Original Article
Integrative analysis of circular RNA regulatory network in papillary thyroid carcinoma

Ji Young Kim[1], Yeongun Lee[1], So Hee Dho[1], Hyo Jin Park[1], Da-Mi Kim[2], Jae Cheong Lim[2], Seok-Mo Kim[3], Lark Kyun Kim[1]

[1]Department of Biomedical Sciences, Graduate School of Medical Science, Brain Korea 21 Project, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea; [2]Radioisotope Research Division, Korea Atomic Energy Research Institute, Daejeon, Republic of Korea; [3]Department of Surgery, Thyroid Cancer Center, Gangnam Severance Hospital, Institute of Refractory Thyroid Cancer, Yonsei University College of Medicine, Seoul, Republic of Korea

Abstract: Papillary thyroid cancer (PTC) is the most common type of endocrine cancer worldwide. Generally, PTC has an excellent prognosis; however, lymph node metastases and recurrences occur frequently. Over the last decade, circular RNAs (circRNAs), a large class of noncoding RNAs (ncRNAs), have emerged as key regulators of various tumor progression pathways. Here, we aimed to identify novel circRNAs as PTC biomarkers. Differentially expressed circRNAs and mRNAs were analyzed using public datasets from the Gene Expression Omnibus and Cancer Genome Atlas. In addition, we screened for target miRNAs using online prediction databases. Based on these results, we established a circRNA-miRNA-mRNA regulatory network associated with PTC, in which protein-protein interaction networks led to the identification of hub genes. Functional enrichment and survival analyses were performed to gain insights into the biological mechanisms of circRNA involvement. As a result, we found that two circRNAs (hsa_circ_0041829 and has_circ_0092299), four miRNAs (miR-369, miR-486, miR-574, and miR-665), and nine hub genes (BBC3, E2F1, FYN, MAG, SDC1, SDC3, SNAP25, TK1, and TYMS) play significant roles in PTC progression. This study provides a novel framework for understanding the roles of circRNA-miRNA-mediated gene regulation in PTC. It also introduces potential therapeutic targets and prognostic biomarkers, which may serve as a basis for developing targeted therapeutic interventions for PTC.

Keywords: Papillary thyroid cancer, circular RNA, biomarkers, regulatory network

Introduction

Thyroid cancer (THCA) is a common endocrine malignancy, particularly in females, and its incidence is increasing worldwide [1]. Follicular cell-derived thyroid malignancies account for majority of the thyroid tumors and are classified as follicular THCA, papillary thyroid cancer (PTC), poorly differentiated THCA, Hürthle cell cancer, and anaplastic THCA [2]. PTC is the most frequent pathological subtype, representing 80%-85% of all thyroid malignancies [3, 4]. Although PTC has a favorable prognosis, with a 10-year survival rate of over 90% [5], long-term follow-up reveals recurrence in 25% of the patients, and approximately 50% of the patients develop cervical lymph node metastasis [6-8]. Several risk factors, such as radiation expo-sure, endocrine disruptors, and diet (iodine), contribute to the progression of PTC. However, limited genomic and proteomic marker discovery analyses have been performed to date [9], making an in-depth study of the molecular mechanisms underlying PTC progression important for identifying potential therapeutic targets and prognostic biomarkers.

Over the past decade, increasing evidence has indicated that noncoding RNAs (ncRNAs), such as microRNAs (miRNAs), long noncoding RNA (lncRNAs), and circular RNAs (circRNAs), play important roles in the progression and patho-genesis of various types of cancer [10-13]. MiRNAs, composed of 21-25 nucleotides, are small noncoding RNAs that regulate gene expression [14]. Dysregulated miRNAs influ-

ence cell death, proliferation signaling, invasion, angiogenesis, and metastasis [15]. LncRNAs are noncoding transcripts with a length of over 200 nucleotides that function as either oncogenes or tumor suppressors in various diseases [16, 17]. Although miRNAs and lncRNAs are potential biomarkers for the molecular diagnosis of human diseases, they do not provide sufficient information related to specific diseases [18]. CircRNAs have gained considerable attention as a novel research direction and approach to overcome these limitations and are being explored as potential diagnostic and prognostic biomarkers, as evidenced by their detectability in liquid biopsy samples, such as blood, plasma, serum, and urine [19]. CircRNAs are a subclass of ncRNAs that are widely expressed in different cellular contexts and play various roles in growth and development [20]. CircRNAs are produced by back-splicing of linear precursor RNA, resulting in a covalently closed-loop structure, and lack a 5' cap and 3' end poly (A) tail [21, 22]. Based on their origin in different regions of the genome, they are classified into three types: (1) exonic circRNAs, which are generated from direct back-splicing or exon skipping; (2) circular intronic circRNAs, which are produced from the intronic region of pre-mRNAs; and (3) exon-intron circRNAs, which consist of exons and introns [23, 24].

The circRNA/miRNA/mRNA axis plays a crucial role in cancer progression. Several studies have demonstrated that circRNAs are dysregulated in various diseases, including cancer [25]. CircRNAs can act as competitive endogenous RNAs by serving as miRNA sponges, competing with miRNAs to bind to miRNA response elements (MREs) and regulate target gene expression. This process ultimately affects the onset and progression of several diseases [26]. Several circRNAs are upregulated or downregulated in PTC tissues compared with adjacent normal tissues [27, 28]. For example, circTIAM1 acts as a sponge for miR-646 and functions in PTC by targeting miR-646 and the heterogeneous ribonucleoprotein A1 [29]. CircRUNX1 promotes PTC progression and metastasis by sponging miR-296-3p and regulating DDHD2 expression [30]. CircNEURL4 inhibits the proliferation and migration of PTC cells by competitively binding to miR-1278, thereby indirectly increasing LATS1 expression [31].

Although a few independent studies have conducted circRNA microarray and sequencing in PTC, a potential limitation remains owing to the small sample size of these datasets [32-34]. Therefore, the unbiased screening of circRNAs in a larger cohort would be highly beneficial for identifying potential PTC biomarkers. In this study, we aimed to identify potential PTC biomarkers with greater confidence and accuracy by collecting circRNA microarray datasets from the Gene Expression Omnibus (GEO) database and analyzing differentially expressed circRNAs (DECs) across datasets by selecting overlapping DECs. This study provides a novel framework for understanding the roles of circRNA-miRNA-mediated gene regulation in PTC. Thus, it may serve as a basis for developing targeted therapeutic interventions and prognostic biomarkers for PTC.

Materials and methods

Datasets and identification of DECs and differentially expressed genes (DEGs)

We selected the GSE93522 and GSE173299 datasets from the GEO database, which provide circRNA microarray data for PTC. Each pair of samples from GSE93522 represents six PTC tumors and six matching contralateral normal samples. GSE173299 represents three pairs of PTC tumors and three matching adjacent normal tissues. Next, we performed a differential analysis (P < 0.05, fold change > 1.5) by comparing tumor tissues to normal tissues in R software using the limma package [35]. The normalized microarray data were analyzed, and the outcomes were compared for further investigation. Overlapping DECs were identified by intersecting DECs from the two datasets. To analyze the DEGs between cancerous and cancer-adjacent tissues of patients with THCA, we downloaded the gene expression data of tumor (n=510) and normal (n=58) tissues of patients with THCA from The Cancer Genome Atlas (TCGA) GDC website (https://portal.gdc.cancer.gov/). We then performed a differential analysis ($|log_2FC| > 2$ and adjusted P < 0.05) by comparing tumor tissues to normal tissues in R software using the Deseq2 package [36]. The results are displayed using volcanic plots.

RNA extraction and qRT-PCR

The total RNA was isolated from tissues using TRIzol reagent (Invitrogen, CA, USA), according

to the manufacturer's protocol. cDNA was synthesized with random primers using a SuperScript™ IV Reverse Transcriptase kit (Invitrogen). qRT-PCR was performed on the Roche Real-Time PCR System (LightCycler 480). Primers for circRNAs were synthesized by Bionics. The relative circRNA expression levels were analyzed using the $2^{-\Delta\Delta Ct}$ method. GAPDH were used as the internal control genes. For the RNase R treatment, 1 µg total RNA was digested at 37°C for 30 min and 70°C for 10 min with 2 U/µg RNase R (BiosearchTechnology, RNR07250). The primers used for qRT-PCR analysis were as follows: hsa_circ_0000253 (F 5'-AAGAGCTTAGGTGGTGTGGG-3', R 5'-TCCATC-TTCCCTCACAGCAG-3'), hsa_circ_0011385 (F 5'-TGACAACAATGAGCCCTACATG-3', R 5'-TGCTT-GTCCGTGGAGAACAT-3'), hsa_circ_0005777 (F 5'-AGCAGTCCTTCATGTCACCA-3', R 5'-GGTAGTG-GCAGTGATTCCCT-3'), hsa_circ_0021549 (F 5'-CTACCGTTTCCCCTCTGTGT-3', R 5'-TGTCATATG-GGATGGGGTCG-3'), hsa_circ_0021553 (F 5'-ATGACACTCCAAAACCAGCG-3', R 5'-AGCAAAAG-ATGGAAGCAGGC-3'), hsa_circ_0041829 (F 5'-CATGGACAACAGCCTGGC-3', R 5'-TGGACACTTG-GACACACTGG-3'), hsa_circ_0044556 (F 5'-GG-TCCTGATGGCAAAACTGG-3', R 5'-CAACACCATC-TGCGCCAG-3'), hsa_circ_0048937 (F 5'-GTG-GATCACTGGGAAGACGA-3', R 5'-GATGAAGAGC-TGGGCAAAGG-3'), hsa_circ_0076092 (F 5'-GTATTGCCTGTGAATGCCGT-3', R 5'-GATGTGGC-AGTTGTCAGTCC-3'), hsa_circ_0079891 (F 5'-CAAGTCCATGTACACGCGAG-3', R 5'-CCACCCAT-CACAGACTTCCT-3'), hsa_circ_0092278 (F 5'-AGCCAGACCTTTTGACTCCA-3', R 5'-GGCTATGC-TTTGTGAGGCTG-3'), hsa_circ_0092299 (F 5'-TCTGTGACTTTGTGCTTGGC-3', R 5'-GAAGCAAG-ACACACACTGCA-3') and GAPDH (F 5'-AATC-CCATCACCATCTTCCA-3', R 5'-TGGACTCCACGA-CGTACTCA-3').

Prediction of MREs and miRNA target genes

MREs in DECs were predicted using two web tools: CircInteractome (https://circinteractome.nia.nih.gov/) [37] and CircBank (www.circbank.cn) [38]. Overlapping miRNAs from CircBank and CircInteractome were used for further analyses. Target genes of the miRNAs were predicted using the open online bioinformatics tool TargetScan (https://www.targetscan.org/vert_80/).

Cancer-specific circRNA database (CSCD) analysis

CSCD is a database developed for cancer-specific circRNAs that collects available RNA sequencing (total RNA with rRNA-depleted or polyA-enriched) datasets from 87 cancer cell line samples (http://gb.whu.edu.cn/CSCD/) [39]. We used CSCD to obtain the structural ring diagram of each candidate circRNA.

Expression and survival analyses using database

The expression levels of interacting miRNAs in THCA were further analyzed using the UALCAN database (http://ualcan.path.uab.edu/index.html) [40]. Statistical significance was set at P < 0.05. Immunohistochemical data for the hub genes in thyroid and PTC tissues were downloaded from the Human Protein Atlas (HPA; https://www.proteinatlas.org/) using the "Tissue" and "Pathology" modules, respectively. The Kaplan-Meier plotter (http://kmplot.com/analysis/) was used to evaluate the influence of different expression levels of hub genes on the relapse-free survival (RFS) and overall survival (OS) of patients with THCA with different clinical factors [41]. Hazard ratios, 95% confidence intervals (95% CI), and log-rank *P* value were calculated and displayed in survival charts. Patients were split based on "automatically select the best cut-off point", and the follow-up threshold covered all patients. Statistical significance was set at P < 0.05.

Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis

Target genes of DECs were analyzed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID; https://david.ncifcrf.gov/) for GO function enrichment and KEGG pathway analyses [42]. Statistical significance was set at P < 0.05. The GO enrichment analysis was performed using the R package clusterProfiler.

Construction of a protein-protein interaction (PPI) network and identification of hub genes

The predicted miRNA target genes and DEGs from TCGA-THCA were intersected to obtain

overlapping genes. The PPI network of the intersecting DEGs was generated using the Search Tool for the Retrieval of Interacting Genes (STRING) online tool (https://string-db.org/) with a confidence score of 0.4. The results were imported and visualized using the Cytoscape software (version 3.9.1) [43, 44]. The Maximal Clique Centrality (MCC) algorithm, an effective method for searching hub nodes, was utilized via the CytoHubba plugin in Cytoscape to select the top 10 genes [45]. The hub gene subnetwork was extracted using CytoHubba to visualize the genes that directly interacted with these top-ranked nodes.

Results

Identification of DECs in PTC

In this study, we aimed to identify novel circRNAs that may be involved in controlling PTC progression. First, we identified DECs between normal thyroid tissues and PTC using a public microarray dataset. Next, we established a putative circRNA/miRNA/mRNA regulatory axis linked to PTC progression by predicting interacting miRNAs and their target genes and performing PPI network analysis, pathway enrichment analysis, survival analysis, and hub gene identification (**Figure 1A**). To identify potential circRNAs and construct a network of circRNA/miRNA/mRNA interactions in PTC, we used two microarray datasets (GSE93522 and GSE173299) from the GEO database to screen for overlapping circRNAs that were differentially expressed in PTC tissues compared to adjacent non-tumor tissues. The volcano plot displays the distribution of differentially expressed circRNAs (fold change > 1.5, P < 0.05) between PTC and non-tumor tissues in the two datasets (**Figure 1B**). In total, 267 DECs, including 179 upregulated and 88 downregulated circRNAs, were identified in GSE93522, whereas 1147 DECs, including 562 upregulated and 585 downregulated circRNAs, were identified in GSE173299. We integrated the DECs into a Venn diagram, which showed 27 upregulated and 19 downregulated circRNAs common to both microarray datasets (**Figure 1C**). To identify significantly differentially expressed circRNAs, we selected overlapping DECs with fold-change ratios greater than two in both datasets for further analyses. The seven overlapping circRNAs (four upregulated and three downregu-

lated), along with their circBase ID, genomic locations, circRNA types, and host genes, are listed in **Table 1**. The structural patterns of candidate circRNAs predicted using the CSCD database are shown in **Figure 1D**. Six exonic circRNAs (hsa_circ_0091710, hsa_circ_0041829, hsa_circ_0011385, hsa_circ_0079891, hsa_circ_0005777, and hsa_circ_0021549) exhibited structural ring diagrams. However, the structural pattern of the upregulated intronic circRNA hsa_circ_0092299 was not observed in CSCD. The expression levels of the seven overlapping circRNAs were assessed using qRT-PCR in 12 pairs of PTC samples along with their corresponding adjacent nontumor tissues. As shown in **Figure 1E**, all upregulated circRNAs demonstrated a significant increase in PTC tissues compared to matched nontumor tissues (P < 0.05), consistent with the microarray data. Similarly, all downregulated circRNAs displayed decreased expression in PTC tissues (P < 0.05).

Generally, it is crucial to recognize the significance of normalization in maintaining the accuracy and reproducibility of the analysis. As a result, we conducted data normalization for the analysis of the GEO dataset and subsequently compared the outcomes. After normalization, we depicted differentially expressed circRNAs (fold change > 1.5, P < 0.05) through a volcano plot comparing PTC and non-tumor tissues (Supplementary Figure 1A). Utilizing a Venn diagram, we illustrated 84 upregulated and 59 downregulated circRNAs shared between both microarray datasets (Supplementary Figure 1B). In accordance with the previous analysis, we selected overlapping DECs with fold-change ratios exceeding two in both datasets. Intriguingly, among the thirteen overlapping circRNAs, seven had been previously identified. The remaining six newly discovered overlapping circRNAs (three upregulated and three downregulated) are enumerated in **Table 1**. Supplementary Figure 1C showcased the ring structures of four exonic circRNAs (hsa_circ_0021553, hsa_circ_0044556, hsa_circ_0048937, and hsa_circ_0076092). However, the structural configurations of the intronic circRNAs, hsa_circ_0000253 and hsa_circ_0092278, were absent in the CSCD dataset. The expression levels of the six circRNAs were assessed using qRT-PCR in 12 pairs of PTC samples along with their corresponding adja-
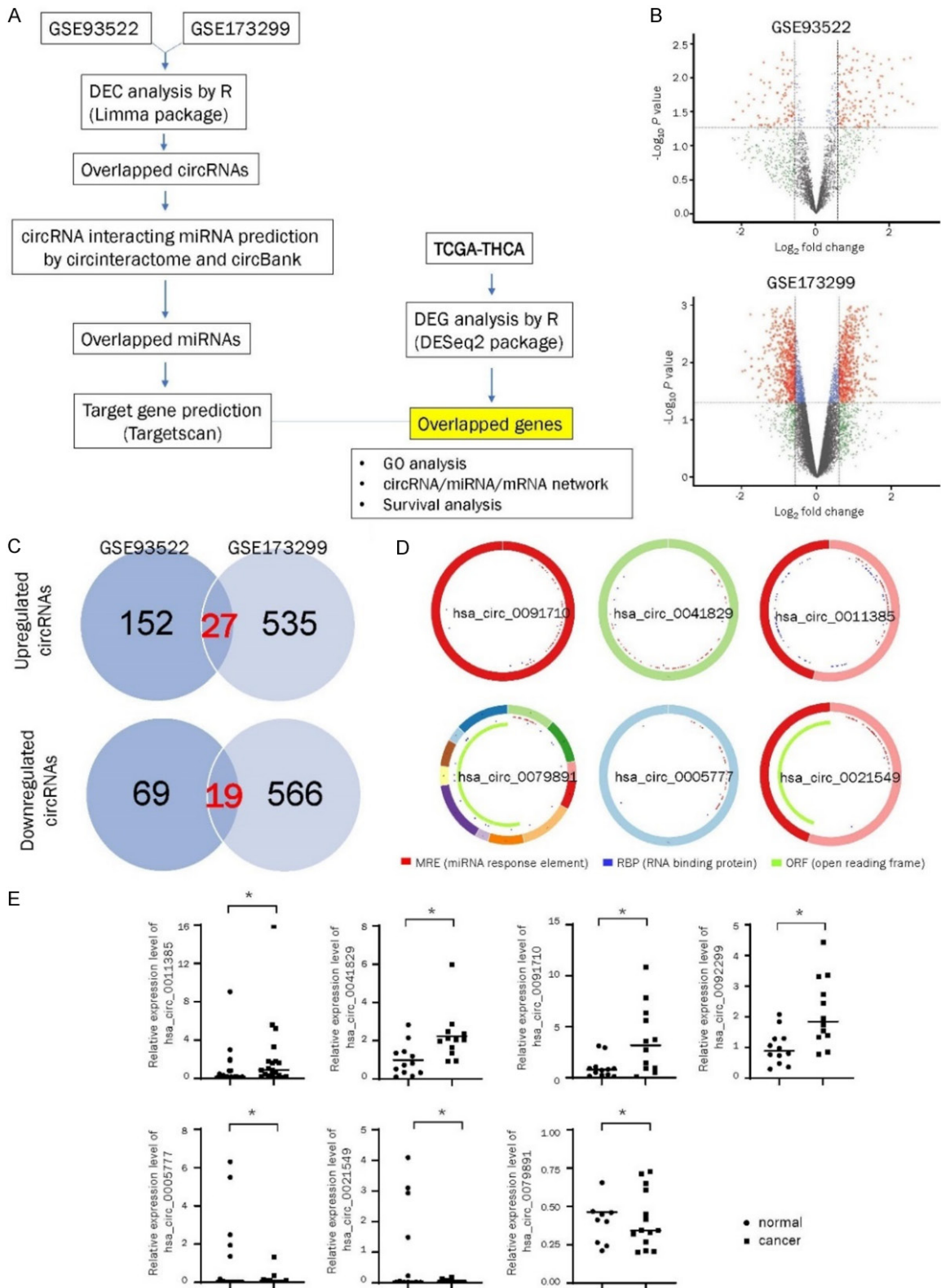
Figure 1. Identification of differentially expressed circRNAs (DECs) in thyroid cancer. A. Flowchart of this study. B. Identification of DECs in two Gene Expression Omnibus (GEO) datasets. Volcano plots depicting the DECs obtained from GEO datasets, where red dots denote significantly upregulated and downregulated circRNAs. A threshold of $P < 0.05$ and fold change > 1.5 were set. C. Venn diagram showing 27 overlapping upregulated DECs (upper) and 19 overlapping downregulated DECs (lower) between the two GEO datasets. D. Structural patterns of the six cir-

cRNAs by the Cancer-specific CircRNA Database (CSCD). MRE: miRNA response element; RBP: RNA binding protein; ORF: open reading frame. E. qRT-PCR validation of seven differentially expressed circRNAs. The relative expression levels of circRNAs in 12 PTC tissues and paired normal tissues. Notably, circ_0011385 was analyzed in 20 pairs, circ_0005777 in 18 pairs, and circ_0079891 in 14 pairs, enhancing the comprehensive assessment of their expression patterns; *P < 0.05 (•: normal tissues, ▪: PTC tissues).

Table 1. Basic information on circRNAs with $|\log_2 FC| > 1$ (upper) or < -1 (lower) in the two GEO datasets

$|\log_2 FC| > 1$								
circRNA	circbase	chrom	strand	txStart	txEnd	type	best_transcript	host gene
hsa_circRNA_105037	hsa_circ_0091710	chrX	-	151130894	1.51E+08	exonic	NM_004961	GABRE
hsa_circRNA_400070	hsa_circ_0092299	chr22	+	35819173	35819173	intronic	ENST00000216122	MCM5
hsa_circRNA_101971	hsa_circ_0041829	chr17	-	7226185	7226185	exonic	NM_032442	NEURL4
hsa_circRNA_100146	hsa_circ_0011385	chr1	+	32692131	32692131	exonic	NM_003757	EIF3I
hsa_circRNA_102121	hsa_circ_0044556	chr17	-	48271490	48272189	exonic	NM_000088	COL1A1*
hsa_circRNA_102430	hsa_circ_0048937	chr19	+	6934997	6937659	exonic	NM_001974	EMR1*
hsa_circRNA_400031	hsa_circ_0092278	chr17	+	4797948	4798348	intronic	ENST00000347992	MINK1*
$|\log_2 FC| < -1$								
circRNA	circbase	chrom	strand	txStart	txEnd	type	best_transcript	host gene
hsa_circRNA_104348	hsa_circ_0079891	chr7	-	37250990	37382367	exonic	NM_014800	ELMO1
hsa_circRNA_103888	hsa_circ_0005777	chr5	+	73136304	73136585	exonic	NM_001080479	ARHGEF28
hsa_circRNA_100775	hsa_circ_0021549	chr11	-	30516842	30557722	exonic	NM_001584	MPPED2
hsa_circRNA_001350	hsa_circ_0000253	chr10	-	97999787	97999925	intronic	NR_047681	BLNK*
hsa_circRNA_100777	hsa_circ_0021553	chr11	-	30557540	30602041	exonic	NM_001145399	MPPED2*
hsa_circRNA_104099	hsa_circ_0076092	chr6	+	35195356	35201078	exonic	NM_152753	SCUBE3*

Asterisks (*) denote a newly identified circRNAs.

cent nontumor tissues. The findings indicated that all three upregulated circRNAs demonstrated a significant increase in PTC tissues compared to matched nontumor tissues (P < 0.05). Similarly, all three downregulated circRNAs displayed decreased expression in PTC tissues (P < 0.001) (Supplementary Figure 1D).

Prediction of circRNA-interacting miRNAs and identification of DEGs

The most commonly reported role of circRNAs in cancer is to act as miRNA sponges [46, 47]. To identify potential miRNAs that interact with the six circRNAs, we used a circRNA-interacting miRNA database from CircBank and CircInteractome and selected the intersection of the prediction results of both databases. We found that 33 miRNAs may interact with the six circRNAs. Using Venn diagram analysis, we identified 14 intersecting miRNAs for hsa_circ_0092299, 1 intersecting miRNA for hsa_circ_0041829, 2 intersecting miRNAs for hsa_circ_0011385 and hsa_circ_0005777, 10 intersecting miRNAs for hsa_circ_0079891, and 4 intersecting miRNAs for hsa_circ_

0021549, whereas hsa_circ_0091710 had zero interacting miRNAs (**Figure 2A** and **Table 2**). Furthermore, our findings reveal potential interactions between 20 miRNAs and the six newly discovered circRNAs. Illustrated in Supplementary Figure 1E are the specific details: we identified 2 intersecting miRNAs for hsa_circ_0044556, 5 intersecting miRNAs for hsa_circ_0048937, 5 intersecting miRNAs for hsa_circ_0092278, 2 intersecting miRNAs for hsa_circ_0000253, 4 intersecting miRNAs for hsa_circ_0021553, and 2 intersecting miRNAs for hsa_circ_0076092 (**Table 2**).

We then examined the differential expression of these 33 intersecting miRNAs in PTC using the TCGA database (TCGA-THCA) and analyzed the statistically significant results using UALCAN (P < 0.05). We identified six miRNAs (miR-361, miR-369, miR-486, miR-574, miR-607, and miR-665) that exhibited a negative correlation and one miRNA (miR-182) that exhibited a positive correlation with circRNA expression. Notably, all six miRNAs were significantly downregulated in the THCA group (**Figure 2B**). We also assessed the expression of 20
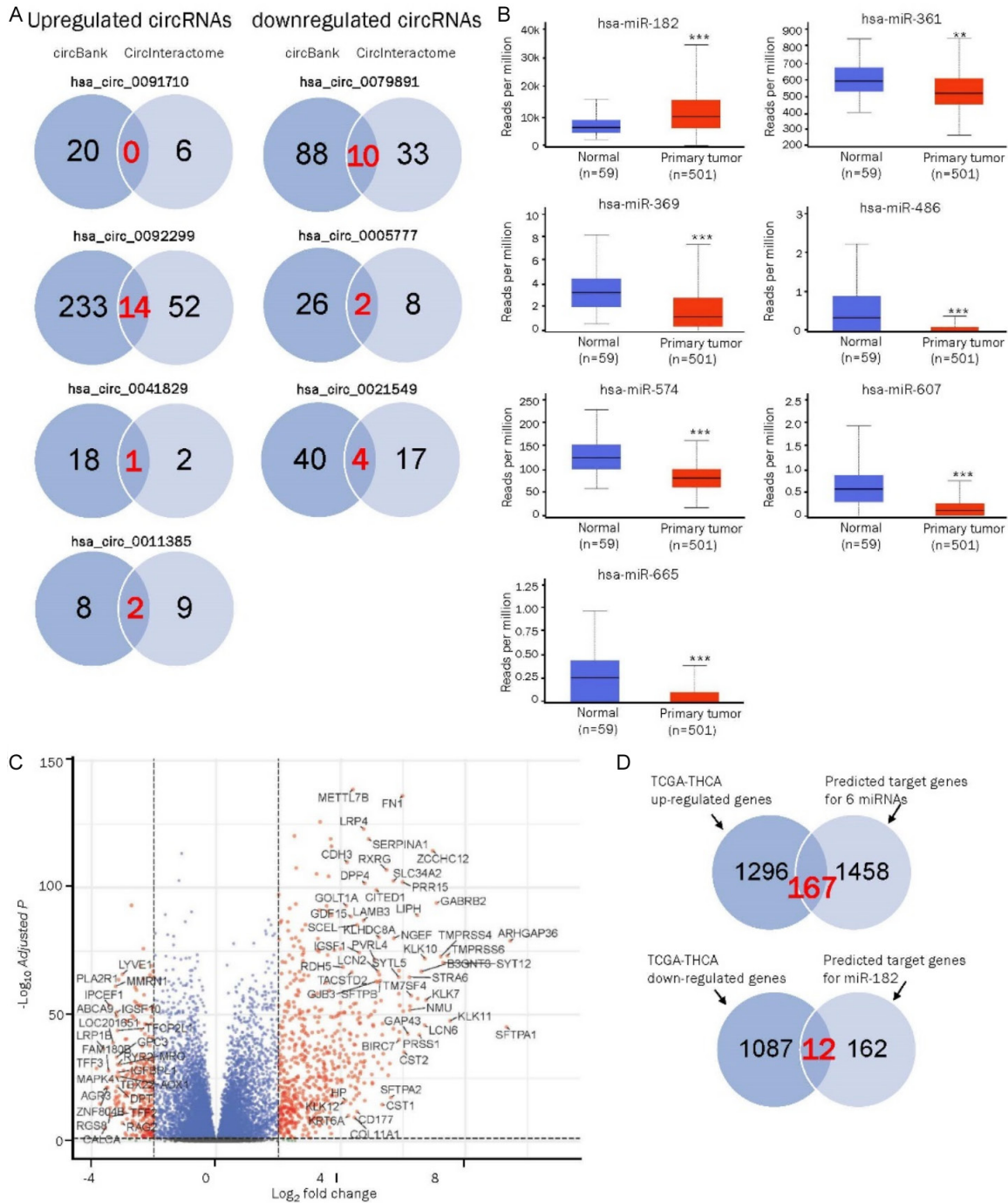
Figure 2. Identification of circRNA-miRNA interactions and differentially expressed genes (DEGs) in the The Cancer Genome Atlas (TCGA) dataset of thyroid cancer. A. Venn diagram between the seven target miRNAs predicted by CircBank and CircInteractome. B. TCGA expression profiles for the seven miRNAs in thyroid cancer samples (n=501, red) and normal samples (n=59, blue). C. Volcano plots of DEGs in the The Cancer Genome Atlas (TCGA)-Thyroid Cancer (THCA) dataset. The cutoff criteria used were |log$_2$FC| ≥ 1.0 and adjusted P < 0.05. D. Venn diagram represent the overlap between miRNA target genes and DEGs from TGCA. Upper; 167 mRNAs were obtained from the intersection of 1,463 upregulated mRNAs (TCGA-THCA dataset) and 1,625 miRNA targets predicted using TargetScan. Lower; 12 mRNAs were obtained from the intersection of 1,099 downregulated mRNAs (TCGA-THCA dataset) and 174 miRNA targets predicted using TargetScan.

newly identified overlapping miRNAs using UALCAN. Among these, we pinpointed four miR-

NAs (miR-502, miR-589, miR-624, and miR-643) that displayed a negative correlation with

Table 2. Identification of miRNAs interacting with the six circRNAs using the CircBank and CircInteractome databases

Up-regulated circRNAs	Interacting miRNAs
hsa_circ_0092299	hsa-miR-324-5p
	hsa-miR-486-3p*
	hsa-miR-515-3p
	hsa-miR-574-5p*
	hsa-miR-607*
	hsa-miR-622
	hsa-miR-635
	hsa-miR-646
	hsa-miR-648
	hsa-miR-665*
	hsa-miR-1182
	hsa-miR-1184
	hsa-miR-1303
	hsa-miR-1324
hsa_circ_0041829	hsa-miR-369-5p*
hsa_circ_0011385	hsa-miR-361-3p*
	hsa-miR-1272
hsa_circ_0044556	hsa-miR-589-5p*
	hsa-miR-634
hsa_circ_0048937	hsa-miR-1205
	hsa-miR-624-3p*
	hsa-miR-630
	hsa-miR-643*
	hsa-miR-665*
hsa_circ_0092278	hsa-miR-1236-3p
	hsa-miR-149-5p
	hsa-miR-331-3p
	hsa-miR-432-5p
	hsa-miR-502-5p*
Down-regulated circRNAs	Interacting miRNAs
hsa_circ_0079891	hsa-miR-361-3p*
	hsa-miR-1272
	hsa-miR-182-5p*
	hsa-miR-384
	hsa-miR-512-5p
	hsa-miR-515-5p
	hsa-miR-558
	hsa-miR-576-3p
	hsa-miR-647
	hsa-miR-665
	hsa-miR-1206
	hsa-miR-1243
	hsa-miR-1256
	hsa-miR-1270
hsa_circ_0005777	hsa-miR-607
	hsa-miR-1225-5p
hsa_circ_0021549	hsa-miR-346
	hsa-miR-553
	hsa-miR-574-5p
	hsa-miR-657
hsa_circ_0000253	hsa-miR-1236-3p
	hsa-miR-1324
hsa_circ_0076092	hsa-miR-1206
	hsa-miR-643
hsa_circ_0021553	hsa-miR-1257
	hsa-miR-548p
	hsa-miR-574-5p
	hsa-miR-661

Asterisks (*) denote a negative correlation between circRNA and miRNA expression.

circRNA expression. Consistent with the aforementioned outcome, all four of these recently discovered miRNAs exhibited significant downregulation in the THCA group (Supplementary Figure 1F).

To investigate the function of these circRNAs in PTC, we used the DESeq2 package to analyze the TCGA-THCA data with the criteria of $|\log_2 FC| > 1$ and adjusted P < 0.05. We identified 2,562 genes, including 1,463 upregulated and 1,099 downregulated genes (**Figure 2C**). The target genes of the seven selected miRNAs were predicted using the TargetScan database. By intersecting the 1,625 predicted target genes for the six downregulated miRNAs with the 1,463 upregulated genes from the TCGA-THCA analysis and the 174 predicted target genes for miR-182 with the 1,099 downregulated genes from the TCGA-THCA analysis, we obtained 179 shared DEGs (167 upregulated and 12 downregulated genes) (**Figure 2D**). Furthermore, we acquired 187 common DEGs by intersecting the 1,921 predicted target genes of the ten downregulated miRNAs (miR-361, miR-369, miR-486, miR-502, miR-574, miR-589, miR-607, miR-624, miR-643, and miR-665) with the 1,463 upregulated genes identified in the TCGA-THCA analysis (Supplementary Figure 1G).

Functional and pathway enrichment analyses

The DEGs targeted by the four circRNA-seven miRNAs were subjected to GO functional annotation and pathway analyses using DAVID software to understand their functions (**Table 3**). In the biological process (BP) category of GO,

Table 3. Gene ontology (GO) enrichment analysis

BP term	*P*-value
GO:0007267~cell-cell signaling	2.18E-06
GO:0044707~single-multicellular organism process	2.32E-05
GO:0048869~cellular developmental process	2.54E-05
GO:0030154~cell differentiation	6.38E-05
GO:0050804~modulation of synaptic transmission	6.53E-05
GO:0051701~interaction with host	7.02E-05
GO:0044700~single organism signaling	7.25E-05
GO:0032501~multicellular organismal process	9.50E-05
GO:0023052~signaling	9.85E-05
GO:0032502~developmental process	1.01E-04
CC term	***P*-value**
GO:0005886~plasma membrane	9.66E-08
GO:0071944~cell periphery	2.34E-07
GO:0032279~asymmetric synapse	7.57E-06
GO:0098984~neuron to neuron synapse	9.27E-06
GO:0043005~neuron projection	2.02E-05
GO:0097458~neuron part	2.83E-05
GO:0016020~membrane	3.05E-05
GO:0014069~postsynaptic density	3.20E-05
GO:0031224~intrinsic component of membrane	6.39E-05
GO:0099572~postsynaptic specialization	7.07E-05
MF term	***P*-value**
GO:0001618~virus receptor activity	4.50E-05
GO:0042802~identical protein binding	5.59E-04
GO:0005509~calcium ion binding	0.001598
GO:0005488~binding	0.005114
GO:0048306~calcium-dependent protein binding	0.007423
GO:0019901~protein kinase binding	0.008912
GO:0005515~protein binding	0.009731
GO:0008324~cation transmembrane transporter activity	0.011631
GO:0034185~apolipoprotein binding	0.011944
GO:0046873~metal ion transmembrane transporter activity	0.012133
KEGG pathway	***P*-value**
hsa04514: cell adhesion molecules	0.003455
hsa04115: p53 signaling pathway	0.030927
hsa04215: apoptosis - multiple species	0.03596
hsa01232: nucleotide metabolism	0.04538

The top 10 terms in biological process (BP), cellular component (CC), molecular function (MF), and KEGG pathways.

DEGs were enriched in functional annotations related to cell-cell signaling, single-multicellular organism processes, cell differentiation, and modulation of synaptic transmission. In the cellular component (CC) category, DEGs were enriched in functional annotations related to the plasma membrane, cell periphery, and neu-ron-to-neuron synapses. In the molecular function (MF) category, DEGs were enriched in functional annotations related to viral receptor activity, calcium ion binding, and protein kinase binding. KEGG pathway analysis revealed that the genes were mainly enriched in cell adhesion molecules, the p53 signaling pathway, and apoptosis-multiple species.

Finally, considering all the data, the DEGs targeted by the seven circRNAs (hsa_circ_0011385, hsa_circ_0041829, hsa_circ_0044556, hsa_circ_0048937, hsa_circ_0079891, hsa_circ_0092278, and hsa_circ_009-2299) and eleven miRNAs (miR-182, miR-361, miR-369, miR-486, miR-574, miR-607, miR-665, miR-502, miR-589, miR-624, and miR-643) underwent GO functional annotation. Within the BP category of GO, the DEGs exhibited enrichment in functional annotations associated with chemical regulation, synaptic transmission, and skin development (Supplementary Figure 1H). In the CC category, the DEGs showed enrichment in functional annotations related to neuron-to-neuron synapses, transport vesicles, and membrane microdomains.

PPI network and hub gene identification in PTC

We created a PPI network using the STRING online database and Cytoscape software (**Figure 3A**). Using the MCC algorithm of the CytoHubba plug-in, the nine highest-scoring genes in the PPI network were selected as hub genes. These hub genes were BCL2-binding component 3 (BBC3), E2F transcription factor 1 (E2F1), Src family tyrosine kinase (FYN), FYN proto-oncogene, myelin-associated glycoprotein (MAG), syndecan 1
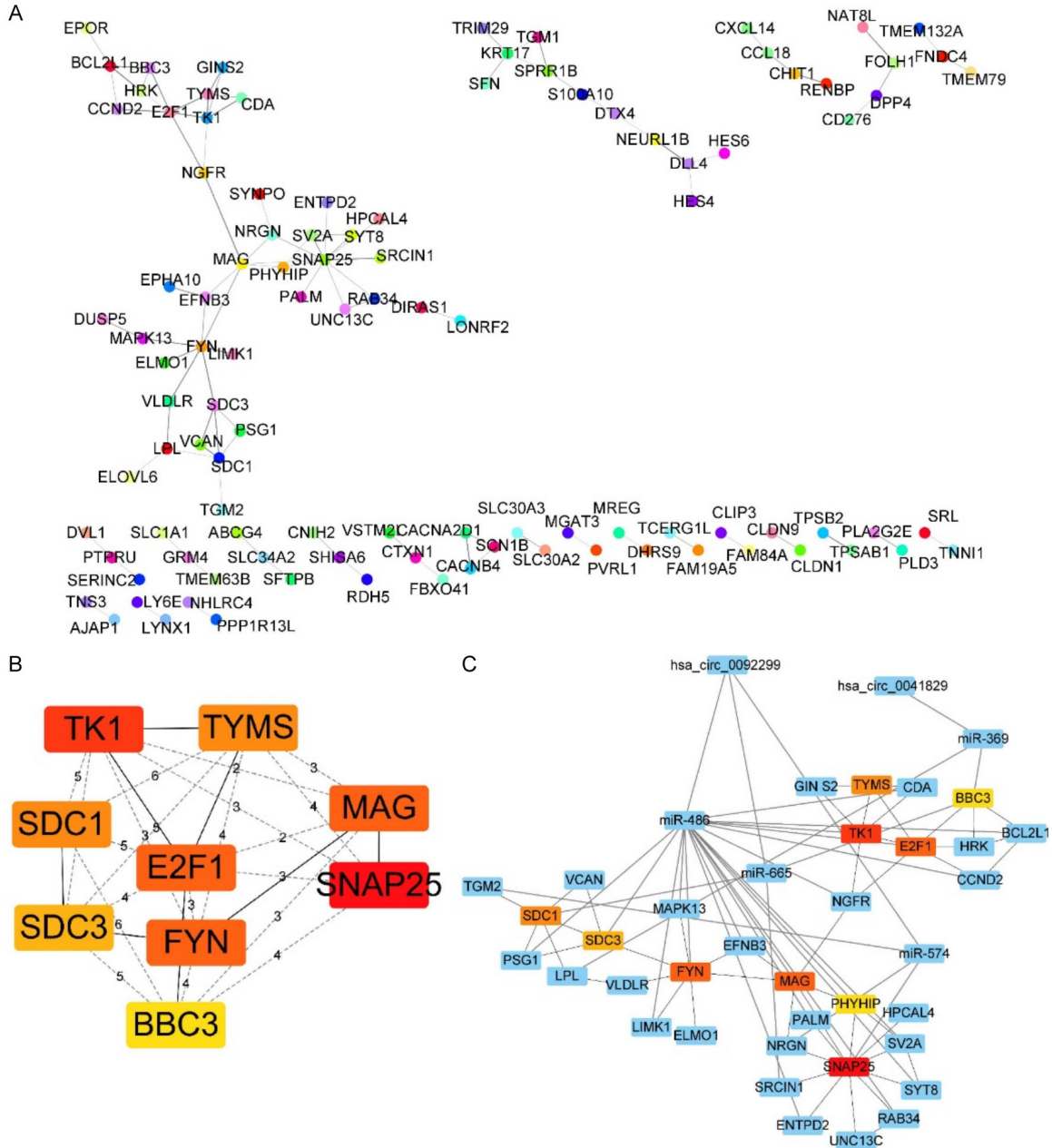
Figure 3. Visualization of the protein-protein interaction (PPI) network and the hub genes. A. PPI network of 179 overlapped genes. The nodes represent the genes, and edges indicate interaction associations between nodes. B. Identification of the hub genes from the PPI network. The genes with the top 9 Maximal Clique Centrality (MCC) values were considered hub genes. Edges represent the protein-protein associations. Red to yellow color gradients indicate the higher MCC ranking of hub genes. Numbers represent path length. C. circRNA-miRNA-hub gene subnetwork. Red-orange-yellow represents hub genes.

(SDC1), syndecan 3 (SDC3), synaptosome-associated protein 25 (SNAP25), thymidine kinase 1 (TK1), and thymidylate synthase (TYMS) (**Figure 3B**). A subnetwork of hub genes extracted from CytoHubba showed interconnections with other genes, including miRNAs and circRNAs (**Figure 3C**). Of the 9 hub genes, miR-486, miR-574, and miR-665 were predicted to be hsa_circ_0092299-interacting miRNAs, indicating that hsa_circ_0092299 may play an important role in THCA (**Figure 3C**). GO and KEGG functional enrichment analyses were

Table 4. Gene ontology enrichment analysis of the nine hub genes

BP term	P-value
GO:0048468~cell development	0.001062
GO:0009157~deoxyribonucleoside monophosphate biosynthetic process	0.002893
GO:0043523~regulation of neuron apoptotic process	0.003991
GO:0050767~regulation of neurogenesis	0.004474
GO:0051402~neuron apoptotic process	0.00489
GO:0090150~establishment of protein localization to membrane	0.005486
GO:0051960~regulation of nervous system development	0.006369
GO:0060284~regulation of cell development	0.006974
GO:1901214~regulation of neuron death	0.00875
CC term	**P-value**
GO:0005737~cytoplasm	0.01826
GO:0098805~whole membrane	0.029556
GO:0098552~side of membrane	0.0306
GO:0000323~lytic vacuole	0.034592
GO:0005764~lysosome	0.034592
GO:0043202~lysosomal lumen	0.037849
GO:0031234~extrinsic component of cytoplasmic side of plasma membrane	0.039726
GO:0005796~Golgi lumen	0.040477
GO:0005773~vacuole	0.043079
MF term	**P-value**
GO:0042802~identical protein binding	8.70E-04
REACTOME Term	**P-value**
R-HSA-202733~cell surface interactions at the vascular wall	1.03E-04
R-HSA-69205~G1/S-specific transcription	1.75E-04
R-HSA-69206~G1/S transition	0.003797
R-HSA-453279~mitotic G1 phase and G1/S transition	0.004884
R-HSA-139915~activation of PUMA and translocation to mitochondria	0.006561
R-HSA-109582~hemostasis	0.008209
R-HSA-3656253~defective EXT1 causes exostoses 1, TRPS2, and CHDS	0.01019
R-HSA-3656237~defective EXT2 causes exostoses 2	0.01019
R-HSA-9694614~attachment and entry	0.013807
R-HSA-4420332~defective B3GALT6 causes EDSP2 and SEMDJL1	0.014529

performed to investigate the functions of the hub genes (**Table 4**). In the BP category of GO, the hub genes were enriched in functional annotations related to cell development, deoxyribonucleoside monophosphate biosynthetic processes, regulation of neuronal apoptotic processes, and regulation of neurogenesis. In the CC category, the hub genes were enriched in functional annotations related to the cytoplasm and whole membrane. In the MF category, the hub genes were enriched in identical protein binding. According to the results of the REACTOME pathway analysis, the hub genes were mainly enriched in cell surface interactions at the vascular wall, G1/S transition, and hemostasis.

We constructed a PPI network involving the DEGs targeted by the eleven miRNAs (Supplementary Figure 2A). Employing the MCC algorithm from the CytoHubba plug-in, we identified ten hub genes. Interestingly, among these ten hub genes, five (BBC3, E2F1, FYN, SNAP25, and TK1) had been previously recognized. The newly discovered hub genes included BCL2 Like 1 (BCL2L1), C-X-C Motif Chemokine Ligand 8 (CXCL8), Harakiri (HRK), Jagged Canonical Notch Ligand 2 (JAG2), and Phorbol-12-Myristate-13-Acetate-Induced Protein 1 (PM-AIP1) (Supplementary Figure 2B). By analyzing a subnetwork of these hub genes, we predicted miR-624, miR-643, and miR-665 to be miRNAs interacting with hsa_circ_0048937, suggest-

ing a potentially significant role for hsa_circ_0048937, along with hsa_circ_0092-299, in the context of THCA (Supplementary Figure 2C).

Verification of prognostic values of hub genes

The expression levels of the nine hub genes BBC3, E2F1, FYN, MAG, DC1, SDC3, SNAP25, TK1, and TYMS were examined in TCGA-THCA using the GEPIA database. All nine hub genes were significantly upregulated in THCA tissues compared with normal tissues (**Figure 4A-J**). Additionally, newly identified five hub genes were also upregulated in THCA tissues (**Figure 4K-N**). Furthermore, immunohistochemical staining results derived from a freely accessible HPA database, enabling examination of the expression of specific protein-coding genes in different cancer types, revealed the upregulation of seven genes in PTC [48]. However, the gene MAG exhibited moderate staining intensity in both normal and cancerous tissues (**Figure 5A-H**). Notably, strong staining for SNAP25 was observed in only one of the three samples, and BBC3 expression was not observed in either the normal or cancerous tissues (data not shown). In cancerous thyroid tissues, we noticed moderate and high staining intensities for BCL2L1 and JAG2, respectively (**Figure 5I**, **5J**). However, CXCL8 expression wasn't identified in either normal or cancerous tissues, and there was no discernible difference in the staining intensity of HRK and PMAIP1 between the normal and cancerous tissues (data not shown).

In addition, using the Kaplan-Meier survival curve, we performed RFS and OS analyses of the nine hub genes to investigate their prognostic value in thyroid cancer patients. Thyroid cancer patients were categorized into the high- and low-expression groups based on the gene's expression levels, using the median as the cut-off point. The results indicated that there was no significant association between RFS of thyroid cancer patients and the expression levels of BBC3 (P=0.48), FYN (P=0.017), MAG (P=0.77), SDC1 (P=0.82), SDC3 (P=0.29), and SNAP25 (P=0.12) as shown in **Figure 6A**. However, high expressions of E2F1 (P=0.042), TK1 (P=0.0071), and TYMS (P=0.0014) were associated with unfavorable RFS in thyroid cancer patients. Nevertheless, there were no significant differences observed between low and high gene expressions in OS of thyroid cancer

patients, including BBC3 (P=0.85), E2F1 (P=0.064), FYN (P=0.28), MAG (P=0.28), SDC1 (P=0.02), SDC3 (P=0.032), SNAP25 (P=0.39), TK1 (P=0.45), and TYMS (P=0.021) (**Figure 6B**). Notably, we found no noteworthy correlation between the OS and RFS of thyroid cancer patients and the expression levels of newly identified hub genes: E2F1, CXCL8, HRK, JAG2, and PMAIP1.

Discussion

This study provides a circRNA-miRNA-mRNA regulatory network in PTC progression. To identify the differentially expressed circRNAs in PTC, we profiled public GEO datasets and identified two significantly upregulated circRNAs: hsa_circ_0041829 and hsa_circ_0092299. Additionally, we predicted four miRNAs (miR-369, miR-486, miR-574, and miR-665) that interact with these circRNAs. Moreover, using the PPI network, we identified nine hub genes, BBC3, E2F1, FYN, MAG, DC1, SDC3, SNAP25, TK1, and TYMS, all of which were upregulated in PTC.

While batch correction and normalization are typically advised for GEO datasets involving RNA-seq data from both normal and cancerous tissues, we entertained the notion that these procedures might exhibit a diminished impact on our analysis. This stems from the fact that, in both instances-GSE93522 and GSE173299-researchers meticulously devised the experimental framework to encompass matched pairs of normal and cancerous tissues within the same individual. Consequently, we posited that the paired nature of the samples might potentially alleviate batch effects, given that any systematic variation affecting both normal and cancerous tissues equally could be effectively eliminated during the analysis. Nonetheless, it's imperative to acknowledge that these measures are still generally recommended to uphold the precision and reproducibility of the analysis. Therefore, we conducted data normalization for the analysis of the GEO dataset and subsequently compared the results. Ultimately, we discovered thirteen overlapping circRNAs, and notably, among these thirteen, seven had been previously identified. This substantiates our speculation that the paired nature of the samples could indeed alleviate batch effects. The expression of BBC3, a pro-apoptotic BH3-only gene, is induced in
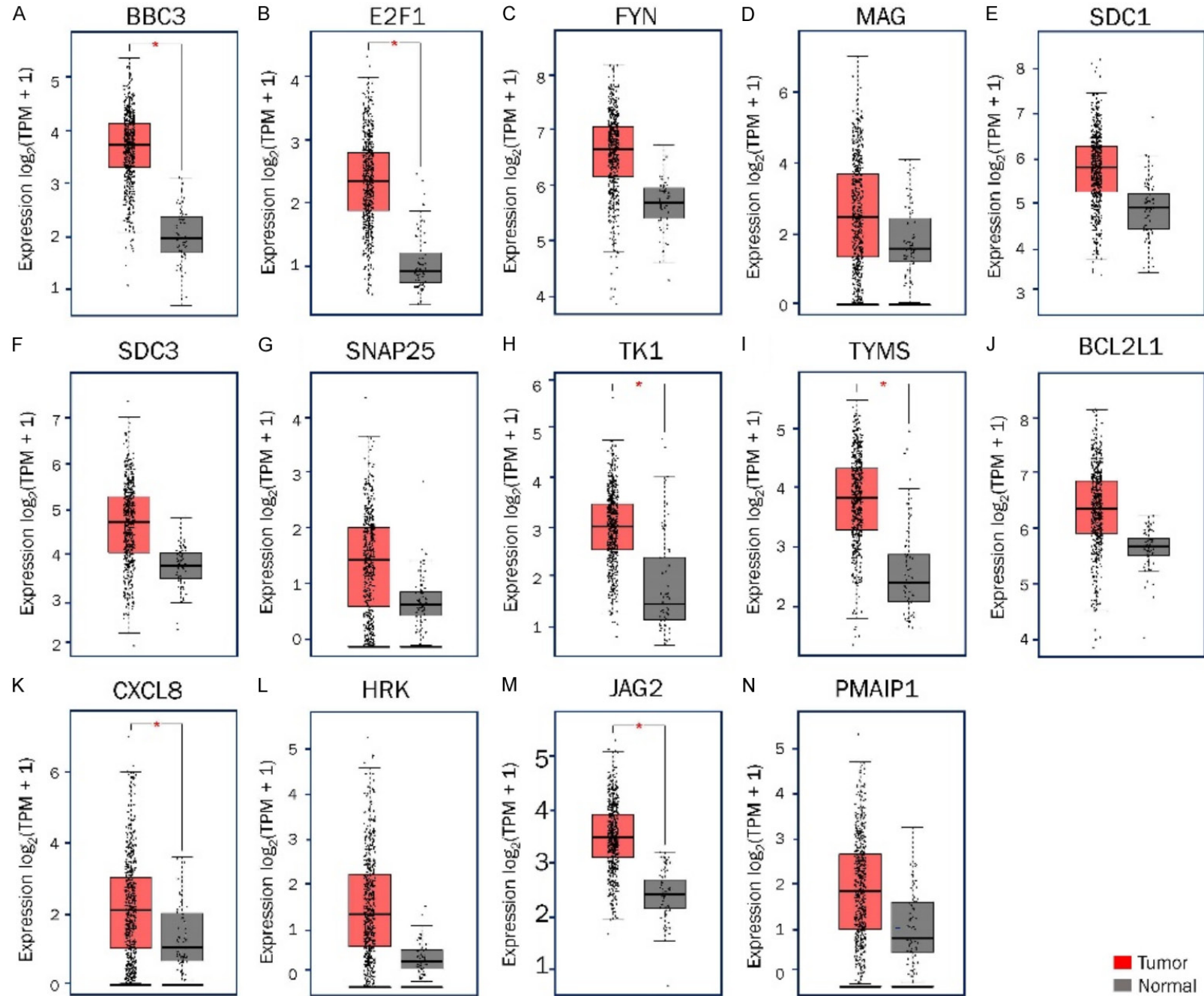
circRNA regulatory network in PTC

Figure 4. Expression of hub genes in thyroid cancer (THCA) using the GEPIA database analysis. Expression of nine hub genes (A-J) and newly identified four hub genes (K-N) in THCA based on sample types of normal (black) and primary tumor (red) from the Cancer Genome Atlas samples.
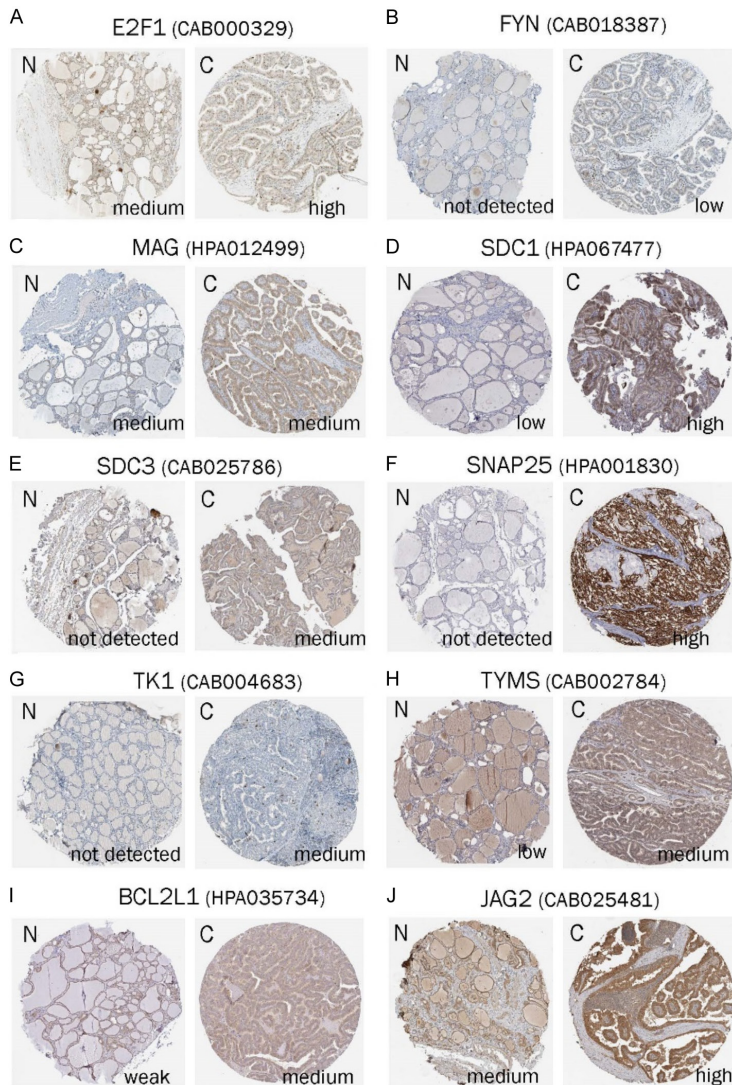

Figure 5. The Protein expression of hub genes in PTC obtained from the HPA online database. The staining intensity was classified into four categories: not detected, low, medium, and high. Image credit: Human Protein Atlas. Images were obtained from: v22. proteinatlas.org, via: https://www.proteinatlas.org/ENSG00000101412-E2F1/pathology/thyroid+cancer. https://www.proteinatlas.org/ENSG00000010810-FYN/pathology/thyroid+cancer. https://www.proteinatlas.org/ENSG0000010-5695-MAG/pathology/thyroid+cancer. https://www.proteinatlas.org/EN-SG00000115884-SDC1/pathology/thyroid+cancer. https://www.proteinatlas.org/ENSG00000162512-SDC3/pathology/thyroid+cancer. https://www.proteinatlas.org/ENSG00000132639-SNAP25/pathology/thyroid+cancer. https://www.proteinatlas.org/ENSG00000167900-TK1/pathology/thyroid+cancer. https://www.proteinatlas.org/ENSG00000176890-TYMS/pathology/thyroid+cancer. https://www.proteinatlas.org/ENSG00-000171552-BCL2L1/pathology/thyroid+cancer. https://www.proteinatlas.org/ENSG00000184916-JAG2/pathology/thyroid+cancer.

response to diverse apoptotic stimuli, including DNA damage, glucocorticoid treatment, and growth factor deprivation [49]. It is also downregulated with increased tumor size, independent of p53 expression, in head and neck cancer [50]. The expression of tyrosine kinase FYN, a member of the SRC family of kinases, is upregulated in THCA and promotes cell proliferation, invasion, and migration [51]. The malignancy-associated gene (MAG) is expressed in various tumors and preexisting conditions [52]. Dendritic cell 1 (DC1) is critical for mediating CD8+ T cell activation and antitumor immune responses via cross-presentation, co-stimulation, and soluble factors [53]. The expression of syndecan-3 is dysregulated in several cancer types and expressed in cancer cells and macrophages because of limited oxygenation in the tumor environment [54]. The expression of SNAP25 is significantly downregulated in prostate cancer and involved in the activation, differentiation, and migration of immune cells [55]. Among these hub genes, E2F1, TK1, and TYMS were significantly associated with unfavorable RFS in thyroid cancer patients. Interestingly, the trends of RFS and OS in the high-expression and low-expression groups, as depicted in **Figure 6**, exhibit complete opposition. Several factors can account for the contradictory findings between RFS and OS for a specific gene. First, the gene's role in different biological processes associated with RFS and OS could contribute to this discrepancy. If the gene significantly
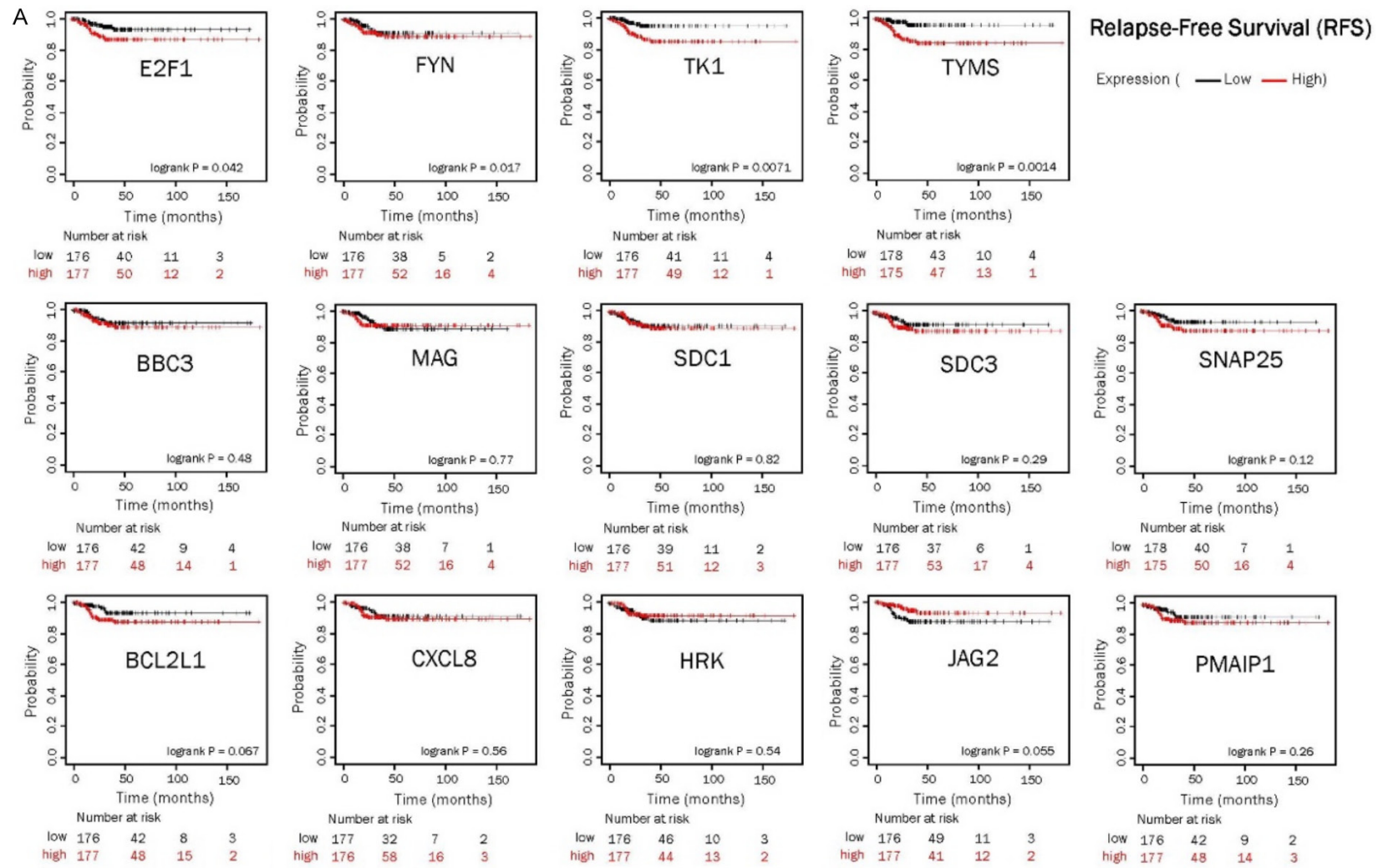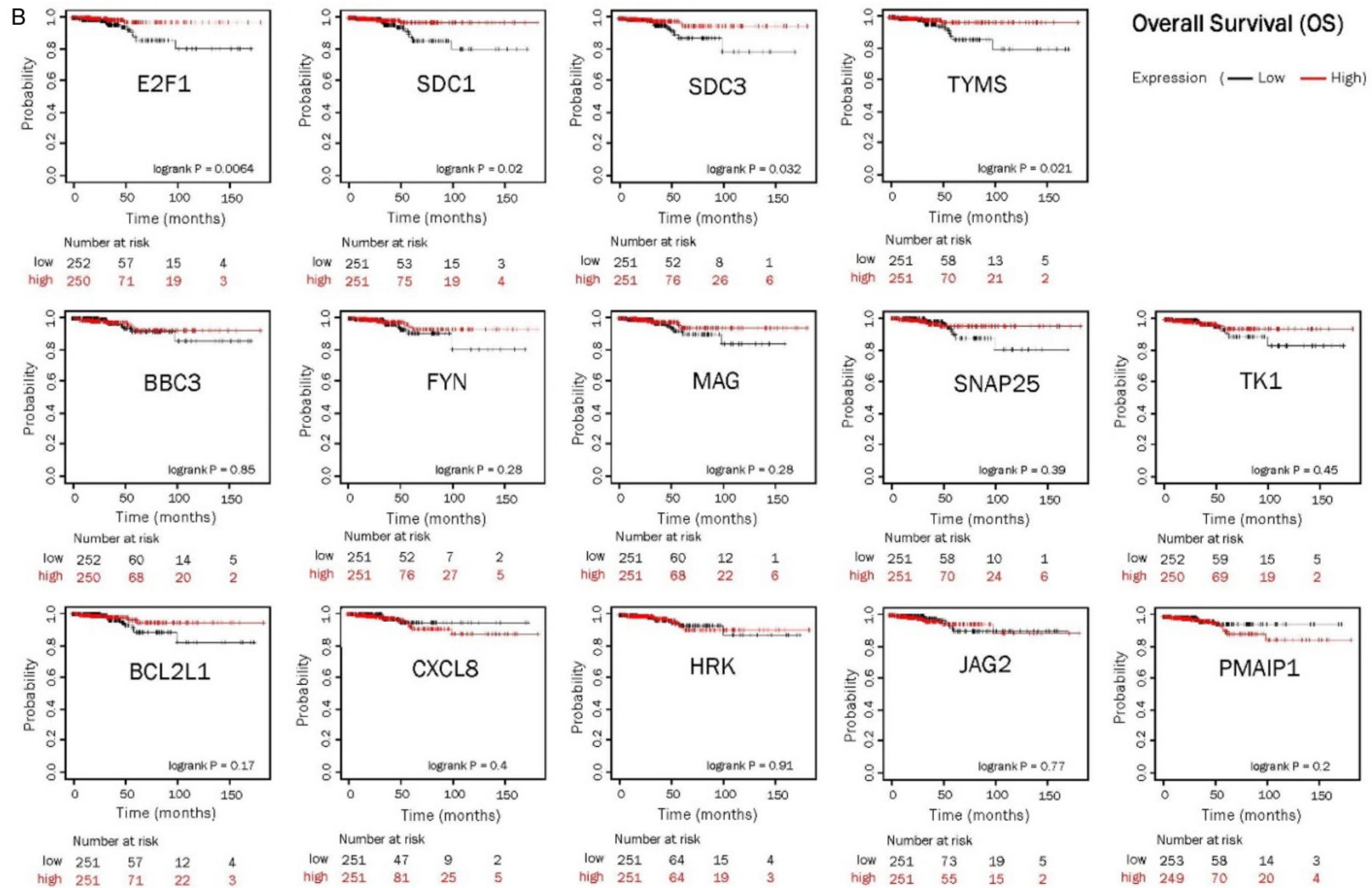
Figure 6. Prognostic effect of nine hub gene expression in thyroid cancer patients. A. The Kaplan-Meier plots illustrating the relapse-free survival (RFS) of patients with thyroid cancer were presented, with *P*-values and numbers at risk indicated on the plots. The low expression group is represented in black, while the high expression group is depicted in red, referring to the hub gene expression. B. The Kaplan-Meier plots illustrating the overall survival (OS) of patients with thyroid cancer were presented, with *P*-values and numbers at risk indicated on the plots. The low expression group is represented in black, while the high expression group is depicted in red, referring to the hub gene expression.

influences tumor growth, metastasis, immune response, and other factors, it can result in divergent RFS and OS outcomes. Second, the impact of other predictive factors needs to be considered. RFS and OS may reflect distinct predictive factors. For instance, the influence of gene expression levels on the likelihood of tumor recurrence might have a stronger effect on RFS compared to the overall survival period. Lastly, it is crucial to acknowledge the limitations of the data. Opposing results between gene expression levels and survival analysis are often associated with small sample sizes, data imbalances, incorrect data collection or analysis, and other factors. These limitations can affect the reliability of the findings, resulting in contradictory outcomes. Recent studies have documented that E2F1 is highly expressed in PTC cells and enhances their proliferation and invasion by upregulating LINC00152 and the PI3K/AKT axis [56]. Immunohistochemical analysis using an antibody against E2F1 revealed a prominent intracellular E2F1 protein in most primary papillary cancers (16 of 18; 89%). These data indicate that increased E2F1 expression might play a significant role in human thyroid carcinogenesis through derangement of the Rb-E2F signaling pathway [57]. FYN, a member of the SRC family, is increased in THCA and significantly associated with an increased THCA risk [51, 58]. TK1, a regulatory factor that modulates the cell cycle, is upregulated in the serum of patients with thyroid nodules. Knockdown of TK1 suppresses THCA cell proliferation, invasion, migration, and epithelial-mesenchymal transition and induces cell apoptosis [59]. High TYMS expression increases the viability and invasion of differentiated THCA cells [60]. Therefore, these hub genes, which are regulated by circRNAs, play important roles in PTC progression. The present study had certain limitations. Firstly, our selection of circRNAs was based solely on their large fold-change ratios. However, circRNAs with large differences may not necessarily be biologically significant. To address this, machine learning methods can be employed to consider the combined effects of multiple circRNAs and their interactions, enabling a more comprehensive understanding of the biological significance of differentially expressed circRNAs. Secondly, our analysis relied solely on microarray data and bioinformatics results, lacking corresponding in vitro and in vivo experiments. To further enhance our understanding, it would be advantageous to investigate two specific circRNAs, four miRNAs, and nine hub genes. Such an analysis could provide novel insights into the molecular mechanisms, potential therapeutic targets, and prognostic biomarkers for PTC.

In conclusion, we identified differentially expressed seven circRNAs between normal thyroid tissues and PTC using a public microarray dataset. Then, we established a potential circRNA/miRNA/mRNA regulatory axis that is linked to PTC progression by predicting interacting miRNAs and their target genes, PPI network analysis, and finally identifying hub genes.

Acknowledgements

All patients had signed a written informed consent to allow the use of their tumor tissue for research purposes.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Lark Kyun Kim, Department of Biomedical Sciences, Graduate School of Medical Science, Brain Korea 21 Project, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul 06230, Republic of Korea. Tel: 82-2-2019-5402; Fax: 82-2-2019-5210; E-mail: LKKIM@yuhs.ac

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185

countries. CA Cancer J Clin 2018; 68: 394-424.

[2] Lloyd RV, Osamura RY, Klöppel G and Rosai J. WHO classification of tumours of endocrine organs. International Agency for Research on Cancer (IARC); 2017.

[3] Dralle H, Machens A, Basa J, Fatourechi V, Franceschi S, Hay ID, Nikiforov YE, Pacini F, Pasieka JL and Sherman SI. Follicular cell-derived thyroid cancer. Nat Rev Dis Primers 2015; 1: 15077.

[4] Li Y, Che W, Yu Z, Zheng S, Xie S, Chen C, Qiao M and Lyu J. The incidence trend of papillary thyroid carcinoma in the United States during 2003-2017. Cancer Control 2022; 29: 10732748221135447.

[5] Daniels GH. Follicular thyroid carcinoma: a perspective. Thyroid 2018; 28: 1229-1242.

[6] Liu FH, Kuo SF, Hsueh C, Chao TC and Lin JD. Postoperative recurrence of papillary thyroid carcinoma with lymph node metastasis. J Surg Oncol 2015; 112: 149-154.

[7] Kim H, Kim TH, Choe JH, Kim JH, Kim JS, Oh YL, Hahn SY, Shin JH, Chi SA, Jung SH, Kim YN, Kim HI, Kim SW and Chung JH. Patterns of initial recurrence in completely resected papillary thyroid carcinoma. Thyroid 2017; 27: 908-914.

[8] Abdullah MI, Junit SM, Ng KL, Jayapalan JJ, Karikalan B and Hashim OH. Papillary thyroid cancer: genetic alterations and molecular biomarker investigations. Int J Med Sci 2019; 16: 450-460.

[9] Drozd V, Branovan DI and Reiners C. Increasing incidence of thyroid carcinoma: risk factors and seeking approaches for primary prevention. Int J Thyroidol 2020; 13: 95-110.

[10] Liyanarachchi S, Li W, Yan P, Bundschuh R, Brock P, Senter L, Ringel MD, de la Chapelle A and He H. Genome-wide expression screening discloses long noncoding RNAs involved in thyroid carcinogenesis. J Clin Endocrinol Metab 2016; 101: 4005-4013.

[11] Grillone K, Riillo C, Scionti F, Rocca R, Tradigo G, Guzzi PH, Alcaro S, Di Martino MT, Tagliaferri P and Tassone P. Non-coding RNAs in cancer: platforms and strategies for investigating the genomic "dark matter". J Exp Clin Cancer Res 2020; 39: 117.

[12] Anastasiadou E, Jacob LS and Slack FJ. Non-coding RNA networks in cancer. Nat Rev Cancer 2018; 18: 5-18.

[13] Cao J, Zhang M, Zhang L, Lou J, Zhou F and Fang M. Non-coding RNA in thyroid cancer - functions and mechanisms. Cancer Lett 2021; 496: 117-126.

[14] Condrat CE, Thompson DC, Barbu MG, Bugnar OL, Boboc A, Cretoiu D, Suciu N, Cretoiu SM and Voinea SC. miRNAs as biomarkers in disease: latest findings regarding their role in diagnosis and prognosis. Cells 2020; 9: 276.

[15] Peng Y and Croce CM. The role of MicroRNAs in human cancer. Signal Transduct Target Ther 2016; 1: 15004.

[16] Qian Y, Shi L and Luo Z. Long non-coding RNAs in cancer: implications for diagnosis, prognosis, and therapy. Front Med (Lausanne) 2020; 7: 612393.

[17] Huarte M. The emerging role of lncRNAs in cancer. Nat Med 2015; 21: 1253-1261.

[18] Van Roosbroeck K, Pollet J and Calin GA. miR-NAs and long noncoding RNAs as biomarkers in human diseases. Expert Rev Mol Diagn 2013; 13: 183-204.

[19] Wang S, Zhang K, Tan S, Xin J, Yuan Q, Xu H, Xu X, Liang Q, Christiani DC, Wang M, Liu L and Du M. Circular RNAs in body fluids as cancer biomarkers: the new frontier of liquid biopsies. Mol Cancer 2021; 20: 13.

[20] Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB and Kjems J. The biogenesis, biology and characterization of circular RNAs. Nat Rev Genet 2019; 20: 675-691.

[21] Hsu MT and Coca-Prados M. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. Nature 1979; 280: 339-340.

[22] Jeck WR and Sharpless NE. Detecting and characterizing circular RNAs. Nat Biotechnol 2014; 32: 453-461.

[23] Chen LL. The expanding regulatory mechanisms and cellular functions of circular RNAs. Nat Rev Mol Cell Biol 2020; 21: 475-490.

[24] Huang A, Zheng H, Wu Z, Chen M and Huang Y. Circular RNA-protein interactions: functions, mechanisms, and identification. Theranostics 2020; 10: 3503-3517.

[25] Liu J, Zhang X, Yan M and Li H. Emerging role of circular RNAs in cancer. Front Oncol 2020; 10: 663.

[26] Mitra A, Pfeifer K and Park KS. Circular RNAs and competing endogenous RNA (ceRNA) networks. Transl Cancer Res 2018; 7 Suppl 5: S624-S628.

[27] Xu X and Jing J. Advances on circRNAs contribute to carcinogenesis and progression in papillary thyroid carcinoma. Front Endocrinol (Lausanne) 2020; 11: 555243.

[28] Yao X and Zhang Q. Function and clinical significance of circular RNAs in thyroid cancer. Front Mol Biosci 2022; 9: 925389.

[29] Zhang D, Tao L, Xu N, Lu X, Wang J, He G, Tang Q, Huang K, Shen S and Chu J. CircRNA circTIAM1 promotes papillary thyroid cancer progression through the miR-646/HNRNPA1 signaling pathway. Cell Death Discov 2022; 8: 21.

[30] Chu J, Tao L, Yao T, Chen Z, Lu X, Gao L, Fang L, Chen J, He G, Shen S and Zhang D. Circular RNA circRUNX1 promotes papillary thyroid cancer progression and metastasis by sponging
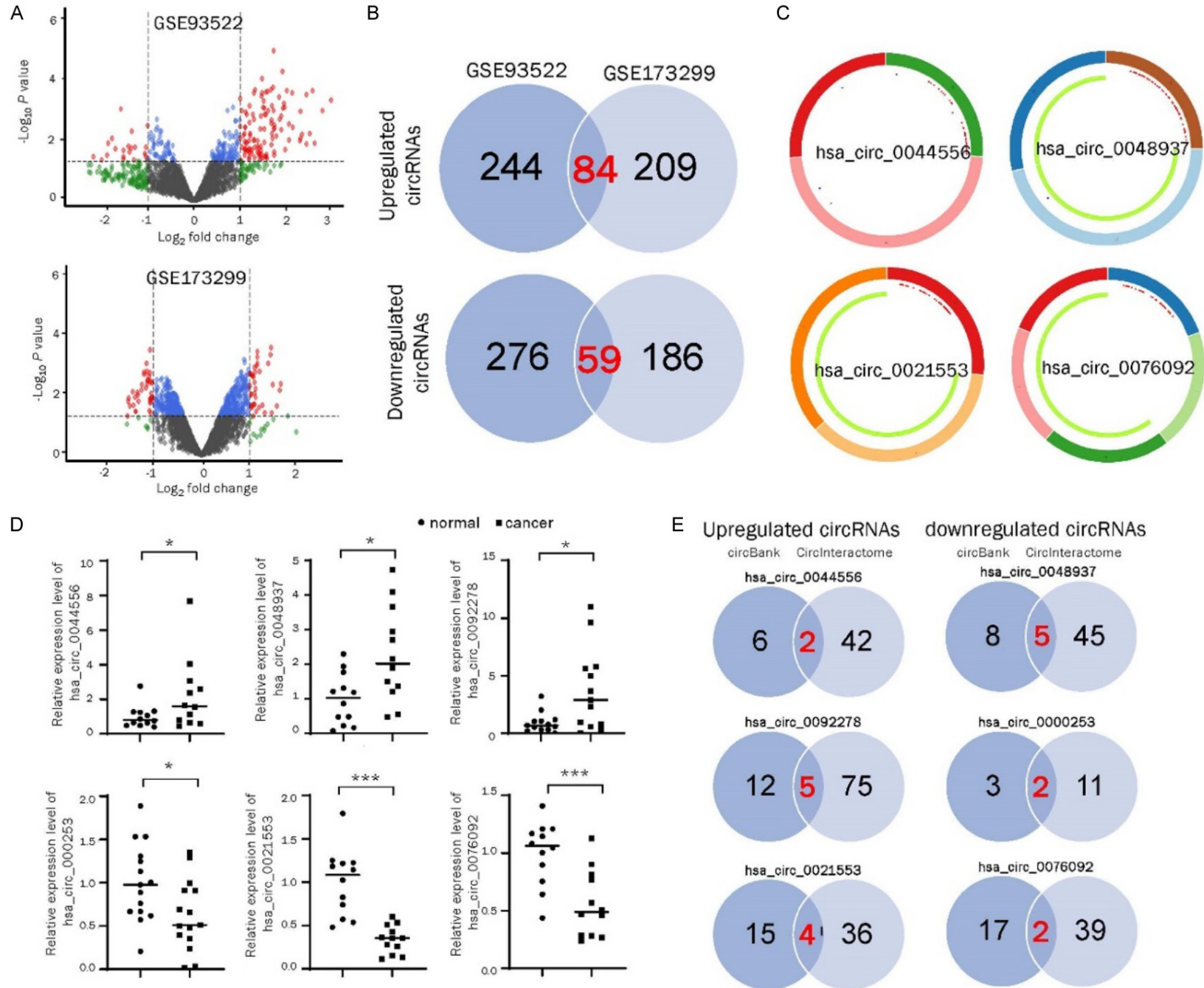
MiR-296-3p and regulating DDHD2 expression. Cell Death Dis 2021; 12: 112.

[31] Ding W, Shi Y and Zhang H. Circular RNA circ-NEURL4 inhibits cell proliferation and invasion of papillary thyroid carcinoma by sponging miR-1278 and regulating LATS1 expression. Am J Transl Res 2021; 13: 5911-5927.

[32] Peng N, Shi L, Zhang Q, Hu Y, Wang N and Ye H. Microarray profiling of circular RNAs in human papillary thyroid carcinoma. PLoS One 2017; 12: e0170287.

[33] Guo M, Sun Y, Ding J, Li Y, Yang S, Zhao Y, Jin X and Li SS. Circular RNA profiling reveals a potential role of hsa_circ_IPCEF1 in papillary thyroid carcinoma. Mol Med Rep 2021; 24: 603.

[34] Lv C, Sun W, Huang J, Qin Y, Ji X and Zhang H. Expression profiles of circular RNAs in human papillary thyroid carcinoma based on RNA deep sequencing. Onco Targets Ther 2021; 14: 3821-3832.

[35] Jombart T. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 2008; 24: 1403-1405.

[36] Love MI, Huber W and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15: 550.

[37] Dudekula DB, Panda AC, Grammatikakis I, De S, Abdelmohsen K and Gorospe M. CircInteractome: a web tool for exploring circular RNAs and their interacting proteins and microRNAs. RNA Biol 2016; 13: 34-42.

[38] Liu M, Wang Q, Shen J, Yang BB and Ding X. Circbank: a comprehensive database for circRNA with standard nomenclature. RNA Biol 2019; 16: 899-905.

[39] Xia S, Feng J, Chen K, Ma Y, Gong J, Cai F, Jin Y, Gao Y, Xia L, Chang H, Wei L, Han L and He C. CSCD: a database for cancer-specific circular RNAs. Nucleic Acids Res 2018; 46: D925-D929.

[40] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.

[41] Nagy Á, Munkácsy G and Győrffy B. Pancancer survival analysis of cancer hallmark genes. Sci Rep 2021; 11: 6047.

[42] Huang da W, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4: 44-57.

[43] Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ and von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks,

made broadly accessible. Nucleic Acids Res 2017; 45: D362-D368.

[44] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.

[45] Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014; 8 Suppl 4: S11.

[46] Kristensen LS, Jakobsen T, Hager H and Kjems J. The emerging roles of circRNAs in cancer and oncology. Nat Rev Clin Oncol 2022; 19: 188-206.

[47] Tang X, Ren H, Guo M, Qian J, Yang Y and Gu C. Review on circular RNAs and new insights into their roles in cancer. Comput Struct Biotechnol J 2021; 19: 910-928.

[48] Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J and Pontén F. Proteomics. Tissue-based map of the human proteome. Science 2015; 347: 1260419.

[49] Han J, Flemington C, Houghton AB, Gu Z, Zambetti GP, Lutz RJ, Zhu L and Chittenden T. Expression of bbc3, a pro-apoptotic BH3-only gene, is regulated by diverse cell death and survival signals. Proc Natl Acad Sci U S A 2001; 98: 11318-11323.

[50] Tajnik M, Stražišar M, Volavšek M, Boštjančič E and Glavač D. BBC3 is down-regulated with increased tumor size independently of p53 expression in head and neck cancer. Cancer Biomark 2012; 11: 197-208.

[51] Zheng J, Li H, Xu D and Zhu H. Upregulation of tyrosine kinase FYN in human thyroid carcinoma: role in modulating tumor cell proliferation, invasion, and migration. Cancer Biother Radiopharm 2017; 32: 320-326.

[52] Ljubimova JY, Wilson SE, Petrovic LM, Ehrenman K, Ljubimov AV, Demetriou AA, Geller SA and Black KL. Novel human malignancy-associated gene (MAG) expressed in various tumors and in some tumor preexisting conditions. Cancer Res 1998; 58: 4475-4479.

[53] Kim CW, Kim KD and Lee HK. The role of dendritic cells in tumor microenvironments and their uses as therapeutic targets. BMB Rep 2021; 54: 31-43.
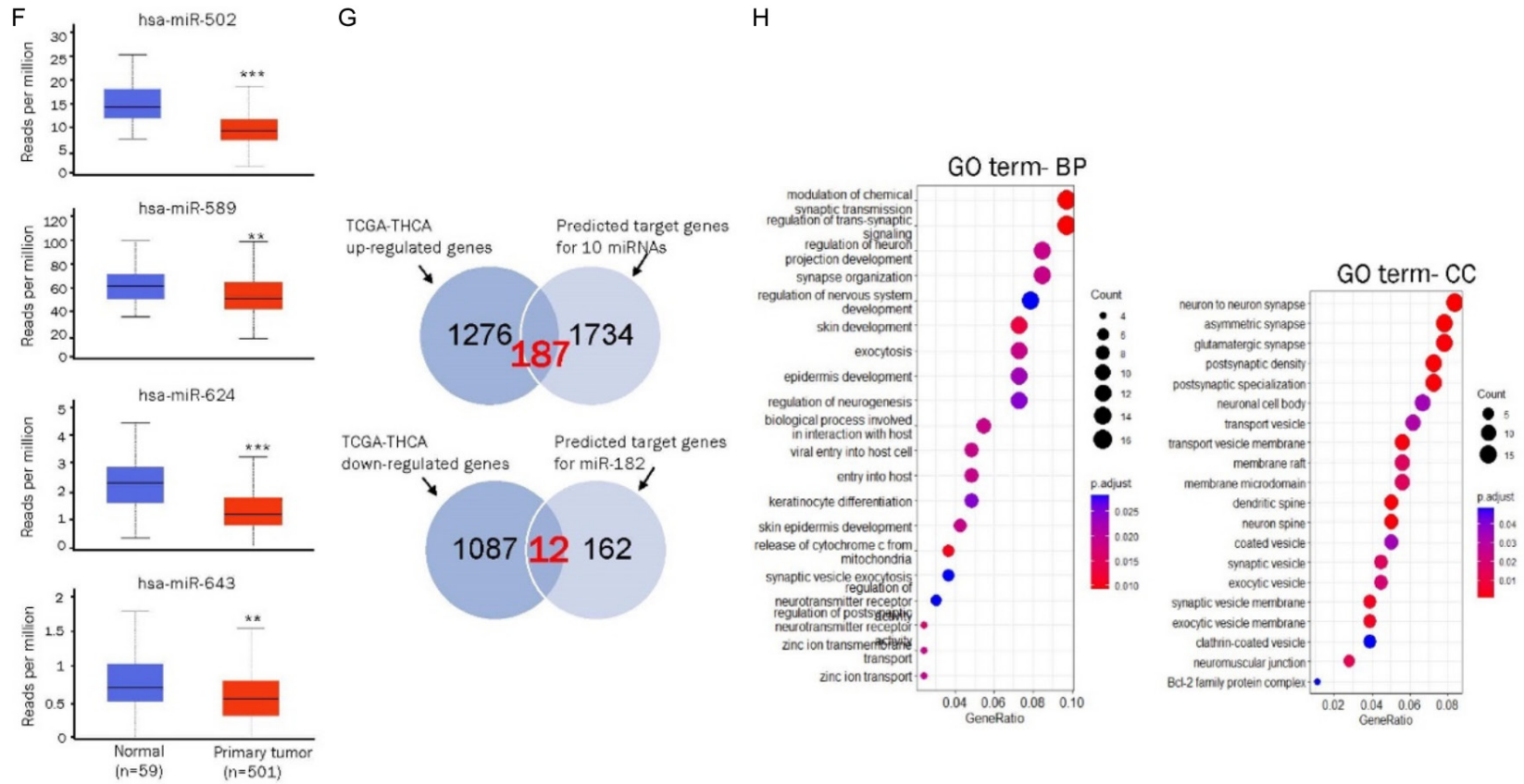
[54] Prieto-Fernández E, Egia-Mendikute L, Bosch A, García Del Río A, Jimenez-Lasheras B, Antoñana-Vildosola A, Lee SY and Palazon A. Hypoxia promotes syndecan-3 expression in the tumor microenvironment. Front Immunol 2020; 11: 586977.

[55] Di L, Gu M, Wu Y, Liu G, Zhang L, Li Y and Zhang W. SNAP25 is a potential prognostic biomarker for prostate cancer. Cancer Cell Int 2022; 22: 144.

[56] Yang J, Ying Y, Zeng X, Liu J, Xie Y, Deng Z, Hu Z and Li Z. Transcription factor E2F1 exacerbates papillary thyroid carcinoma cell growth and invasion via upregulation of LINC00152. Anal Cell Pathol (Amst) 2022; 2022: 7081611.

[57] Onda M, Nagai H, Yoshida A, Miyamoto S, Asaka SI, Akaishi J, Takatsu K, Nagahama M, Ito K, Shimizu K and Emi M. Up-regulation of transcriptional factor E2F1 in papillary and anaplastic thyroid cancers. J Hum Genet 2004; 49: 312-318.

[58] Nisar A, Kayani MA, Nasir W, Mehmood A, Ahmed MW, Parvez A and Mahjabeen I. Fyn and Lyn gene polymorphisms impact the risk of thyroid cancer. Mol Genet Genomics 2022; 297: 1649-1659.

[59] Liu C, Wang J, Zhao L, He H, Zhao P, Peng Z, Liu F, Chen J, Wu W, Wang G and Dong F. Knockdown of thymidine kinase 1 suppresses cell proliferation, invasion, migration, and epithelial-mesenchymal transition in thyroid carcinoma cells. Front Oncol 2020; 9: 1475.

[60] Wu M, Ou-Yang DJ, Wei B, Chen P, Shi QM, Tan HL, Huang BQ, Liu M, Qin ZE, Li N, Hu HY, Huang P and Chang S. A prognostic model of differentiated thyroid cancer based on upregulated glycolysis-related genes. Front Endocrinol (Lausanne) 2022; 13: 775278.

circRNA regulatory network in PTC

A GSE93522 / GSE173299 volcano plots

B Upregulated circRNAs Venn: 244 | 84 | 209; Downregulated circRNAs Venn: 276 | 59 | 186

C hsa_circ_0044556, hsa_circ_0048937, hsa_circ_0021553, hsa_circ_0076092

D Relative expression levels (normal ● vs cancer ■): hsa_circ_0044556 (*), hsa_circ_0048937 (*), hsa_circ_0092278 (*), hsa_circ_000253 (*), hsa_circ_0021553 (***), hsa_circ_0076092 (***)

E Upregulated circRNAs / downregulated circRNAs (circBank / CircInteractome):
hsa_circ_0044556: 6 | 2 | 42; hsa_circ_0048937: 8 | 5 | 45;
hsa_circ_0092278: 12 | 5 | 75; hsa_circ_0000253: 3 | 2 | 11;
hsa_circ_0021553: 15 | 4 | 36; hsa_circ_0076092: 17 | 2 | 39

1

Supplementary Figure 1. Data normalization for the analysis of the GEO dataset and identification of DECs in thyroid cancer. A. Identification of DECs in two GEO datasets. Volcano plots depicting the DECs obtained from normalized GEO datasets, where red dots denote significantly upregulated and downregulated circRNAs. A threshold of P < 0.05 and fold change > 1.5 were set. B. Venn diagram showing 84 overlapping upregulated DECs (upper) and 59 overlapping downregulated DECs (lower) between the two GEO datasets. C. Structural patterns of the four circRNAs by the CSCD. D. qRT-PCR validation of six differentially expressed circRNAs. The relative expression levels of circRNAs in 12 PTC tissues and paired normal tissues. Notably, circ_0000253 was analyzed in 15 pairs enhancing the comprehensive assessment of expression patterns; *P < 0.05, ***P < 0.001 (•: normal tissues, •: PTC tissues). E. Venn diagram between the seven target miRNAs predicted by CircBank and CircInteractome. F. TCGA expression profiles for the four miRNAs in thyroid cancer samples (n=501, red) and normal samples (n=59, blue). G. Venn diagram represent the overlap between miRNA target genes and DEGs from TGCA. Upper; 187 mRNAs were obtained from the intersection of 1,463 upregulated mRNAs (TCGA-THCA dataset) and 1,921 miRNA targets predicted using TargetScan. Lower; 12 mRNAs were obtained from the intersection of 1,099 downregulated mRNAs (TCGA-THCA dataset) and 174 miRNA targets predicted using TargetScan. H. Visualizing Functional Enrichment Analysis Results for DEGs using Dot Plots. Dot size corresponds to gene count within a pathway, while dot color indicates the *P*-value of the pathway. "BP" signifies Biological Process, and "CC" signifies Cellular Component.
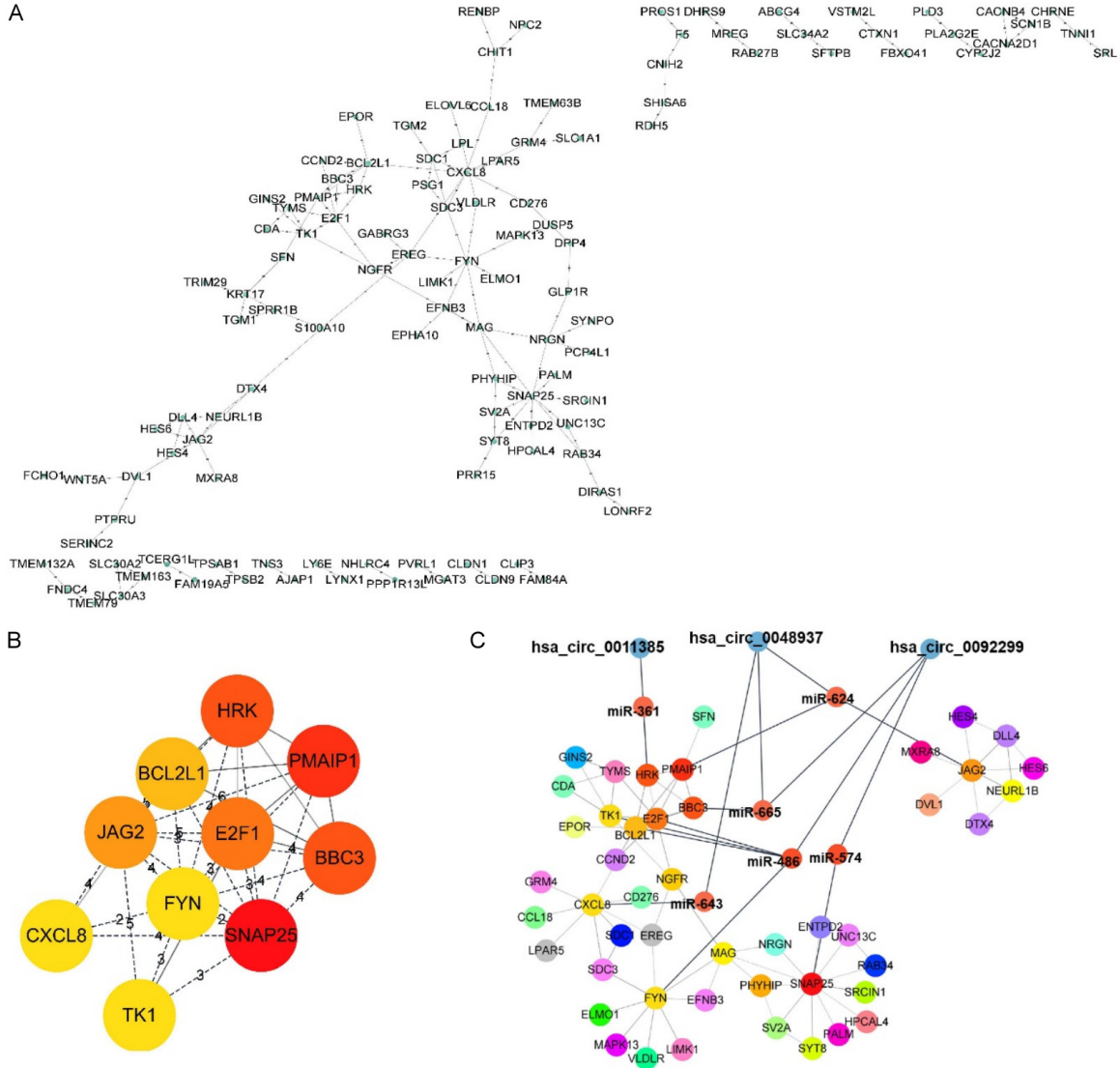
Supplementary Figure 2. Visualization of the PPI network and the hub genes. A. PPI network of 187 overlapped genes. The nodes represent the genes, and edges indicate interaction associations between nodes. B. Identification of the hub genes from the PPI network. The genes with the top 10 MCC values were considered hub genes. Edges represent the protein-protein associations. Red to yellow color gradients indicate the higher MCC ranking of hub genes. Numbers represent path length. C. circRNA-miRNA-hub gene subnetwork. Red-orange-yellow represents hub genes.