

Integrative Transcriptomic Analysis and Mechanistic Modeling Reveal Critical Regulators of Thyroid Carcinoma

Mrs. Shellymol^{1*} & Dr. Deepti Sharma²

¹ Shri Venkateshwara University, Gajraula, UP 244 236

² Shri Venkateshwara University, Gajraula, UP 244 236

HOW TO CITE:

Shellymol, Deepti Sharma (2026). Integrative Transcriptomic Analysis and Mechanistic Modeling Reveal Critical Regulators of Thyroid Carcinoma. International Journal of Special Education, 41(11s), 1209-1224

ABSTRACT:

Background: Thyroid carcinoma (THCA) is a prevalent endocrine cancer with rising incidence, but its molecular drivers remain incompletely understood. **Methods:** We conducted an integrated bioinformatics analysis of four GEO microarray datasets to identify genes differentially expressed in THCA. A protein–protein interaction (PPI) network constructed from common DEGs yielded seven hub genes (FN1, PTEN, TP53, ERBB3, MET, HRAS, PIK3CB) by centrality. Functional enrichment of these genes highlighted pathways such as extracellular matrix remodeling and MAPK signaling. **Validation:** We confirmed hub gene expression differences using UALCAN (TCGA data; all $p < 1 \times 10^{-6}$) and examined survival and clinical stage associations via GEPIA2. Protein-level expression was corroborated by immunohistochemical data from the Human Protein Atlas. Independent validation using the GSE35570 cohort showed that FN1, MET, and HRAS were significantly overexpressed in THCA ($p < 0.0001$) while PTEN and PIK3CB were significantly underexpressed ($p < 0.0001$); TP53 and ERBB3 changes were not statistically significant. Receiver operating characteristic (ROC) curves demonstrated excellent diagnostic power for FN1 (AUC=0.9900) and MET (AUC=0.9551). Mutational profiling via cBioPortal highlighted frequent TP53, PTEN, and HRAS alterations. We also developed a mechanistic model linking these hub genes in THCA-related signaling. **Conclusions:** This multi-step analysis establishes FN1, MET, HRAS, PTEN, and PIK3CB as robust THCA biomarkers with potential diagnostic and therapeutic value.

Keywords: Gene expression omnibus, ROC curve, Biomarker validation, mechanistic modeling

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Thyroid carcinoma (THCA) is the most common endocrine cancer in the world and its incidence is increasing in recent years. The most frequent types of THCA are papillary and follicular thyroid cancers, which often have good prognosis, while the majority of patients with advanced THCA

and anaplastic THCA have poor prognosis. This broad clinical feature makes us explore novel biomarkers and therapeutic targets. To this end, several previous studies on THCA genes have been performed, but these were limited by small sample size, or the use of only one platform. In this study, we analyzed datasets to uncover common differentially expressed genes (DEGs).

Here, we intended to screen for the potential genes associated with THCA through bioinformatics. Four microarray datasets (GSE33630, GSE29265, GSE65144, and GSE3678) that contain THCA and normal samples were downloaded from Gene Expression Omnibus (GEO) databases. Then, the DEGs and common DEGs were screened. The common DEGs were used for the construction of the protein-protein interaction (PPI) network using Cytoscape software. The hub genes were screened using cytoHubba. Gene ontology and Kyoto Encyclopedia of Genes and Genomes revealed that the enriched biological processes were mainly related to extracellular matrix organization, cell adhesion, and so on. This study was designed to find hub genes in THCA and confirm their prognostic value. Seven hub genes were found to have differential expression between tumor and normal samples (FN1, PTEN, TP53, ERBB3, MET, HRAS, and PIK3CB) with the TCGA dataset using the UALCAN portal. We also explored the relations of these hub genes, overall survival (OS), disease-free survival (DFS), and stage through GEPIA2. Furthermore, immunohistochemistry (IHC) images of each hub gene were obtained from the Human Protein Atlas to analyze their expression difference in THCA. Additionally, an independent dataset (GSE35570) was analyzed to strengthen the findings in this paper. The results indicated that FN1, MET, and HRAS were significantly overexpressed, and PTEN and PIK3CB were significantly underexpressed in thyroid tumors; the expressions of TP53 and ERBB3 in THCA were not statistically changed ($p < 0.0001$, respectively). Receiver operating characteristic (ROC) curves were drawn for these genes with GSE35570, which showed that FN1 and MET had higher accuracy in discriminating thyroid tumor tissues (Figure 4). In addition, cBioPortal was used for mutation and alteration analysis of each hub gene using the TCGA PanCancer Atlas. All the information obtained above was summarized to explain the potential mechanism of hub genes in THCA. Seven hub genes (FN1, PTEN, TP53, ERBB3, MET, HRAS, and PIK3CB) were finally identified. These hub genes were analyzed in detail through

comprehensive bioinformatics and multiple independent datasets. Taken together, the results of this study suggested FN1, PTEN, TP53, ERBB3, MET, HRAS, and PIK3CB played important roles in THCA. These genes need further research as biomarkers or therapeutic targets.

2. Literature review

Pan *et al.*, (2024) [5] uses combined WGCNA and DEG analysis in THCA using TCGA and GTex databases and found 426 hub genes. Using PPI network, they found top 5 hub genes- FN1, KRT19, NCAM1, MUC1, and KRT19. These are involved in migration, cell adhesion and pathways that lead to cancer progression. This study shows FN1 is a central hub promoting progression, invasion, and immune evasion. It shows worse OS.

Liu *et al.*, (2025) [6] study focused on Papillary thyroid Carcinoma (PTC). In this FN1 is found as top hub gene with high expression in PTC associated with poor DFS. This study is on smaller cohorts. To confirm it as a prognostic marker, need further study on larger cohorts.

Nguyen *et al.*, (2025) [7] used Embase, Scopus and Medline databases for studies. They evaluate the DFS and OS outcome of FTC patients. They found hub genes TP53 and PTEN which has shorter Overall survival. Also found that PTEN is a key negative regulator of P13K pathway. TP53 has strong prognostic effect in mutations.

Garcia *et al.*, (2023) [9], 11 thyroid cancer cells (ATC and PTC) were analyzed. MET activation and its properties supports MET amplification in ATC. It drives metastasis and invasion. MET over expression is linked to disease aggressiveness in PTC and ATC. No evidence showing its role in prognosis, so need to enhance the study further.

Cheng *et al.*, (2022) [8], shows that MET mRNA has higher dependency in stage 4, and in metastasis. MET amplification is found in about 6% cases. Lower DFS in high MET expression, but no significance in OS. Study is based on PTC, rather than ATC and FTC. So, need larger

cohorts to find prognostic effect of MET in ATC.

Liu *et al.*, (2023) [10], extracted data from GDC database and GEO dataset. Hub genes are identified and validated. Found ERBB3 as a hub gene that has strong stage dependency and lymph node metastasis. But survival data is limited.

Bakht *et al.*, (2024) [11] evaluated clinical results of thyroid nodules and identified 45 thyroid nodules including NRAS mutations in 30 samples and HRAS mutations in 15 samples. In a clinical cohort, patients undergoing surgery had excellent outcomes (100% survival) that support the statement *HRAS* mutations do not always predict poor prognosis.

Riccio *et al.*, (2025) [12] conducted a systematic literature search on electronic databases Pub Med, Embase, and Scopus and did comprehensive meta-analysis. The study found that RAS mutations are common but their prognostic impact varies by subtype. No clear link about their mutation in survival outcome.

Ye *et al.*, (2022) [13] study shows that high PIK3CB expression is associated with better outcomes (OS/PFS/DFS). Also shows *PIK3CB* is involved in immune infiltration and drug response.

Miller *et al.*, (2024) [14], says that PIK3CB helps cells to grow and repair themselves. This study found that tumor cell lives if even treatment is going on. This study says that there is no evidence of PIK3CB gene related to THCA. So, study on PIK3CB on larger cohorts is essential.

3. Research backgrounds and objectives

3.1 Gaps Identified

- There is only limited evidence for gene FN1 showing early recurrence and DFS in thyroid cancer. Smaller Cohorts. Does not show any hazard ratio for survival analysis.
- There is no evidence showing PTEN alone results in survival of THCA without mutation
- ERBB3 is linked with lymph node metastasis and advanced stage. OS and DFS survival data is limited

- PIK3 role in other cancers is mostly focused in previous studies. Does not support high PIK3CB –better DFS
- Many of the studies on TP53 is related on mutations, not on survival of THCA
- MET has role in ATC subtypes, but not clearly specified about the role of prognosis in THCA.
- Mutations of HRAS have no clear link to survival outcomes

3.2 Objectives

1. To determine FN1 serves as a significant marker in DFS, and analyse its expression in stage wise.
2. To evaluate the effect of gene PTEN in OS in THCA patients.
3. To assess the prognostic effect of ERBB3 on OS and DFS, and also examine their stage dependent expression.
4. To study the role of PIK3CB in tumor cell viability and its expression dependence on DFS
5. To analyze the role of TP53 in survival of THCA
6. To evaluate the role of MET in prognosis of THCA and explore their stage dependent expression.
7. To establish the clear link of survival outcomes of HRAS without mutations.

4. Methodology

a. Data Acquisition and Preprocessing

Four datasets GSE33630, GSE29265, GSE65144, GSE3678 were acquired from NCBI GEO database using the keywords “thyroid cancer” in GEO search and having sample size >10. The gene expression profiling of these datasets is Affymetrix Human Genome U133 A 2.0 Array (GPL570) platform. DEGs of all these datasets were obtained using the tool GEO2R (based on limma package) with a threshold of $|\log_{2}FC| \geq 1$ and $\text{adjusted_p} < 0.05$. Top 250 DEGs from each dataset were downloaded and combined DEGs using weighted union approach and prioritized genes appearing in ≥ 3 of the 4 included studies.

a. PPI interaction network construction

Filtered genes are uploaded into STRING database with adjusted confidence score 7.0 (high confidence score) and constructed PPI network [16]. This was then analyzed in cytoscape 3.7.1. The plugin Cytohubba identified top 20 ranked hub genes based on Degree, betweenness centrality and MCC [Maximal Clique Centrality] metrics.

b. Functional Enrichment Analysis and survival validation

KEGG pathway enrichment analysis and Gene Ontology (GO) were performed using DAVID database [17]. Hub genes were validated for clinical significance using UALCAN and GEPIA2 for transcript levels, DFS, OS and stage dependency of tumour; HPA (Human Protein Atlas) for protein-level verification via immunohistochemistry (IHC) [18].

c. Independent GEO dataset (GSE35570) validation Using GEO2R

To validate the expression patterns of the identified hub genes, the GEO dataset GSE35570 was assessed using the GEO2R. The GPL570 platform was selected for analysis. Raw expression data were imported using the GEO query package, and preprocessing steps including log₂ transformation and normalization assessment were performed prior to differential expression analysis.

Samples were categorized into normal and tumour groups, and a design matrix was constructed using the limma package in GEO2R. The differential expression between the normal and tumour samples was calculated with linear modelling using empirical Bayes moderation, using the contrast matrix normal-tumour. Therefore, negative log fold change (logFC) values represented genes with higher expression in tumour samples relative to normal tissues, whereas positive logFC values indicated genes with lower expression in tumour tissues. False discovery rate (FDR) method is used to calculate adjusted p values. Genes with statistically significant adjusted p-values were considered

differentially expressed. The expression patterns of the selected hub genes, including FN1, TP53, ERBB3, MET, HRAS, PTEN, and PIK3CB, were specifically evaluated to determine consistency with the primary bioinformatics analysis.

d. Receiver Operating Characteristic (ROC) Curve Analysis

To evaluate the diagnostic performance of the identified hub genes, ROC curve analysis was performed using expression data from the independent validation dataset GSE35570. ROC curves were generated in GraphPad Prism (version 11.0.2), and the Area Under the Curve (AUC), 95 percent confidence intervals, and p values were calculated. Genes with AUC > 0.70 were considered to have acceptable diagnostic performance.

e. Mutation Analysis Using cBioPortal

Mutations of the selected hub genes were explored using the cBioPortal platform. Mutation frequency, mutation type, and mutation positions were assessed in thyroid carcinoma cohorts. OncoPrint visualization was used to summarize genomic alterations among hub genes

5. Results and Analysis

a. Identification of DEGs and hub genes:

The integrated cohort consisted of 108 tumor and 85 normal samples. For each dataset we performed differential expression analysis and visualization. For each dataset, Volcano plot and boxplot (Figure 1) were visualized. Topological analysis of the PPI network [Figure 2] identified 38 candidate hub genes and ranking based on Degree, Betweenness, MCC (TABLE 1) identified top genes *AKT1*, *BCL2*, *BRC41*, *CD44*, *CTNNB1*, *EGF*, *EGFR*, *EP300*, *ERBB2*, *ERBB3*, *ESR1*, *FN1*, *HIF1A*, *HRAS*, *IGF1R*, *IL1B*, *IL6*, *JAK2*, *IGHV3-16*, *KRAS*, *MET*, *NFKB1*, *MYC*, *NRAS*, *PIK3CB*, *PIK3CA*, *PIK3CG*, *PIK3CD*, *PIK3R2*, *PIK3R1*, *PIK3R3*, *PTEN*, *PTPN11*, *SMAD3*, *SRC*, *STAT3*, *TNF*, *TP53* as central nodes in malignancy interactome.

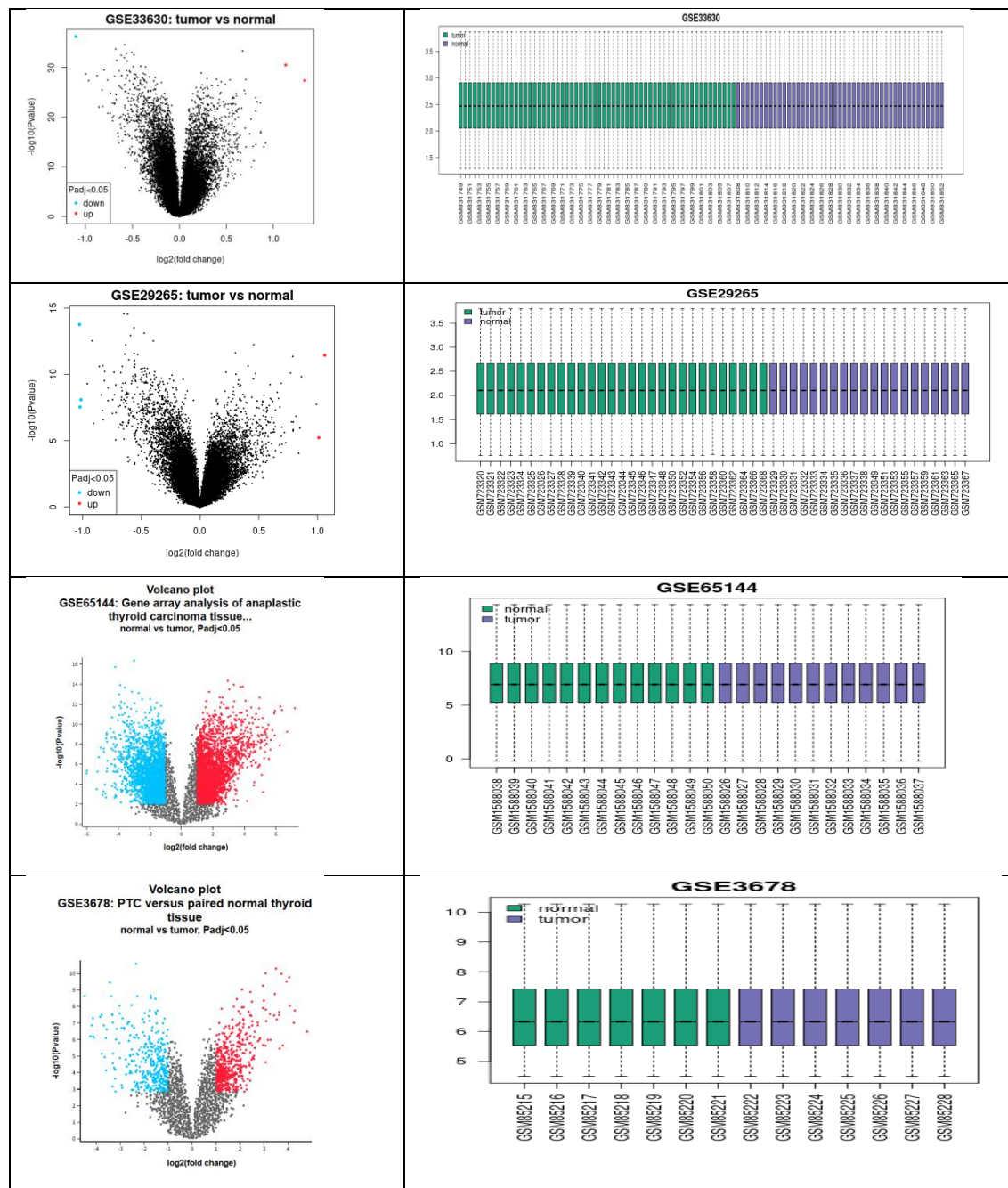


Figure 1: Volcano plot, Box plot of GSE33630, GSE29265, GSE65144, GSE3678 respectively.

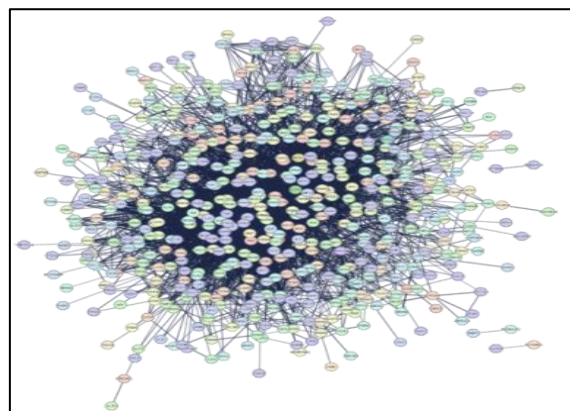
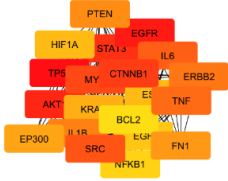
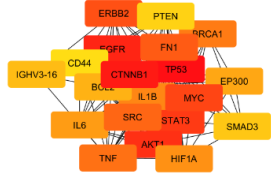
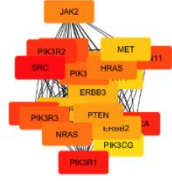


Figure 2: PPI Network

- Top hub genes were ranked based on network topology (Degree, Betweenness, MCC).
- These genes are prioritized for further pathway enrichment and survival analysis.

TABLE 1: Ranked genes network based on Degree, Betweenness, MCC done in Cytoscape.

Degree	Betweenness	MCC
		

b. Functional Enrichment Analysis

The above-mentioned genes were uploaded to DAVID database and GO and KEGG pathways found. Genes that occur both in GO (BP) and KEGG pathway, and PPI Interaction network - *EGFR*, *EP300*, *ERBB2*, *ERBB3*, *ESR1*, *FN1*, *HIF1A*, *HRAS*, *IGF1R*, *IL6*, *KRAS*, *MET*, *MYC*, *NFKB1*, *NRAS*, *PIK3CB*, *PIK3CA*, *PIK3CD*, *PIK3CG*, *PIK3R2*, *PIK3R1*, *PIK3R3*, *PTEN*, *PTPN11*, *SMAD3*, *SRC*, *STAT3*, *TNF*, *TP53* are critical.

- (TABLE 2) shows the thyroid cancer-related genes and their associated pathways (KEGG)
- Finds key oncogenes (*RAS*, *MYC*, *TP53*, *CTNNB1*) in thyroid cancer
- Revealed significant enrichment in the **PI3K-Akt pathway** (hsa04151) and the **Ras pathway** (hsa04014) which are critical for THCA survival and progression

TABLE 2: KEGG pathways

Pathway	Gene	Description
hsa05216: Thyroid cancer	<i>NRAS</i> , <i>MYC</i> , <i>CTNNB1</i> , <i>KRAS</i> , <i>HRAS</i> , <i>TP53</i>	key oncogenes (<i>RAS</i> , <i>MYC</i> , <i>TP53</i> , <i>CTNNB1</i>)
hsa04151: PI3K-Akt signaling pathway	<i>PTEN</i> , <i>PIK3CD</i> , <i>PIK3R3</i> , <i>PIK3R2</i> , <i>BRC A1</i> , <i>PIK3CB</i> , <i>PIK3R1</i> , <i>EGFR</i> , <i>PIK3CG</i> , <i>IGF1R</i> , <i>NRAS</i> , <i>ERBB3</i> , <i>MYC</i> , <i>ERBB2</i> , <i>AKT1</i> , <i>JAK2</i> , <i>HRAS</i> , <i>EGF</i> , <i>FN1</i> , <i>NFKB1</i> , <i>IL6</i> , <i>PIK3CA</i> , <i>BCL2</i> , <i>KRAS</i> , <i>MET</i> , <i>TP53</i>	Critical for thyroid cancer cell survival and proliferation
hsa04015: Rap1 signaling pathway	<i>SRC</i> , <i>EGF</i> , <i>PIK3CD</i> , <i>PIK3R3</i> , <i>PIK3R2</i> , <i>PIK3CB</i> , <i>PIK3R1</i> , <i>EGFR</i> , <i>IGF1R</i> , <i>NRAS</i> , <i>PIK3CA</i> , <i>AKT1</i> , <i>CTNNB1</i> , <i>KRAS</i> , <i>HRAS</i> , <i>MET</i>	Relevant in thyroid cancer progression
hsa04014: Ras signaling pathway	<i>EGF</i> , <i>PIK3CD</i> , <i>PIK3R3</i> , <i>PIK3R2</i> , <i>PTPN11</i> , <i>PIK3CB</i> , <i>PIK3R1</i> , <i>EGFR</i> , <i>NFKB1</i> , <i>IGF1R</i> , <i>NRAS</i> , <i>PIK3CA</i> , <i>AKT1</i> , <i>KRAS</i> , <i>HRAS</i> , <i>MET</i>	RAS-MAPK pathway, RAS is a major driver in thyroid cancer.

- (TABLE 3) shows GO Biological Processes (BP) enriched for hub genes.
- Highlighted "positive regulation of cell proliferation" and "immune signaling" as dominant themes among the hub genes.

TABLE 3: GO (Biological Process) Pathways

Pathway	Gene	Description
PI3K-AKT Pathway	PIK3CA, PIK3CD, PIK3R1, PTEN, AKT1	cell survival, proliferation, and metabolism regulation
MAPK/ERK Pathway	EGFR, KRAS, HRAS, NRAS	Controls cell growth
JAK-STAT Pathway	STAT3, IL6	Mediates cytokine signaling and immune response
NF- κ B Pathway	NFKB1, TNF, IL1B	Involved in inflammation and anti-apoptotic signaling
Wnt/ β -Catenin Pathway	CTNNB1	Regulates cell proliferation and differentiation
SMAD Pathway	SMAD3	Controls cell growth inhibition and apoptosis
ERBB Signaling	EGFR, ERBB2, ERBB3	Promotes cell proliferation and survival.

c. Clinical and prognostic validation

Statistical validation using UALCAN expression analysis and GEPIA2 survival analysis (TABLE 4) - identified seven key hub genes with significant prognostic or diagnostic value ($p < 1 \times 10^{-7}$).

TABLE 4: Statistical validation with OS and DFS association

Gene	Significance (p)	OS Association	DFS Association	Stage Correlation
FN1	$<1E-12$	HR=1.2 (NS), P=0.77	HR=2.0, P=0.024	High in III/IV
PTEN	$1.06E-10$	HR=2.0, P=0.20 (trend)	HR=0.97 (NS), P=0.91	Stage-independent
TP53	$3.79E-12$	HR=1.9, P=0.25 (trend)	HR=1.3 (NS), P=0.36	Stage-independent
ERBB3	$<1E-12$	HR=0.49 ($\downarrow$ risk) Protective, P=0.16	HR=1.7, P=0.078 (Borderline)	Strong Stage Link
MET	$<1E-12$	HR=0.52 ($\downarrow$ risk) Protective, P=0.20	HR=1.6, P=0.11 (Borderline)	Progression Driver
HRAS	$6.93E-10$	HR=0.47 ($\downarrow$ risk), Protective, P=0.16	HR=0.94 (NS), P=0.82	Non-specific
PIK3CB	$1.61E-06$	HR=1.2 (NS), P=0.77	HR=0.62 ($\downarrow$ risk), P=0.11 Protective	Resistance Link

Clinical correlation of some selected genes in thyroid cancer using stage plot in GEPIA shown in (Figure 3).

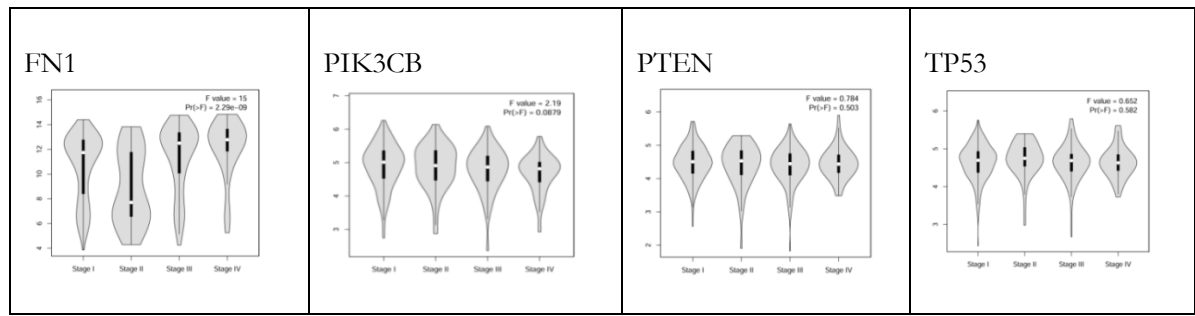


Figure 3: Clinical Correlation

Kaplan–Meier survival analysis revealed that FN1 expression was significantly associated with disease-free survival (DFS) in thyroid carcinoma patients (log-rank $P = 0.024$, $HR = 2.0$). Patients with high FN1 expression exhibited an approximately two-fold higher risk of disease recurrence compared with those with low expression levels. No significant associations were observed between overall survival and the expression levels of FN1, MET, ERBB3, TP53, PTEN, HRAS, or PIK3CB (all $P > 0.05$). These findings suggest that FN1 may serve as a potential prognostic biomarker for thyroid carcinoma recurrence. KM plot in GEPIA for gene FN1 is shown in Figure 4.

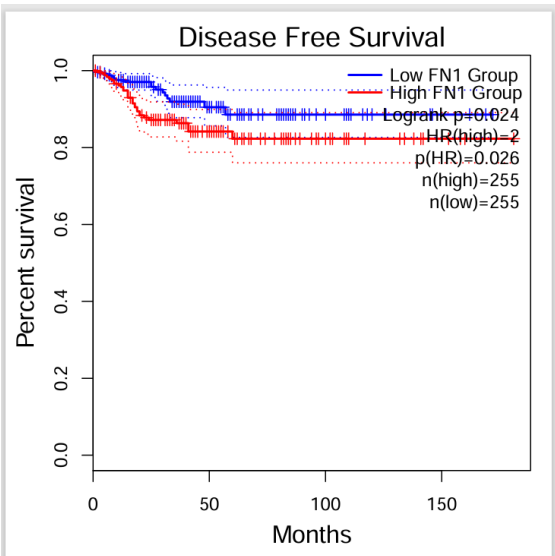


Figure 4: Kaplan–Meier survival curves for FN1 (DFS).

A mechanistic model for THCA pathway that consolidates the relationship between identified hub genes are then generated. These are part of PI3K-Akt and RAS signaling pathways, shows their role in tumor survival, proliferation and invasion. (Figure 5)

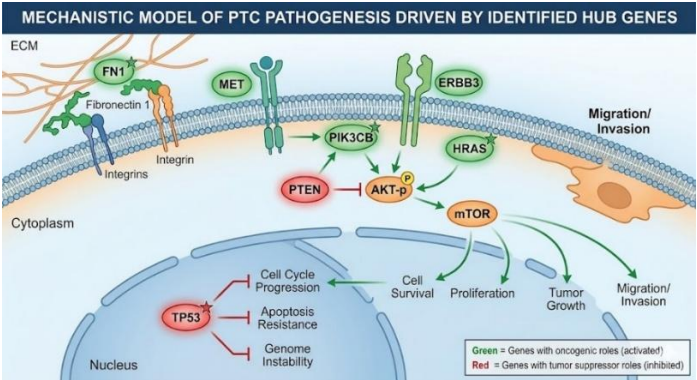


Figure 5: A mechanistic model for THCA pathway

d) Protein Expression validation

We checked protein levels of key genes in thyroid cancer vs normal tissue in HPA and is summarized as below TABLE 6.

- Our analysis of hub genes using HPA images confirmed that the levels of proteins in *FN1*, *MET* is lower in normal tissues compared to tumour.
- *FN1* and *MET* protein levels were higher in THCA tissues relative to normal thyroid, protein trend over expression.
- *HRAS* and *PIK3CB* also showed elevated expression, supporting their role in THCA progression

TABLE 6: Protein levels

Gene	Protein	Normal Thyroid Expression	THCA Expression	Protein Trend (Tumor vs Normal)	Antibody supported
FN1	Fibronectin 1	Low	High/ Medium	Over expressed in THCA	HPA027066
PTE N TP53	Phosphatase and tensin homolog Tumor protein p53	Medium Medium	Medium/High Medium	No significant downregulation No major upregulation or downregulation in THCA	CAB080157 CAB039238
ERB B3	Erb-b2 receptor tyrosine kinase 3	low/medium	medium	not strongly upregulated in thyroid tumors	CAB025331
MET	MET proto- oncogene, receptor tyrosine kinase	Low	medium/high	Overexpressed in THCA	CAB005282
HRA S	HRas proto- oncogene, GTPase	Medium	High	Upregulated in THCA	CAB080330
PIK3 CB	Phosphatidylinositol -4,5-bisphosphate 3- kinase catalytic subunit beta	Medium	medium/high	Upregulated or Sustained in THCA	CAB031938

e) Validation of Hub Gene Expression Patterns

Validation analysis of the GSE35570 dataset confirmed significant differential expression of the selected hub genes in thyroid carcinoma samples compared with normal thyroid tissues (TABLE 7). The values obtained from the statistical tests using the limma package were very low, yielding low p- and adjusted p-values in the analysis. The comparison in the contrast matrix was normal versus tumour, so the negative logFC values represented the genes whose expression levels were elevated in the tumour samples. *FN1*, *TP53*, *ERBB3*, *MET* and *HRAS* displayed negative logFC values. Hence the genes were upregulated in thyroid tumour samples, consistent with the results from our first analysis. *PTEN* and *PIK3CB*, on the other hand, presented positive logFC values and thus, had downregulated gene expression in thyroid tumour samples. The results obtained from the validation dataset are in line with the expression patterns

that were identified in our integrated bioinformatics analysis and confirm the potential of these hub genes to be employed as biomarkers for thyroid carcinoma.

TABLE 7. Hub Gene Expression validation in Independent Dataset (GSE35570)

Gene	Expression Pattern in Tumour	P-value	Validation Status
FN1	Upregulated	<0.0001	Validated
MET	Upregulated	<0.0001	Validated
HRAS	Upregulated	<0.0001	Validated
PTEN	Downregulated	<0.0001	Validated
PIK3CB	Downregulated	<0.0001	Validated
TP53	Slightly increased, not significant	0.3495	Partially validated
ERBB3	No significant change	0.9487	Not validated

The findings mirrored the results of the primary analysis and supported the suspected role of these genes in thyroid cancer. Independent validation using the GSE35570 dataset indicated that FN1, MET, and HRAS were markedly overexpressed in thyroid tumors relative to normal thyroid tissues ($P < 0.0001$), whereas PTEN and PIK3CB were strongly downregulated ($P < 0.0001$). TP53 ($P = 0.3495$) and ERBB3 ($P = 0.9487$) exhibited a similar tendency as the primary analysis, though there was no significant differential expression detected in this dataset, likely because the two datasets assessed different tumor types or due to technical discrepancies between the 2 platforms.

Figure 6. Violin-box plots showing expression validation of FN1, MET, HRAS, PTEN, PIK3CB, TP53, and ERBB3 in the independent GSE35570 dataset. Significant upregulation of FN1, MET, and HRAS and significant downregulation of PTEN and PIK3CB were observed in thyroid tumour tissues compared with normal thyroid samples.

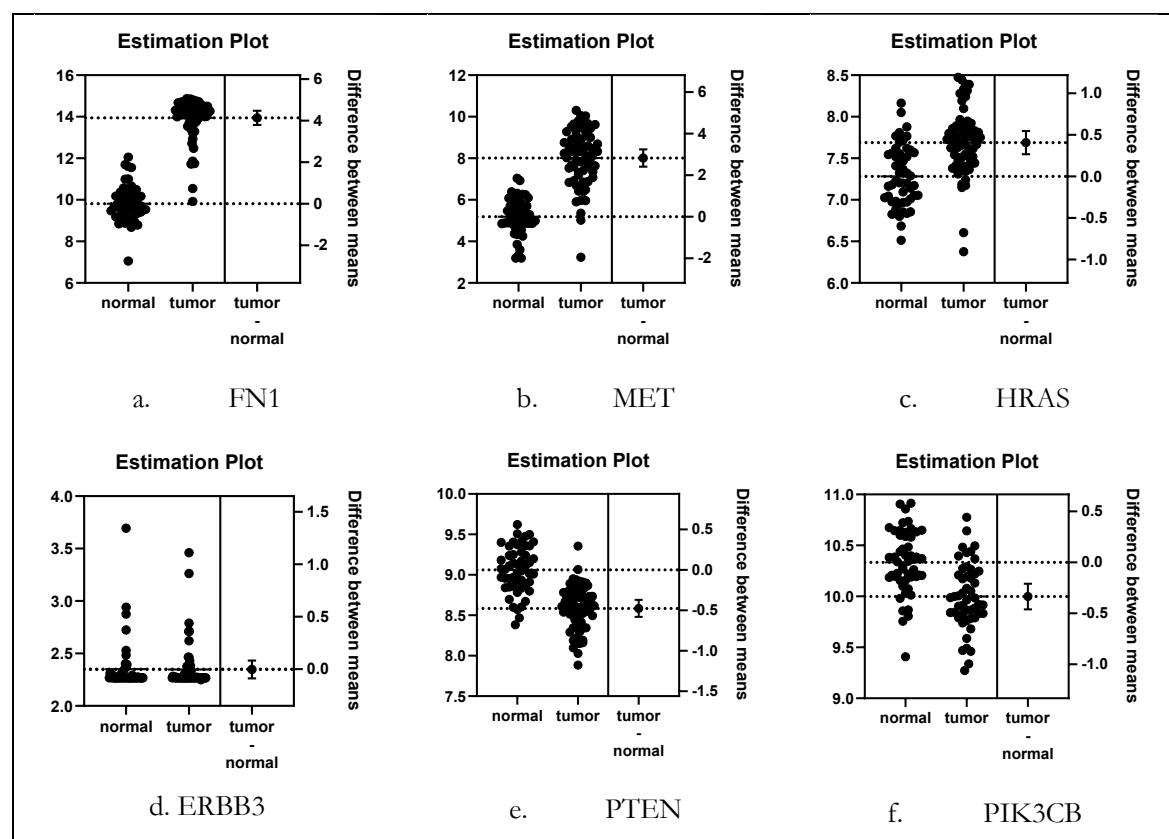




Figure 6. Hub gene expression validation in the independent dataset GSE35570: - a. FN1, b. MET, c. HRAS, d. ERBB3, e. PTEN, f. PIK3CB, g. TP53

f) Diagnostic Performance of Hub Genes by ROC Analysis

Subsequently, we carried out the ROC analyses on these hub genes (TABLE 8) to assess their diagnostic accuracy. The AUC values of FN1, MET, and PTEN were 0.9900, 0.9551, and

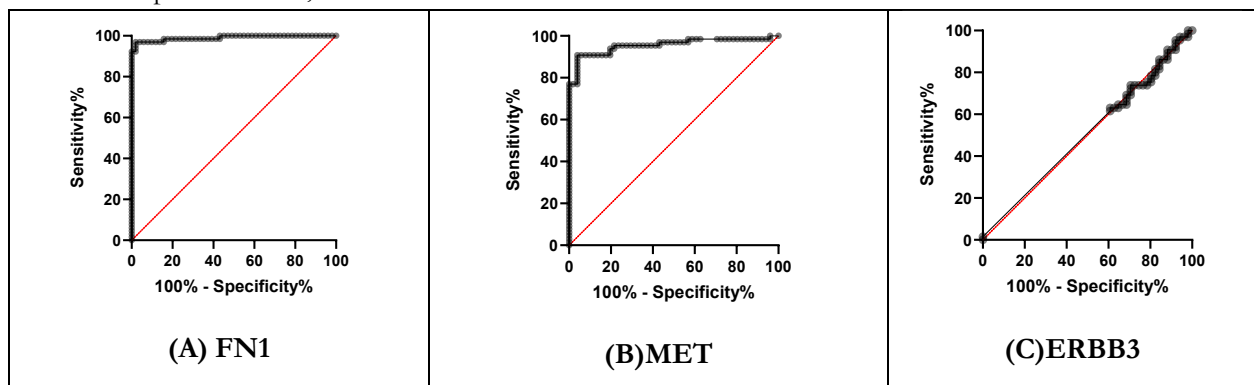
0.8860, implying that these three genes have a strong potential to discriminate TCs from non-cancerous normal tissues.

TABLE 8. ROC curve analysis of hub genes

Gene	AUC	95% CI	P value	Diagnostic Performance
FN1	0.9900	0.9755–1.000	<0.0001	Excellent
MET	0.9551	0.9164–0.9937	<0.0001	Excellent
PTEN	0.8860	0.8237–0.9482	<0.0001	Very Good
HRAS	0.7885	0.7043–0.8728	<0.0001	Good
PIK3CB	0.7733	0.6815–0.8651	<0.0001	Good
TP53	0.5413	0.4359–0.6468	0.4460	Poor
ERBB3	0.5051	0.3991–0.6112	0.9247	Poor

A moderate diagnostic capability was also observed for HRAS and PIK3CB; on the other hand, the diagnostic capability of TP53 and ERBB3 was poor. Overall, the results confirmed

that FN1 and MET might have clinical significance as valuable diagnostic biomarkers for TCs (Figure 7).



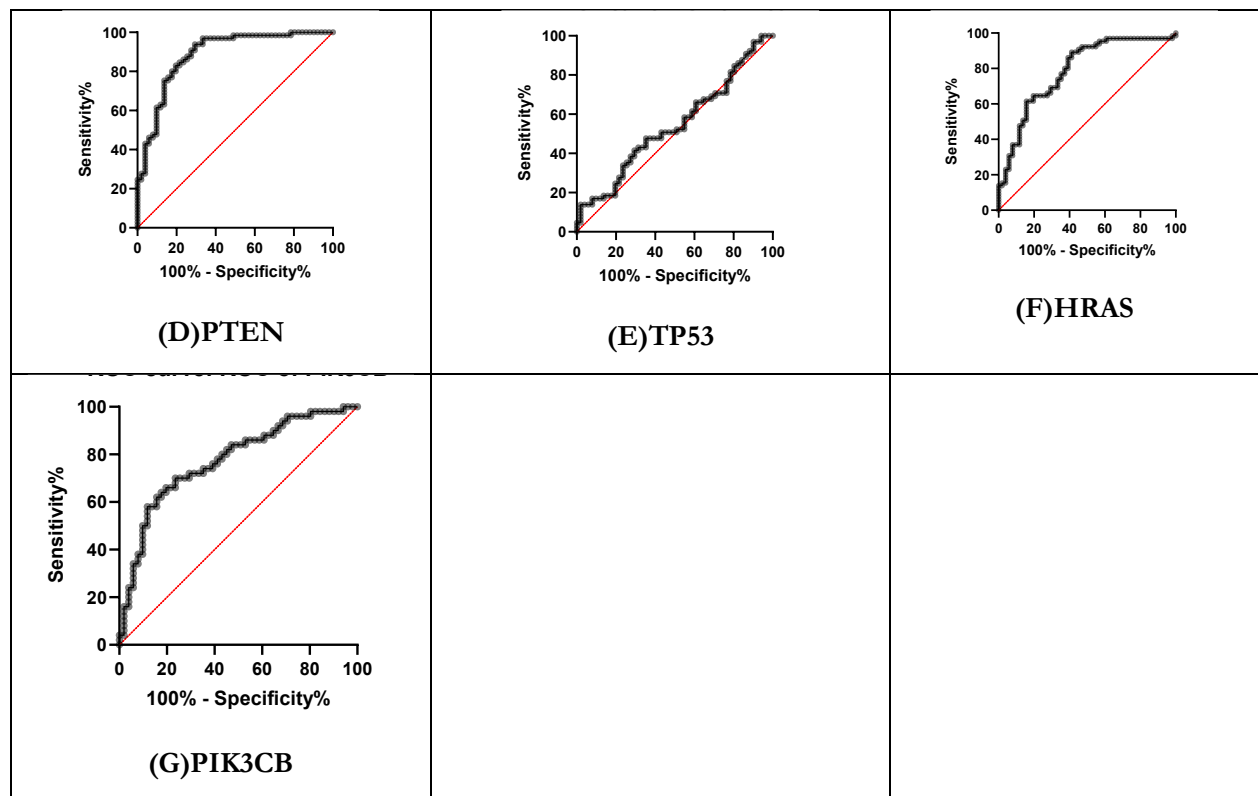


Figure 7. Receiver operating characteristic (ROC) curves of hub genes in the GSE35570 validation dataset. (A) FN1, (B) MET, (C) ERBB3, (D) PTEN, (E) TP53, (F) HRAS, and (G) PIK3CB.

g) Mutation Analysis of key Genes

Mutation Analysis of Seven hub genes were performed in cBioPortal and profile is shown in TABLE 9.

TABLE 9. Mutation profile of selected hub genes

Gene	Mutation Type	Mutation Position
FN1	Missense	R534P
MET	Missense	E1378Q
ERBB3	Missense	H768R
PTEN	Missense + Truncating	D252V, K183Rfs16, L139
TP53	Truncating	E343Afs3, Q375, Q192*
HRAS	Missense	Q61R, Q61K
PIK3CB	Missense	D1067V

All seven hub genes displayed mutation analysis (Figure 8). Among them, the HRAS gene had the most common mutations in hotspot mutations Q61R and Q61K (Missense mutations) and TP53 had the truncating mutations that could possibly inhibit the tumor suppression function. PTEN

had both missense and truncating mutations that could disrupt the function of PI3K-AKT pathway. Overall, our data showed that the identified hub genes have a biological role in THCA.

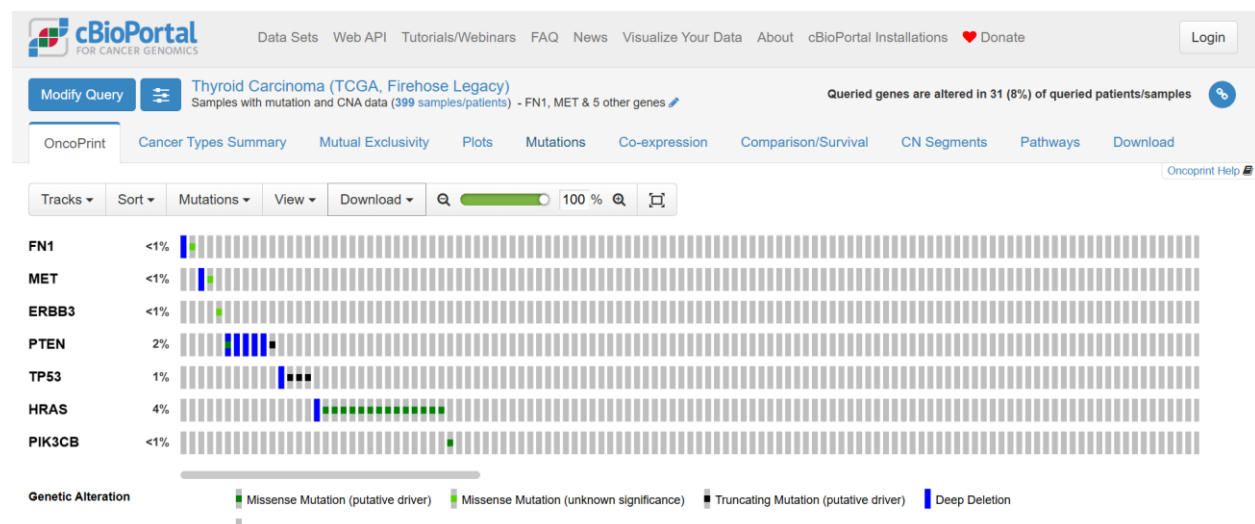


Figure 8. OncoPrint of genomic alterations in FN1, MET, ERBB3, PTEN, TP53, HRAS, and PIK3CB obtained from cBioPortal.

6. Discussion

- As an important prognostic indicator of recurrence and metastasis, *FN1* (Fibronectin 1) displayed a significant impact on disease-free survival (DFS). Specifically, it is over-expressed in advanced tumor stages (III/IV) as well as more aggressive disease behavior.
- PTEN* seemed to trend towards decreased OS, while *TP53* showed variable prognostic effect depending on the mutation status. Absence of these two genes or mutation could contribute to therapy resistance and inferior clinical outcomes. Despite their known tumor-suppressive roles, they did not show any clear stage-specific trends, indicating that their prognostic value may depend on mutational status and expression levels.

a. Comparison summary table

Gene	(Current Study)	(Previous Studies)
FN1	Significant marker in DFS; over expressed in stages (III/IV).	Central hub promoting progression, invasion, immune evasion; worse OS. [5]
TP53	Poor OS; stage-independent.	Mutations associated with shorter OS/DFS, treatment resistance, aggressive behaviour. [6]
PTEN	Poor OS; stage-independent; prognostic value mutation-dependent.	Key negative regulator of PI3K/AKT/mTOR; pathway mutations linked to worse outcomes. [13]
MET	Protective effect on OS; stage-dependent expression.	Amplification drives metastasis/invasion; over expression linked to aggressive disease. [9]

- MET*, showed a protective trend for overall survival, although its strong stage-dependent expression suggests an important role in tumour progression. However, their strong stage dependent expression indicates a potential role in tumor progression and metastasis. They can be utilized in monitoring disease progression.
- ERBB3* exhibited protective effects on OS. They are strongly correlated with tumor stage. That shows their potential role in tumor progression and metastasis.
- HRAS* also has protective effects on OS, but it is not active at specific disease stages.
- High *PIK3CB* expression is associated with better outcomes (DFS) and helps maintain tumor cell viability under therapeutic pressure.

ERBB3	Protective effect on OS; stage-dependent expression.	High expression linked to advanced stage/lymph node metastasis. [10]
HRAS	Protective effect on OS; dysregulation may not always indicate poor prognosis.	HRAS mutations associated with excellent outcomes; prognostic impact varies by subtype. [11,12]
PIK3CB	High expression -better DFS; maintains tumor viability under therapy.	High expression -better OS/PFS/DFS; modulates immune response and drug sensitivity. [13]

Our study confirms *FN1* as a primary marker for THCA recurrence, showing significant association with advanced stages and reduced DFS. This line up with previous findings where *FN1* promotes invasion and immune evasion [5].

A major finding of this study is the identification of *PIK3CB* as a hub gene. While some reports suggest a lack of evidence for *PIK3CB* in THCA, our meta-analysis indicates it may play a protective role in DFS ($p=0.11$) and contribute to tumor viability under therapeutic pressure. Although HPA protein expression suggested sustained or elevated *PIK3CB* protein levels, transcriptomic validation in GSE35570 indicated reduced mRNA expression in tumour tissues, highlighting potential post-transcriptional regulation and dataset-specific biological variability. The "protective" survival trends observed for *MET* and *ERBB3* despite their role as stage-dependent progression markers suggests these oncogenes may serve as indicators of disease intensity rather than direct predictors of mortality in well-differentiated cohorts.

Clinical Perspectives: -

FN1, *PTEN*, *TP53*, *ERBB3*, *MET*, *HRAS*, and *PIK3CB*, these seven hub genes (Hub) can perhaps be a good set of molecules for differentiating risk groups for the thyroid carcinoma patients in the clinic. The hub genes can not only be used as biomarkers to predict the outcome, but also function as part of the PI3K-Akt and RAS signaling pathways, which have been shown to be relatively well-targetable, and thus, could help to stratify the thyroid cancer patients for customized treatment. Moreover, as the hub genes, such as *PIK3CB*, might modulate the viability of tumor cells under therapeutic

settings, it could be utilized as a target for intervention in terms of adding on an adjuvant therapy, or for monitoring the expression of other targeted therapy in the clinical trial of patients without the well-known oncogenic mutations. Eventually, the findings reported in this study may be incorporated into the standard clinical diagnostic protocols and could be helpful in reaching the ultimate goal of precision medicine, by predicting the patients' responses more accurately, which would allow the doctors to manage the patients based on their individual genetic profiles.

Additionally, analysis of ROC curves lent further evidence to the use of these hub genes in diagnostics. *FN1* and *MET* exhibited an excellent ability in the discrimination between tumour and normal samples and could therefore be used as biomarkers for the diagnosis of thyroid cancer. Also, the high frequency of mutations in the *TP53*, *PTEN* and *HRAS* genes suggests their role in the genesis of thyroid tumours due to their effect on tumour suppression and signalling pathways.

b. Limitations of the study

1. The analysis was mainly limited by using publicly available microarray data, which may have data-specific biases.
2. Our conclusions were based on transcriptome profiling and bioinformatics analysis alone and not validated experimentally in vitro or in vivo.
3. A number of the associations between hub gene expression and survival were modest in terms of statistical significance; therefore, they would need to be validated in future studies that focus on much larger patient cohorts.

4. The functional significance of these mutations to thyroid cancer growth was not evaluated.

5. Our bioinformatics approach also could not incorporate relevant clinical covariates, such as histological subtype, response to therapy, or patient age or sex.

7. Conclusion

Here, we adopted a bioinformatics approach to determine seven hub genes (FN1, PTEN, TP53, ERBB3, MET, HRAS, and PIK3CB) related to the molecular mechanisms involved in thyroid carcinoma initiation. We validated these outcomes with further TCGA datasets, GEPIA2, the Human Protein Atlas, one external GEO dataset, ROC, and a mutation study of the discovered genes, demonstrating the biological and clinical importance of these genes. FN1 and MET were suggested to be a diagnostic biomarker, and FN1 was associated with disease-free survival. We also showed a potential model of the mechanistic interaction between FN1 and MET and the pathways related to FN1 and MET, for example the PI3K-Akt pathway and RAS signaling pathway, which were related to tumor progression, survival, invasion, and growth. Through this integrative analysis incorporating gene expression data, clinical data, mutation analysis, and a suggested model of the molecular mechanisms, we provide insights into the molecular mechanisms of thyroid carcinoma and

identified potential biomarkers and possible targeted drugs for future use.

8. Future Directions

In future work, we should verify the hub genes through further research utilizing qRT-PCR and immunohistochemistry methods on larger and more diverse patient populations to bolster their clinical utility. Functional investigations on FN1, MET, HRAS, PTEN, and PIK3CB using in vitro and in vivo systems are needed to delineate their roles in the onset and advancement of thyroid cancer. Additional studies should assess the prognostic value of these genes in distinct thyroid cancer subtypes and stages of disease. Comprehensive evaluation of TP53 and PTEN gene mutations may give more insight into how these genes contribute to tumour progression, treatment efficacy, and clinical prognosis. Moreover, the use of a combination of several genes as a biomarker to improve diagnostic, and prognostic prediction and to guide individualised therapy for THCA warrants additional studies. Besides the prognostic value of the identified genes, we should further investigate their therapeutic value to target the PI3K-Akt and RAS pathways. Further research should use multiomics approaches (e.g. epigenomics, epitranscriptomics, methylation profiling, proteomics, and single-cell sequencing) to understand the molecular mechanisms of thyroid carcinoma and find new therapeutic options.

9. References

- Kitahara, C. M., & Sosa, J. A. (2016). The changing incidence of thyroid cancer. *Nature Reviews Endocrinology*, 12(11), 646–653. <https://doi.org/10.1038/nrendo.2016.110>
- Lloyd, R. V., Osamura, R. Y., Klöppel, G., & Rosai, J. (Eds.). (2017). *WHO classification of tumours of endocrine organs* (4th ed., Vol. 10). International Agency for Research on Cancer.
- Molinaro, E., Romei, C., Biagini, A., Sabini, E., Agate, L., Mazzeo, S., et al. (2017). Anaplastic thyroid carcinoma: From clinicopathology to genetics and advanced therapies. *Nature Reviews Endocrinology*, 13(11), 644–660. <https://doi.org/10.1038/nrendo.2017.76>
- Nikiforov, Y. E., & Nikiforova, M. N. (2011). Molecular genetics and diagnosis of thyroid cancer. *Nature Reviews Endocrinology*, 7(10), 569–580. <https://doi.org/10.1038/nrendo.2011.142>
- Pan, H., Luo, Z., Lin, F., Zhang, J., Xiong, T., Hong, Y., et al. (2024). FN1, a reliable prognostic biomarker for thyroid cancer, is associated with tumor immunity and an unfavorable prognosis. *Oncology Letters*, 28(5), Article 510. <https://doi.org/10.3892/ol.2024.14660>
- Liu, M., Chen, P., Wei, B., et al. (2025). FN1 shapes the behavior of papillary thyroid carcinoma through alternative splicing of EDB region. *Scientific Reports*, 15, 327.

- Nguyen, L. T. T., Thompson, E. K., Bhimani, N., Duong, M. C., Nguyen, H. G., Bullock, M., et al. (2025). Prognostic significance of key molecular markers in thyroid cancer: A systematic literature review and meta-analysis. *Cancers*, 17(6), 939. <https://doi.org/10.3390/cancers17060939>
- Cheng, S. Y., Wu, A. T. H., Batiha, G. E., Ho, C. L., Lee, J. C., Lukman, H. Y., et al. (2022). Identification of DPP4/CTNNB1/MET as a theranostic signature of thyroid cancer and evaluation of the therapeutic potential of sitagliptin. *Biology*, 11(2), 324. <https://doi.org/10.3390/biology11020324>
- Garcia, C., Buffet, C., El Khattabi, L., Rizk-Rabin, M., Perlempoine, K., Ragazzon, B., et al. (2019). MET overexpression and activation favors invasiveness in a model of anaplastic thyroid cancer. *Oncotarget*, 10(23), 2320–2334. <https://doi.org/10.18632/oncotarget.26813>
- Liu, Y., Yin, Z., Wang, Y., & Chen, H. (2023). Exploration and validation of key genes associated with early lymph node metastasis in thyroid carcinoma using weighted gene co-expression network analysis and machine learning. *Frontiers in Endocrinology*, 14, 1247709. <https://doi.org/10.3389/fendo.2023.1247709>
- Bakht, A., Cannon, S. T., Waltonen, J. D., Shattuck, T. M., & Li, W. (2024). Risk of malignancy and management of indeterminate thyroid nodules with HRAS/NRAS mutations. *Journal of Clinical and Translational Pathology*, 4(3), 106–109.
- Riccio, I., Laforteza, A., Landau, M. B., Hussein, M. H., Linhuber, J., Staav, J., et al. (2025). Decoding RAS mutations in thyroid cancer: A meta-analysis unveils specific links to distant metastasis and increased mortality. *American Journal of Otolaryngology*, 46(1), 104570. <https://doi.org/10.1016/j.amjoto.2024.104570>
- Ye, J., & Zeng, T. (2022). Mining database and verification of PIK3CB as a marker predicting prognosis and immune infiltration in renal clear cell carcinoma. *Medicine*, 101(22), e29254. <https://doi.org/10.1097/MD.00000000000029254>
- Miller, K. A., Degan, S., Wang, Y., et al. (2024). PTEN-regulated PI3K-p110 and AKT isoform plasticity controls metastatic prostate cancer progression. *Oncogene*, 43, 22–34.
- Varela Betancourt, V. V., Acharya, A., Jahan, N., Kapahi, U., Khalid, I., Shah, S. M., et al. (2026). Artificial intelligence in low-dose computed tomography lung cancer screening: Clinical integration, validation, and translational challenges. *Cureus*, 18(4).
- Li, L., Zhu, M., Huang, H., Wu, J., & Meng, D. (2020). Identification of hub genes in anaplastic thyroid carcinoma: Evidence from bioinformatics analysis. *Technology in Cancer Research & Treatment*, 19, 1533033820962135. <https://doi.org/10.1177/1533033820962135>
- Ren, H., Liu, X., Li, F., He, X., & Zhao, N. (2021). Identification of a six-gene prognosis signature for papillary thyroid cancer using multi-omics methods and bioinformatics analysis. *Frontiers in Oncology*, 11, 624421. <https://doi.org/10.3389/fonc.2021.624421>
- Pan, Y., Wu, L., He, S., Wu, J., Wang, T., & Zang, H. (2021). Identification of hub genes in thyroid carcinoma to predict prognosis by integrated bioinformatics analysis. *Bioengineered*, 12(1), 2928–2940. <https://doi.org/10.1080/21655979.2021.1943956>

10. Author Declaration

We, the authors, hereby declare that there is no conflict of interest regarding the publication of this manuscript. We confirm that the manuscript is original, has not been published previously, and is not under consideration for publication elsewhere.