. 10.1089/thy.2004.14.1020.
30. Garcia-Rostan, G., Tallini, G., Herrero, A., D'Aquila, T.G., Carcangiu, M.L., and Rimm, D.L. (1999). Frequent mutation and nuclear localization of beta-catenin in anaplastic thyroid carcinoma. Cancer Res *59*, 1811-1815.

31. Garcia-Rostan, G., Camp, R.L., Herrero, A., Carcangiu, M.L., Rimm, D.L., and Tallini, G. (2001). Beta-catenin dysregulation in thyroid neoplasms: down-regulation, aberrant nuclear expression, and CTNNB1 exon 3 mutations are markers for aggressive tumor phenotypes and poor prognosis. Am J Pathol *158*, 987-996. 10.1016/s0002-9440(10)64045-x.

32. Kumari, S., Adewale, R., and Klubo-Gwiezdzinska, J. (2020). The Molecular Landscape of Hürthle Cell Thyroid Cancer Is Associated with Altered Mitochondrial Function-A Comprehensive Review. Cells *9*. 10.3390/cells9071570.

33. Gillespie, J.W., Nasir, A., and Kaiser, H.E. (2000). Loss of heterozygosity in papillary and follicular thyroid carcinoma: a mini review. In Vivo *14*, 139-140.

34. Kadota, M., Tamaki, Y., Sekimoto, M., Fujiwara, Y., Aritake, N., Hasegawa, S., Kobayashi, T., Ikeda, T., Horii, A., and Monden, M. (2003). Loss of heterozygosity on chromosome 16p and 18q in anaplastic thyroid carcinoma. Oncol Rep *10*, 35-38.

35. Hemmer, S., Wasenius, V.M., Knuutila, S., Franssila, K., and Joensuu, H. (1999). DNA copy number changes in thyroid carcinoma. Am J Pathol *154*, 1539-1547. 10.1016/S0002-9440(10)65407-7.

36. Yakushina, V.D., Lerner, L.V., and Lavrov, A.V. (2018). Gene Fusions in Thyroid Cancer. Thyroid *28*, 158-167. 10.1089/thy.2017.0318.

37. Galdiero, M.R., Varricchi, G., Loffredo, S., Bellevicine, C., Lansione, T., Ferrara, A.L., Iannone, R., di Somma, S., Borriello, F., Clery, E., et al. (2018). Potential involvement of neutrophils in human thyroid cancer. PLoS One *13*, e0199740. 10.1371/journal.pone.0199740.

38. Fang, W., Ye, L., Shen, L., Cai, J., Huang, F., Wei, Q., Fei, X., Chen, X., Guan, H., Wang, W., et al. (2014). Tumor-associated macrophages promote the metastatic potential of thyroid papillary cancer by releasing CXCL8. Carcinogenesis *35*, 1780-1787. 10.1093/carcin/bgu060.

39. Lu, L., Wang, J.R., Henderson, Y.C., Bai, S., Yang, J., Hu, M., Shiau, C.K., Pan, T.Y., Yan, Y., Tran, T.M., et al. (2023). Anaplastic transformation in thyroid cancer revealed by single cell transcriptomics. J Clin Invest. 10.1172/JCI169653.

40. Hao, N.B., Lü, M.H., Fan, Y.H., Cao, Y.L., Zhang, Z.R., and Yang, S.M. (2012). Macrophages in tumor microenvironments and the progression of tumors. Clin Dev Immunol *2012*, 948098. 10.1155/2012/948098.

41. Kalluri, R. (2016). The biology and function of fibroblasts in cancer. Nat Rev Cancer *16*, 582-598. 10.1038/nrc.2016.73.

42. Chen, Y., McAndrews, K.M., and Kalluri, R. (2021). Clinical and therapeutic relevance of cancer-associated fibroblasts. Nat Rev Clin Oncol *18*, 792-804. 10.1038/s41571-021-00546-5.

43. Wen, S., Qu, N., Ma, B., Wang, X., Luo, Y., Xu, W., Jiang, H., Zhang, Y., Wang, Y., and Ji, Q. (2021). Cancer-Associated Fibroblasts Positively Correlate with Dedifferentiation and Aggressiveness of Thyroid Cancer. Onco Targets Ther *14*, 1205-1217. 10.2147/OTT.S294725.

44. Hu, M.I., Waguespack, S.G., Dosiou, C., Ladenson, P.W., Livhits, M.J., Wirth, L.J., Sadow, P.M., Krane, J.F., Stack, B.C., Zafereo, M.E., et al. (2021). Afirma Genomic Sequencing Classifier and Xpression Atlas Molecular Findings in Consecutive Bethesda III-VI Thyroid Nodules. J Clin Endocrinol Metab *106*, 2198-2207. 10.1210/clinem/dgab304.

45. Steward, D.L., Carty, S.E., Sippel, R.S., Yang, S.P., Sosa, J.A., Sipos, J.A., Figge, J.J., Mandel, S., Haugen, B.R., Burman, K.D., et al. (2019). Performance of a Multigene Genomic Classifier in Thyroid Nodules With Indeterminate Cytology: A Prospective Blinded Multicenter Study. JAMA Oncol *5*, 204-212. 10.1001/jamaoncol.2018.4616.

46. Xing, M. (2013). Molecular pathogenesis and mechanisms of thyroid cancer. Nat Rev Cancer *13*, 184-199. 10.1038/nrc3431.

47. Lemoine, N.R., Mayall, E.S., Wyllie, F.S., Farr, C.J., Hughes, D., Padua, R.A., Thurston, V., Williams, E.D., and Wynford-Thomas, D. (1988). Activated ras oncogenes in human thyroid cancers. Cancer Res *48*, 4459-4463.

48. Nikiforova, M.N., Mercurio, S., Wald, A.I., Barbi de Moura, M., Callenberg, K., Santana-Santos, L., Gooding, W.E., Yip, L., Ferris, R.L., and Nikiforov, Y.E. (2018). Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. Cancer *124*, 1682-1690. 10.1002/cncr.31245.

49. Finkelstein, S.D., Sistrunk, J.W., Malchoff, C., Thompson, D.V., Kumar, G., Timmaraju, V.A., Repko, B., Mireskandari, A., Evoy-Goodman, L.A., Massoll, N.A., and Lupo, M.A. (2022). A Retrospective Evaluation of the Diagnostic Performance of an Interdependent Pairwise MicroRNA Expression Analysis with a Mutation Panel in Indeterminate Thyroid Nodules. Thyroid *32*, 1362-1371. 10.1089/thy.2022.0124.

50. Lupo, M.A., Walts, A.E., Sistrunk, J.W., Giordano, T.J., Sadow, P.M., Massoll, N., Campbell, R., Jackson, S.A., Toney, N., Narick, C.M., et al. (2020). Multiplatform molecular test performance in indeterminate thyroid nodules. Diagn Cytopathol *48*, 1254-1264. 10.1002/dc.24564.

51. Khan, T.M., and Zeiger, M.A. (2020). Thyroid Nodule Molecular Testing: Is It Ready for Prime Time? Front Endocrinol (Lausanne) *11*, 590128. 10.3389/fendo.2020.590128.

52. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.

53. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data.

54. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589-595. 10.1093/bioinformatics/btp698.

55. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res *20*, 1297-1303. 10.1101/gr.107524.110.

56. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res *38*, e164. 10.1093/nar/gkq603.

57. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285-291. 10.1038/nature19057.

58. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434-443. 10.1038/s41586-020-2308-7.

59. Li, M.M., Datto, M., Duncavage, E.J., Kulkarni, S., Lindeman, N.I., Roy, S., Tsimberidou, A.M., Vnencak-Jones, C.L., Wolff, D.J., Younes, A., and Nikiforova, M.N. (2017). Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. J Mol Diagn *19*, 4-23. 10.1016/j.jmoldx.2016.10.002.

60. Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., and Koeffler, H.P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res *28*, 1747-1756. 10.1101/gr.239244.118.

61. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847-2849. 10.1093/bioinformatics/btw313.
62. Koelsche, C., Renner, M., Hartmann, W., Brandt, R., Lehner, B., Waldburger, N., Alldinger, I., Schmitt, T., Egerer, G., Penzel, R., et al. (2014). TERT promoter hotspot mutations are recurrent in myxoid liposarcomas but rare in other soft tissue sarcoma entities. J Exp Clin Cancer Res *33*, 33. 10.1186/1756-9966-33-33.
63. Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Armstrong, J., Barnes, I., et al. (2021). GENCODE 2021. Nucleic Acids Res *49*, D916-D923. 10.1093/nar/gkaa1087.
64. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21. 10.1093/bioinformatics/bts635.
65. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930. 10.1093/bioinformatics/btt656.
66. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. 10.1186/s13059-014-0550-8.
67. Blighe, K., Rana, S., and Lewis, M. (2018). EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling. https://bioconductor.org/packages/devel/bioc/vignettes/EnhancedVolcano/inst/doc/EnhancedVolcano.html.
68. Zhao, S., Guo, Y., Sheng, Q., and Shyr, Y. (2014). Heatmap3: an improved heatmap package with more powerful and convenient features.
69. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet *25*, 25-29. 10.1038/75556.
70. Aleksander, S.A., Balhoff, J., Carbon, S., Cherry, J.M., Drabkin, H.J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N.L., Hill, D.P., et al. (2023). The Gene Ontology knowledgebase in 2023. Genetics *224*. 10.1093/genetics/iyad031.
71. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550. 10.1073/pnas.0506580102.
72. Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., Li, B., and Liu, X.S. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res *48*, W509-W514. 10.1093/nar/gkaa407.
73. Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat Methods *12*, 453-457. 10.1038/nmeth.3337.
74. Becht, E., Giraldo, N.A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W.H., and de Reyniès, A. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol *17*, 218. 10.1186/s13059-016-1070-5.
75. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. https://ggplot2.tidyverse.org.

76. Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., Li, Z., Traugh, N., Bu, X., Li, B., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med *24*, 1550-1558. 10.1038/s41591-018-0136-1.

77. Gu, Z. (2022). Complex Heatmap Visualization. *1*, e43.

78. Hass, B., Dobin, A., Stransky, N., Li, B., Yang, X., Tickle, T., Bankapur, A., Ganote, C., Doak, T., and Pochet, N. (2017). STAR-Fusion: fast and accurate fusion transcript detection from RNA-seq. bioRxiv.

79. Stransky, N., Cerami, E., Schalm, S., Kim, J.L., and Lengauer, C. (2014). The landscape of kinase fusions in cancer. Nat Commun *5*, 4846. 10.1038/ncomms5846.

80. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26. 10.1038/nbt.1754.

81. Network, C.G.A.R. (2014). Integrated genomic characterization of papillary thyroid carcinoma. Cell *159*, 676-690. 10.1016/j.cell.2014.09.050.

82. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417-425. 10.1016/j.cels.2015.12.004.

83. Pratilas, C.A., Taylor, B.S., Ye, Q., Viale, A., Sander, C., Solit, D.B., and Rosen, N. (2009). (V600E)BRAF is associated with disabled feedback inhibition of RAF-MEK signaling and elevated transcriptional output of the pathway. Proc Natl Acad Sci U S A *106*, 4519-4524. 10.1073/pnas.0900780106.

84. Patel, J., Klopper, J., and Cottrill, E.E. (2023). Molecular diagnostics in the evaluation of thyroid nodules: Current use and prospective opportunities. Front Endocrinol (Lausanne) *14*, 1101410. 10.3389/fendo.2023.1101410.

85. Xiao, Y., and Yu, D. (2021). Tumor microenvironment as a therapeutic target in cancer. Pharmacol Ther *221*, 107753. 10.1016/j.pharmthera.2020.107753.

86. Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.P., and Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci *31*, 8-22. 10.1002/pro.4218.

87. Kosmidis, I., Kenne Pagui, E.C., Konis, K., and Sartori, N. (2023). brglm2: Bias Reduction in Generalized Linear Models https://CRAN.R-project.org/package=brglm2.

88. Hiltzik, D., Carlson, D.L., Tuttle, R.M., Chuai, S., Ishill, N., Shaha, A., Shah, J.P., Singh, B., and Ghossein, R.A. (2006). Poorly differentiated thyroid carcinomas defined on the basis of mitosis and necrosis: a clinicopathologic study of 58 patients. Cancer *106*, 1286-1295. 10.1002/cncr.21739.

89. Volante, M., Collini, P., Nikiforov, Y.E., Sakamoto, A., Kakudo, K., Katoh, R., Lloyd, R.V., LiVolsi, V.A., Papotti, M., Sobrinho-Simoes, M., et al. (2007). Poorly differentiated thyroid carcinoma: the Turin proposal for the use of uniform diagnostic criteria and an algorithmic diagnostic approach. Am J Surg Pathol *31*, 1256-1264. 10.1097/PAS.0b013e3180309e6a.

90. Ragazzi, M., Ciarrocchi, A., Sancisi, V., Gandolfi, G., Bisagni, A., and Piana, S. (2014). Update on anaplastic thyroid carcinoma: morphological, molecular, and genetic features of the most aggressive thyroid cancer. Int J Endocrinol *2014*, 790834. 10.1155/2014/790834.

91. Pu, W., Shi, X., Yu, P., Zhang, M., Liu, Z., Tan, L., Han, P., Wang, Y., Ji, D., Gan, H., et al. (2021). Single-cell transcriptomic analysis of the tumor ecosystems underlying initiation and progression of papillary thyroid carcinoma. Nat Commun *12*, 6058. 10.1038/s41467-021-26343-3.

92. Jolly, L.A., Novitskiy, S., Owens, P., Massoll, N., Cheng, N., Fang, W., Moses, H.L., and Franco, A.T. (2016). Fibroblast-Mediated Collagen Remodeling Within the Tumor

Microenvironment Facilitates Progression of Thyroid Cancers Driven by BrafV600E and Pten Loss. Cancer Res *76*, 1804-1813. 10.1158/0008-5472.CAN-15-2351.

93. Dierks, C., Seufert, J., Aumann, K., Ruf, J., Klein, C., Kiefer, S., Rassner, M., Boerries, M., Zielke, A., la Rosee, P., et al. (2021). Combination of Lenvatinib and Pembrolizumab Is an Effective Treatment Option for Anaplastic and Poorly Differentiated Thyroid Carcinoma. Thyroid *31*, 1076-1085. 10.1089/thy.2020.0322.

94. Dierks, C., Ruf, J., Seufert, J., Kreissl, M., Klein, C., Spitzweg, C., Kroiss, M., Thomusch, O., Lorenz, K., Zielke, A., and Miething, C. (2022). 1646MO Phase II ATLEP trial: Final results for lenvatinib/pembrolizumab in metastasized anaplastic and poorly differentiated thyroid carcinoma Annals of Oncology *33* (*7*), S1295.

95. Study of Cemiplimab Combined With Dabrafenib and Trametinib in People With Anaplastic Thyroid Cancer.   https://ClinicalTrials.gov/show/NCT04238624.

96. Pembrolizumab, Dabrafenib, and Trametinib Before Surgery for the Treatment of BRAF-Mutated Anaplastic Thyroid Cancer.   https://ClinicalTrials.gov/show/NCT04675710.

97. Lenvatinib and Pembrolizumab for the Treatment of Stage IVB Locally Advanced and Unresectable or Stage IVC Metastatic Anaplastic Thyroid Cancer. https://ClinicalTrials.gov/show/NCT04171622.

98. Atezolizumab With Chemotherapy in Treating Patients With Anaplastic or Poorly Differentiated Thyroid Cancer.   https://ClinicalTrials.gov/show/NCT03181100.

99. Racle, J., and Gfeller, D. (2020). EPIC: A Tool to Estimate the Proportions of Different Cell Types from Bulk Gene Expression Data. Methods Mol Biol *2120*, 233-248. 10.1007/978-1-0716-0327-7_17.

100. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587.e3529. 10.1016/j.cell.2021.04.048.

101. Ru, B., Huang, J., Zhang, Y., Aldape, K., and Jiang, P. (2023). Estimation of cell lineages in tumors from spatial transcriptomics data. Nat Commun *14*, 568. 10.1038/s41467-023-36062-6.

102. Landa, I., Ibrahimpasic, T., Boucai, L., Sinha, R., Knauf, J.A., Shah, R.H., Dogan, S., Ricarte-Filho, J.C., Krishnamoorthy, G.P., Xu, B., et al. (2016). Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. J Clin Invest *126*, 1052-1066. 10.1172/JCI85271.

103. Giannini, R., Moretti, S., Ugolini, C., Macerola, E., Menicali, E., Nucci, N., Morelli, S., Colella, R., Mandarano, M., Sidoni, A., et al. (2019). Immune Profiling of Thyroid Carcinomas Suggests the Existence of Two Major Phenotypes: An ATC-Like and a PDTC-Like. J Clin Endocrinol Metab *104*, 3557-3575. 10.1210/jc.2018-01167.

104. Mao, X., Xu, J., Wang, W., Liang, C., Hua, J., Liu, J., Zhang, B., Meng, Q., Yu, X., and Shi, S. (2021). Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. Mol Cancer *20*, 131. 10.1186/s12943-021-01428-1.

105. Tessler, F.N., Middleton, W.D., Grant, E.G., Hoang, J.K., Berland, L.L., Teefey, S.A., Cronan, J.J., Beland, M.D., Desser, T.S., Frates, M.C., et al. (2017). ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. J Am Coll Radiol *14*, 587-595. 10.1016/j.jacr.2017.01.046.

106. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, G.P.D.P. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079. 10.1093/bioinformatics/btp352.

107. Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices

pipeline. Curr Protoc Bioinformatics *43*, 11.10.11-11.10.33.
10.1002/0471250953.bi1110s43.

108. Yang, H., and Wang, K. (2015). Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc *10*, 1556-1566. 10.1038/nprot.2015.105.

109. Subbiah, V., Kreitman, R.J., Wainberg, Z.A., Cho, J.Y., Schellens, J.H.M., Soria, J.C., Wen, P.Y., Zielinski, C.C., Cabanillas, M.E., Boran, A., et al. (2022). Dabrafenib plus trametinib in patients with BRAF V600E-mutant anaplastic thyroid cancer: updated analysis from the phase II ROAR basket study. Ann Oncol *33*, 406-415. 10.1016/j.annonc.2021.12.014.

110. Castro-Mondragon, J.A., Riudavets-Puig, R., Rauluseviciute, I., Lemma, R.B., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Manosalva Pérez, N., et al. (2022). JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res *50*, D165-D173. 10.1093/nar/gkab1113.

111. Nassar, L.R., Barber, G.P., Benet-Pagès, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C., Gonzalez, J.N., Hinrichs, A.S., Lee, B.T., et al. (2023). The UCSC Genome Browser database: 2023 update. Nucleic Acids Res *51*, D1188-D1195. 10.1093/nar/gkac1072.

112. Kramer, N., Schmöllerl, J., Unger, C., Nivarthi, H., Rudisch, A., Unterleuthner, D., Scherzer, M., Riedl, A., Artaker, M., Crncec, I., et al. (2017). Autocrine WNT2 signaling in fibroblasts promotes colorectal cancer progression. Oncogene *36*, 5460-5472. 10.1038/onc.2017.144.

113. Hwang, P.Y., Mathur, J., Cao, Y., Almeida, J., Ye, J., Morikis, V., Cornish, D., Clarke, M., Stewart, S.A., Pathak, A., and Longmore, G.D. (2023). A Cdh3-β-catenin-laminin signaling axis in a subset of breast tumor leader cells control leader cell polarization and directional collective migration. Dev Cell *58*, 34-50.e39. 10.1016/j.devcel.2022.12.005.

114. Givant-Horwitz, V., Davidson, B., and Reich, R. (2005). Laminin-induced signaling in tumor cells. Cancer Lett *223*, 1-10. 10.1016/j.canlet.2004.08.030.

115. Zhang, J., Woodhead, G.J., Swaminathan, S.K., Noles, S.R., McQuinn, E.R., Pisarek, A.J., Stocker, A.M., Mutch, C.A., Funatsu, N., and Chenn, A. (2010). Cortical neural precursors inhibit their own differentiation via N-cadherin maintenance of beta-catenin signaling. Dev Cell *18*, 472-479. 10.1016/j.devcel.2009.12.025.

116. Sun, L., Hu, H., Peng, L., Zhou, Z., Zhao, X., Pan, J., Yang, Z., and Ran, Y. (2011). P-cadherin promotes liver metastasis and is associated with poor prognosis in colon cancer. Am J Pathol *179*, 380-390. 10.1016/j.ajpath.2011.03.046.

117. Sridhar, S., Rajesh, C., Jishnu, P.V., Jayaram, P., and Kabekkodu, S.P. (2020). Increased expression of P-cadherin is an indicator of poor prognosis in breast cancer: a systematic review and meta-analysis. Breast Cancer Res Treat *179*, 301-313. 10.1007/s10549-019-05477-5.

118. Saito, S., Okabe, H., Watanabe, M., Ishimoto, T., Iwatsuki, M., Baba, Y., Tanaka, Y., Kurashige, J., Miyamoto, Y., and Baba, H. (2013). CD44v6 expression is related to mesenchymal phenotype and poor prognosis in patients with colorectal cancer. Oncol Rep *29*, 1570-1578. 10.3892/or.2013.2273.

119. Liu, S., Cong, Y., Wang, D., Sun, Y., Deng, L., Liu, Y., Martin-Trevino, R., Shang, L., McDermott, S.P., Landis, M.D., et al. (2014). Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. Stem Cell Reports *2*, 78-91. 10.1016/j.stemcr.2013.11.009.

120. Park, W.S., Chung, K.W., Young, M.S., Kim, S.K., Lee, Y.J., and Lee, E.K. (2013). Differential protein expression of lymph node metastases of papillary thyroid carcinoma harboring the BRAF mutation. Anticancer Res *33*, 4357-4364.

121. Wang, Q., Zhu, G., Lin, C., Lin, P., Chen, H., He, R., Huang, Y., Yang, S., and Ye, J. (2021). Vimentin affects colorectal cancer proliferation, invasion, and migration via regulated by activator protein 1. J Cell Physiol *236*, 7591-7604. 10.1002/jcp.30402.
122. Sampson, C., Wang, Q., Otkur, W., Zhao, H., Lu, Y., Liu, X., and Piao, H.L. (2023). The roles of E3 ubiquitin ligases in cancer progression and targeted therapy. Clin Transl Med *13*, e1204. 10.1002/ctm2.1204.
123. Mani, A., and Gelmann, E.P. (2005). The ubiquitin-proteasome pathway and its role in cancer. J Clin Oncol *23*, 4776-4789. 10.1200/JCO.2005.05.081.
124. Park, H.B., Kim, J.W., and Baek, K.H. (2020). Regulation of Wnt Signaling through Ubiquitination and Deubiquitination in Cancers. Int J Mol Sci *21*. 10.3390/ijms21113904.
125. Hanson, A.J., Wallace, H.A., Freeman, T.J., Beauchamp, R.D., Lee, L.A., and Lee, E. (2012). XIAP monoubiquitylates Groucho/TLE to promote canonical Wnt signaling. Mol Cell *45*, 619-628. 10.1016/j.molcel.2011.12.032.
126. Ng, V.H., Hang, B.I., Sawyer, L.M., Neitzel, L.R., Crispi, E.E., Rose, K.L., Popay, T.M., Zhong, A., Lee, L.A., Tansey, W.P., et al. (2018). Phosphorylation of XIAP at threonine 180 controls its activity in Wnt signaling. J Cell Sci *131*. 10.1242/jcs.210575.
127. Lee, K.K., Rajagopalan, D., Bhatia, S.S., Tirado-Magallanes, R., Chng, W.J., and Jha, S. (2021). The oncogenic E3 ligase TRIP12 suppresses epithelial-mesenchymal transition (EMT) and mesenchymal traits through ZEB1/2. Cell Death Discov *7*, 95. 10.1038/s41420-021-00479-z.
128. Gatti, M., Imhof, R., Huang, Q., Baudis, M., and Altmeyer, M. (2020). The Ubiquitin Ligase TRIP12 Limits PARP1 Trapping and Constrains PARP Inhibitor Efficiency. Cell Rep *32*, 107985. 10.1016/j.celrep.2020.107985.
129. Brunet, M., Vargas, C., Larrieu, D., Torrisani, J., and Dufresne, M. (2020). E3 Ubiquitin Ligase TRIP12: Regulation, Structure, and Physiopathological Functions. Int J Mol Sci *21*. 10.3390/ijms21228515.
130. andrew_aguirre@dfci.harvard.edu, C.G.A.R.N.E.a., and Network, C.G.A.R. (2017). Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell *32*, 185-203.e113. 10.1016/j.ccell.2017.07.007.
131. Silvestrini, V.C., Thomé, C.H., Albuquerque, D., de Souza Palma, C., Ferreira, G.A., Lanfredi, G.P., Masson, A.P., Delsin, L.E.A., Ferreira, F.U., de Souza, F.C., et al. (2020). Proteomics analysis reveals the role of ubiquitin specific protease (USP47) in Epithelial to Mesenchymal Transition (EMT) induced by TGFβ2 in breast cells. J Proteomics *219*, 103734. 10.1016/j.jprot.2020.103734.
132. Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., Stuart, J.M., and Network, C.G.A.R. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet *45*, 1113-1120. 10.1038/ng.2764.
133. Tang, Z., Li, C., Kang, B., Gao, G., and Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res *45*, W98-W102. 10.1093/nar/gkx247.
134. Shi, J., Liu, Y., Xu, X., Zhang, W., Yu, T., Jia, J., and Liu, C. (2015). Deubiquitinase USP47/UBP64E Regulates β-Catenin Ubiquitination and Degradation and Plays a Positive Role in Wnt Signaling. Mol Cell Biol *35*, 3301-3311. 10.1128/MCB.00373-15.
135. Pan, K., Fu, J., and Xu, W. (2021). Role of Ubiquitin-Specific Peptidase 47 in Cancers and Other Diseases. Front Cell Dev Biol *9*, 726632. 10.3389/fcell.2021.726632.
136. Wu, J., Jiao, Y., Dal Molin, M., Maitra, A., de Wilde, R.F., Wood, L.D., Eshleman, J.R., Goggins, M.G., Wolfgang, C.L., Canto, M.I., et al. (2011). Whole-exome sequencing of neoplastic cysts of the pancreas reveals recurrent mutations in components of ubiquitin-dependent pathways. Proc Natl Acad Sci U S A *108*, 21188-21193. 10.1073/pnas.1118046108.

137.   Gopal, R.K., Kübler, K., Calvo, S.E., Polak, P., Livitz, D., Rosebrock, D., Sadow, P.M., Campbell, B., Donovan, S.E., Amin, S., et al. (2018). Widespread Chromosomal Losses and Mitochondrial DNA Alterations as Genetic Drivers in Hürthle Cell Carcinoma. Cancer Cell *34*, 242-255.e245. 10.1016/j.ccell.2018.06.013.

138.   Yoo, S.K., Song, Y.S., Lee, E.K., Hwang, J., Kim, H.H., Jung, G., Kim, Y.A., Kim, S.J., Cho, S.W., Won, J.K., et al. (2019). Integrative analysis of genomic and transcriptomic characteristics associated with progression of aggressive thyroid cancer. Nat Commun *10*, 2764. 10.1038/s41467-019-10680-5.