



Research article

Five genes involved in circular RNA-associated competitive endogenous RNA network correlates with metastasis in papillary thyroid carcinoma

Jie Qiu^{1,†}, Maolin Sun^{1,†}, Chuanshan Zang¹, Liwei Jiang¹, Zuorong Qin¹, Yan Sun¹, Mingbo Liu^{2,*} and Wenwei Zhang^{3,*}

¹ Department of Otolaryngology, The Affiliated Hospital of Qingdao University, Qingdao 266003, China

² Department of Otolaryngology, Hainan Hospital of PLA General Hospital, Sanya 572000, China

³ Radiology Department, The Affiliated Hospital of Qingdao University, Qingdao 266003, China

* **Correspondence:** Tel: +86053282913093; E-mail: zhao57665765183@163.com, Mingbo666@vip.163.com.

† These authors contributed equally to this work.

Abstract: This study aimed to identify potential circular RNA (circRNA), microRNA (miRNA) and mRNA biomarkers as well as their underlying regulatory mechanisms in papillary thyroid carcinoma (PTC). Three microarray datasets from the Gene Expression Omnibus database as well as expression data and clinical phenotype from The Cancer Genome Atlas (TCGA) were downloaded, followed by differential expression, functional enrichment, protein–protein interaction (PPI), and module analyses. The support vector machine (SVM)-recursive feature elimination (RFE) algorithm was used to screen the key circRNAs. Finally, the mRNA-miRNA-circRNA regulatory network and competitive endogenous RNA (ceRNA) network were constructed. The prognostic value and clinical correlations of key mRNAs were investigated using TCGA dataset, and their expression was validated using the UALCAN database. A total of 1039 mRNAs, 18 miRNAs and 137 circRNAs were differentially expressed in patients with PTC. A total of 37 key circRNAs were obtained using the SVM-RFE algorithm, whereas 46 key mRNAs were obtained from significant modules in the PPI network. A total of 11 circRNA-miRNA pairs and 40 miRNA-mRNA pairs were predicted. Based on these interaction pairs, 46 circRNA-miRNA-mRNA regulatory pairs were integrated, of which 8 regulatory pairs in line with the ceRNA hypothesis were obtained, including two circRNAs

(circ_0004053 and circ_0028198), three miRNAs (miR-199a-5p, miR-199b-5p, and miR-7-5p), and five mRNAs, namely *APOA2*, *CCL20*, *LPAR5*, *MFGE8*, and *TIMP1*. Survival analysis showed that *LPAR5* expression was associated with patient survival. *APOA2* expression showed significant differences between metastatic and non-metastatic tumors, whereas *CCL20*, *LPAR5*, *MFGE8* and *TIMP1* showed significant differences between metastatic and non-metastatic lymph nodes. Overall, we identified several potential targets and regulatory mechanisms involved in PTC. *APOA2*, *CCL20*, *LPAR5*, *MFGE8*, and *TIMP1* may be correlated with PTC metastasis.

Keywords: papillary thyroid carcinoma; circular RNAs; ceRNA; metastasis

1. Introduction

Papillary thyroid carcinoma (PTC) is the most common malignant thyroid tumor, accounting for 80–90% of thyroid cancers. PTC is often characterized to have an early metastasis to the lymph node of the neck [1,2]. Although the prognosis of most patients with PTC is good, special histological subtypes of papillary carcinoma, iodine-resistant PTC, thyroid invasion outside the membrane, vascular invasion, and distant metastasis are often considered risk factors for postoperative PTC recurrence, leading to relatively poor prognosis [3,4]. At present, the molecular mechanism of the occurrence and development of PTC has not been clarified; therefore, an increasing number of studies are being performed to identify a molecular marker that can predict the occurrence, metastasis, and recurrence of PTC.

Circular RNA (circRNA) is a type of non-coding RNA that was first discovered in viroids in plants in 1976. circRNAs are single-stranded and covalently closed RNAs that lack a free cap end and poly (A) tail [5,6] and are widely distributed in mammalian cells [7]. circRNAs are stabilized in cells [8] and act as miRNA sponges to competitively bind to mRNA [9]. In recent years, the correlation between circRNAs and the occurrence and development of malignant tumors has become a focus of academic research. For example, the correlation between circRNA and the occurrence and development of bladder cancer [10], correlation between circRNA and the early diagnosis of breast adenocarcinoma [11], correlation between circRNA and the occurrence and development of colon cancer [12], and correlation between circRNA and the proliferation of liver cancer cells [13] have been discovered. Hence, the relationship between circRNAs and human malignancies has been evaluated and established; however, the diagnostic value and biological function of circRNAs in PTC have not been extensively studied.

In this study, microarray datasets for PTC and normal samples were downloaded from a public database. Differentially expressed mRNAs (dif-mRNAs), differentially expressed miRNAs (dif-miRNAs), and differentially expressed circRNAs (dif-circRNAs) between PTC and control tissues were investigated, followed by an enrichment analysis of the dif-mRNAs. To further explore the mechanism of key mRNAs, circRNAs, and miRNAs in PTC, protein–protein interaction (PPI) network and regulatory network analyses were performed. Therefore, this study provides novel therapeutic targets for PTC.

2. Materials and methods

2.1. Data sources

The microarray datasets GSE93522 (platform: Agilent-069978 Arraystar Human CircRNA microarray V1), GSE73182 (platform: Agilent-035758 Human miRBASE 16.0 plus 031181), and GSE66783 (GPL19850 Agilent-060228 Human LncRNA v5 4X180K [Probe Name Version]) from the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) database were downloaded. The dataset GSE93522 includes circRNA expression data of 6 PTC and 6 paired adjacent noncancerous thyroid tissues, GSE73182 includes miRNA expression data of 19 primary PTC and 5 normal thyroids, and GSE66783 includes mRNA expression data of 5 PTC and 5 paired adjacent noncancerous thyroid tissues.

In addition, three microarray datasets were downloaded for external validation, including GSE173299 (circRNA expression data of three PTC and three normal samples), GSE104006 (miRNAs and mRNA expression data of 24 PTC and 5 normal samples), and GSE151181 (miRNA and mRNA expression data of 39 PTC and 13 normal samples).

2.2. Differential expression analysis

The original data were read using the limma package (Version 3.10.3, <http://www.bioconductor.org/packages/2.9/bioc/html/limma.html>) [14] and preprocessed using the robust multi-array average (RMA). Data preprocessing included background correction, normalization, and expression calculations. Gene expression values were calculated by mapping probes to symbols using the microarray dataset and annotation files from the chip platform. The probes were removed when no genes were matched. The mean value was selected as the expression value when multiple probes matched one gene symbol. The dif-mRNAs, dif-miRNAs, and dif-circRNAs between samples from patients with PTC and control tissues were analyzed using a typical Bayesian method. The dif-miRNAs and dif-circRNAs were defined as miRNAs or circRNAs with $|\log_2(\text{FC})|$ greater than 1 and a P-value less than 0.05. The dif-mRNAs were defined as mRNAs with $|\log_2(\text{FC})|$ greater than 2 and a P-value less than 0.05.

2.3. Pathway enrichment analysis of dif-mRNAs

The biological process (BP) terms of the Gene Ontology (GO) [15] annotation and the Kyoto Encyclopedia of Genes and Genomes (KEGG) [16] pathways were enriched for the upregulated and downregulated overlapping dif-mRNAs using the enrichment analysis tool DAVID. The number of enriched genes was set as $\text{count} \geq 2$. Statistical significance was set at $P < 0.05$. The REViGO online tool was utilized to remove redundancy among the GO-BP terms, with the following parameter settings: allowed similarity, small (0.5); species, *Homo sapiens*; and semantic similarity measure, Lin.

2.4. PPI network and module analyses for dif-mRNAs

Using the STRING database (version: 10.0, <http://www.string-db.org/>) [17], the interactions between proteins encoded by dif-mRNAs were retrieved. The gene set consisted of all differentially expressed genes, and humans were used as the species. A PPI score of 0.7 was defined as having high confidence. Cytoscape software [18] (version: 3.2.0, <http://www.cytoscape.org>) was used to

construct the PPI network based on the obtained interactions. Significant network modules were analyzed using the Molecular Complex Detection (MCODE) plug-in (version: 1.4.2, <http://apps.cytoscape.org/apps/MCODE>) [19] in the Cytoscape software. Modules with scores ≥ 10 were considered significant. The BP terms of the GO annotation and KEGG pathways were enriched for the genes in the significant clustering modules. The number of enriched genes was set as count ≥ 2 . Statistical significance was set at $P < 0.05$.

2.5. Screening of key dif-circRNAs

The support vector machine (SVM)-recursive feature elimination (RFE) algorithm was used to screen the key dif-circRNAs. In brief, five-fold cross-validation was conducted for all the dif-circRNAs using the Caret package (version 6.0-86). Thereafter, SVM-RFE was performed for dimensionality reduction screening of dif-circRNAs, and the circRNAs contained in the model with the lowest RMSE (standard error) were selected as the key dif-circRNAs.

2.6. Construction of a competing endogenous RNA (ceRNA) network

The circRNA-miRNA interaction pairs and miRNA-mRNA interaction pairs were predicted for all dif-miRNAs by utilizing the Encyclopedia of RNA Interactomes (ENCORI, previous version: starBase v2.0, <http://starbase.sysu.edu.cn/>) with the following parameters: CLIP-Data ≥ 1 , pan-Cancer ≥ 0 and Degrado-me-Data ≥ 0 . The circRNA-miRNA and miRNA-mRNA interaction pairs were first screened using the key dif-circRNAs (obtained from SVM-RFE analysis) and mRNAs in PPI modules; subsequently, the circRNA-miRNA and miRNA-mRNA interaction pairs that involved the same miRNAs were retained. Based on the ceRNA hypothesis, two RNA molecules regulated by the same miRNA are co-expressed, and an miRNA usually inhibits the expression of its target gene (mRNA). Therefore, positive circRNA-mRNA interaction pairs and negative miRNA-mRNA interaction pairs were selected to construct ceRNA regulatory pairs.

2.7. Associations of key mRNAs with survival and clinical features

The expression data and clinical phenotypes of thyroid cancer were downloaded from TCGA database, and only the TPC samples with clinical phenotypes were selected. A total of 336 samples were obtained. The expression of key mRNAs (mRNAs in the ceRNA network) was extracted, and samples were assigned into high- and low-expression groups based on their median expression value. Kaplan-Meier survival analysis was performed to screen for prognostic mRNAs. In addition, associations between the expression of key mRNAs with clinical phenotype were analyzed, including age, sex, pathologic T/N/M stages, primary thyroid gland neoplasm location anatomic site, and tumor stage. The *t*-test was used for comparison between groups.

2.8. Validation of the expression of key RNAs

The expression of key circRNAs, miRNAs, and mRNAs between PTC and normal tissues was validated using external validation datasets and the UALCAN online tool. The *t*-test was used for comparison between groups.

3. Results

3.1. Differential expression analysis

According to the screening criteria, there were 137 dif-circRNAs (115 upregulated and 22 downregulated), 18 dif-miRNAs (9 upregulated and 9 downregulated), and 1039 dif-mRNAs (454 upregulated and 585 downregulated) between PTC and normal tissues (Figure 1A,B). Principal component analysis showed that samples in the different groups could be clearly distinguished (Figure 1C).

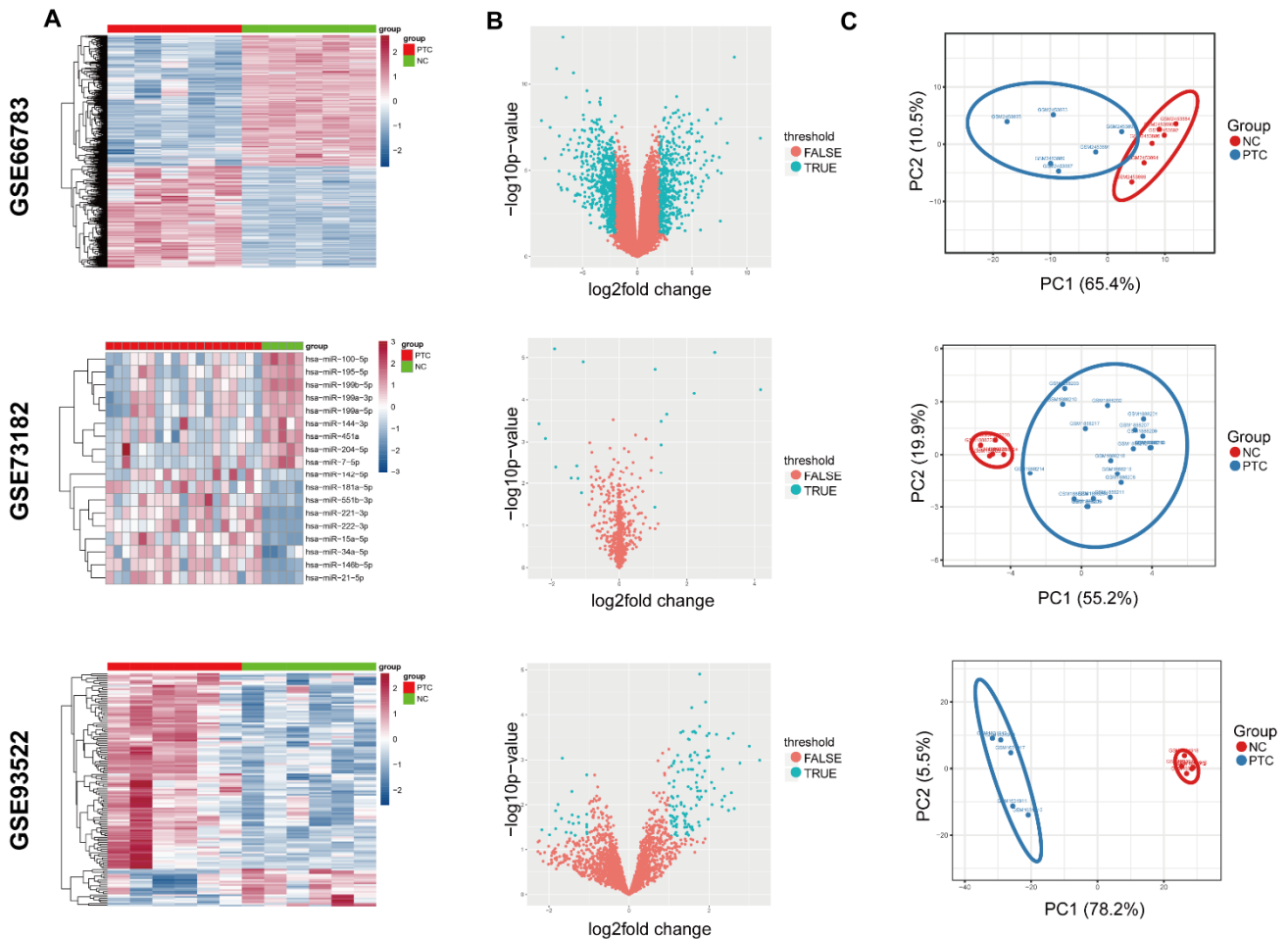


Figure 1. Differential expression analysis. Heatmap (A), volcano plot (B) and principal component analysis (C) of differentially expressed circRNAs (GSE93522), miRNAs (GSE73182) and mRNAs (GSE66783). The X-axis in the heatmaps represents the samples (green blocks represent normal samples, and red blocks represent tumor samples), and Y-axis represents the differentially expressed RNAs. The dots in the volcano plot represent RNAs; blue dots represent differentially expressed RNAs, and red dots represent RNAs without significant expression changes; the X-axis and Y-axis represent expression fold changes and P-values, respectively. Dots in the principal component analysis plot represent individual samples; red dots represent the normal samples, and blue dots represent the tumor samples.

3.2. Enrichment analysis of dif-mRNAs

The functions of dif-mRNAs were explored using enrichment analysis. The upregulated and downregulated mRNAs were enriched in 81 and 114 GO-BP terms, respectively. The significantly enriched GO-BP terms obtained after removing redundancy among the GO-BP terms using REViGO are shown in Figures 2A,B. It could be seen that the upregulated mRNAs were mainly involved in extracellular matrix disassembly, collagen catabolic process, leukocyte migration, regulation of cell adhesion mediated by integrin, and others (Figure 2A). The downregulated mRNAs were mainly implicated in negative regulation of growth, cell adhesion, cellular response to cadmium ions, among others (Figure 2B). Additionally, the significantly enriched KEGG pathways are shown in Figure 2C,D. The upregulated mRNAs were mainly involved in transcriptional misregulation in cancer, p53 signaling pathway, and cytokine–cytokine receptor interaction (Figure 2C) while the downregulated mRNAs were mainly involved in tyrosine/retinol metabolism, drug metabolism-cytochrome P450, TGF–beta signaling pathway, and the Hippo signaling pathway (Figure 2D).

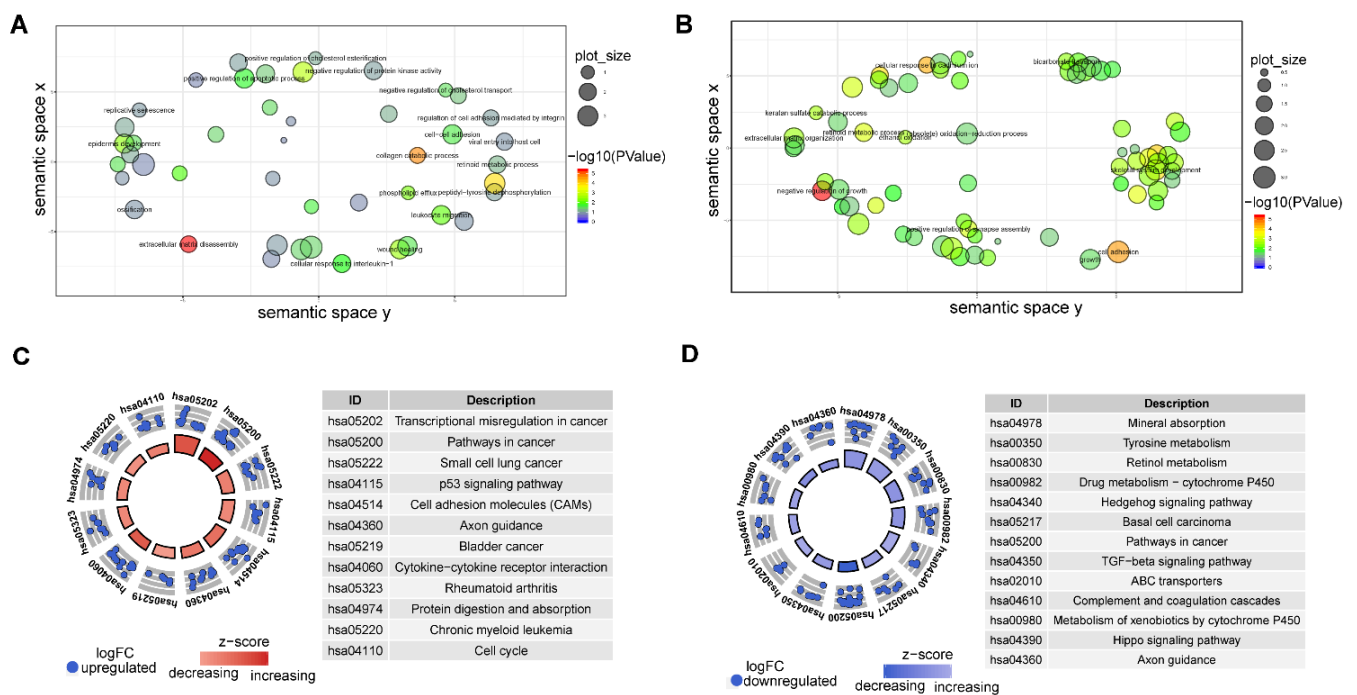


Figure 2. Functional enrichment analysis. Significantly enriched biological process terms of the Gene Ontology annotations after eliminating redundancy for upregulated (A) and downregulated (B) mRNAs. The bubble sizes indicate the number of background genes in the pathway and whether the pathway is general or special. The closer the bubbles are, the more similar are the pathways. The significantly enriched KEGG pathways for upregulated (C) and downregulated (D) mRNAs. The dots in the outer ring represent genes, and the dots in each gray block represent the number of genes enriched in the particular pathway. The colored blocks in the inner ring represent the z-score of each pathway. The colors from light to dark represent statistical significance, where a darker color corresponds to a higher statistical significance.

3.3. PPI network and module analyses

The PPI network of dif-mRNAs comprised 361 nodes and 805 interactions (Figure 3A). *FN1* (degree = 28), *APOA1* (degree = 23), *SERPINA1* (degree = 23), *TIMP1* (degree = 22), *APOE* (degree = 22), *GPC3* (degree = 21), *APOA2* (degree = 20), *NMU* (degree = 20), and *LPAR5* (degree = 20) were hub nodes with the highest degrees in the network. Three significant modules were identified separately from the PPI network (Figure 3B). Module A had 15 nodes and 105 interactions, of which *FN1*, *APOA1*, *SERPINA1*, *APOE*, *TIMP1*, *GPC3* and *APOA2* were also hub nodes. The mRNAs in module A were mainly enriched in GO-BP terms such as high-density lipoprotein particle assembly/clearance, positive regulation of cholesterol esterification, and lipoprotein biosynthetic process and KEGG pathways such as ECM–receptor interaction and focal adhesion. Module B had 21 nodes and 122 interactions, and the mRNAs in this module were mainly involved in GO-BP terms such as G-protein coupled receptor signaling pathway, positive regulation of ERK1 and ERK2 cascade, and cell–cell signaling and in KEGG pathways such as calcium signaling pathway and cytokine–cytokine receptor interaction. Module C had 10 nodes and 45 interactions, and the mRNAs in this module were mainly implicated in 4 GO-BP terms, namely negative regulation of viral genome replication, antibacterial humoral response, response to mechanical stimulus, and cellular response to interleukin-1 (Table 1).

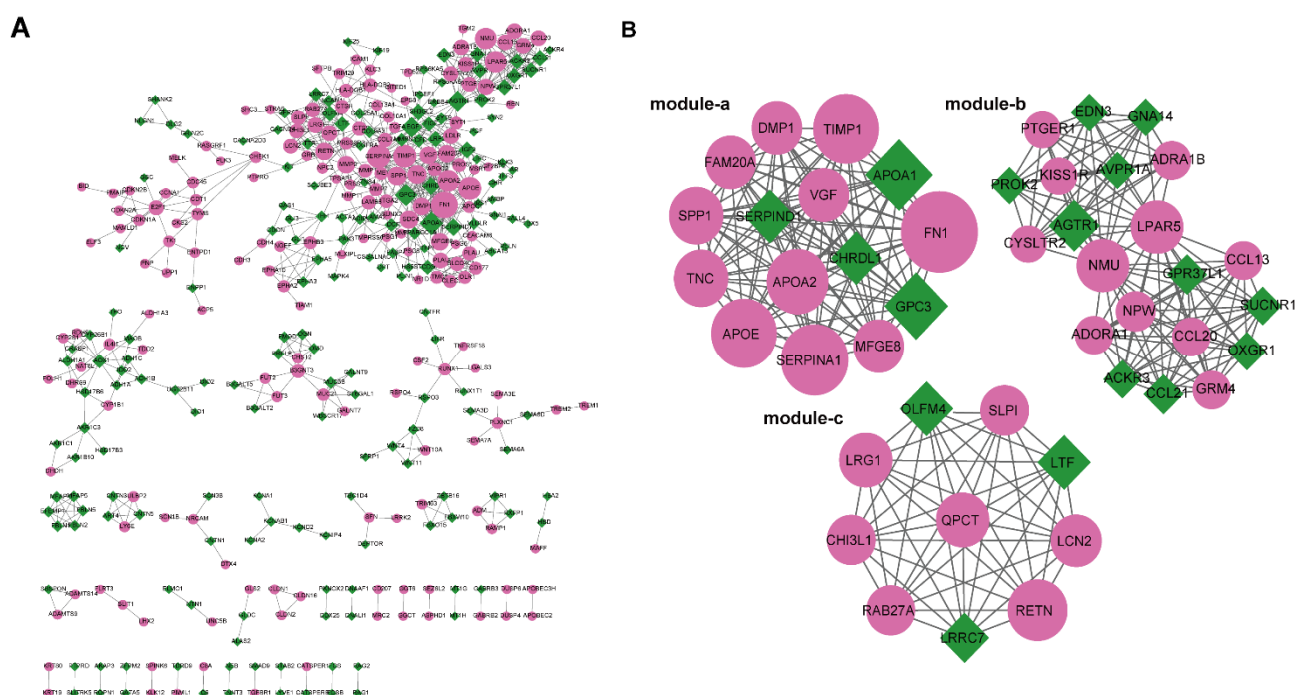


Figure 3. Protein–protein interaction (PPI) network and significant modules. (A) PPI interactions among dif-mRNAs were analyzed using the STRING database, and the PPI network was visualized using Cytoscape. (B) Three significant modules with scores ≥ 10 were identified from the PPI network using the MCODE plug-in. Pink circles represent upregulated mRNAs, and green diamonds represent downregulated mRNAs. The lines between two nodes represent interactions. The node size and degree are proportional, and a large node size corresponds to a high degree in the network.

3.4. Key dif-circRNAs screened via SVM-RFE

The expression of 137 dif-circRNAs was extracted, and SVM-RFE was utilized for the dimensionality reduction screening of dif-circRNAs. As described in the Methods section, the circRNAs with the lowest RMSE (standard error) contained in the model were selected as the key dif-circRNAs, and consequently, 37 dif-circRNAs were obtained (Figure S1, Table S1).

Table 1. KEGG pathways and the top five GO-BP terms of the mRNAs in significant modules.

Module gene KEGG pathway	Category	Term	Count	P-Value
	Module A	hsa04512: ECM–receptor interaction	3	5.37E–03
		hsa04510: Focal adhesion	3	2.80E–02
	Module B	hsa04080: Neuroactive ligand-receptor interaction	8	4.29E–07
		hsa04020: Calcium signaling pathway	6	1.87E–05
		hsa04060: Cytokine–cytokine receptor interaction	4	1.19E–02
		hsa04270: Vascular smooth muscle contraction	3	2.28E–02
		hsa04022: cGMP–PKG signaling pathway	3	3.98E–02
Module gene GO-BP (top 5)	Category	Term	Count	P-Value
Module A		GO:0034384~high-density lipoprotein particle clearance	3	9.66E–06
		GO:0001523~retinoid metabolic process	4	1.61E–05
		GO:0034380~high-density lipoprotein particle assembly	3	1.80E–05
		GO:0010873~positive regulation of cholesterol esterification	3	2.32E–05
Module B		GO:0042158~lipoprotein biosynthetic process	3	2.32E–05
		GO:0007186~G-protein-coupled receptor signaling pathway	16	9.00E–16
		GO:0007218~neuropeptide signaling pathway	5	5.55E–06
		GO:0007267~cell–cell signaling	6	9.80E–06
		GO:0007204~positive regulation of cytosolic calcium ion concentration	5	1.70E–05
Module C		GO:0070374~positive regulation of ERK1 and ERK2 cascade	5	4.85E–05
		GO:0045071~negative regulation of viral genome replication	2	1.89E–02
		GO:0019731~antibacterial humoral response	2	2.08E–02
		GO:0009612~response to mechanical stimulus	2	2.78E–02
		GO:0071347~cellular response to interleukin-1	2	3.33E–02

*Note: KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; BP: Biological Process; DEGs: Differentially expressed gene.

3.5. Construction of a ceRNA regulatory network

A total of 11 circRNA-miRNA interaction pairs were predicted between the key circRNAs and dif-miRNAs. Overall, 40 miRNA-mRNA interaction pairs were predicted between the mRNAs in modules and dif-miRNAs. Finally, 46 circRNA-miRNA-mRNA regulatory networks were constructed, which contained 7 circRNAs (6 upregulated and 1 downregulated), 11 miRNAs (6

upregulated and 5 downregulated), and 18 mRNAs (12 upregulated and 6 downregulated) (Figure 4A).

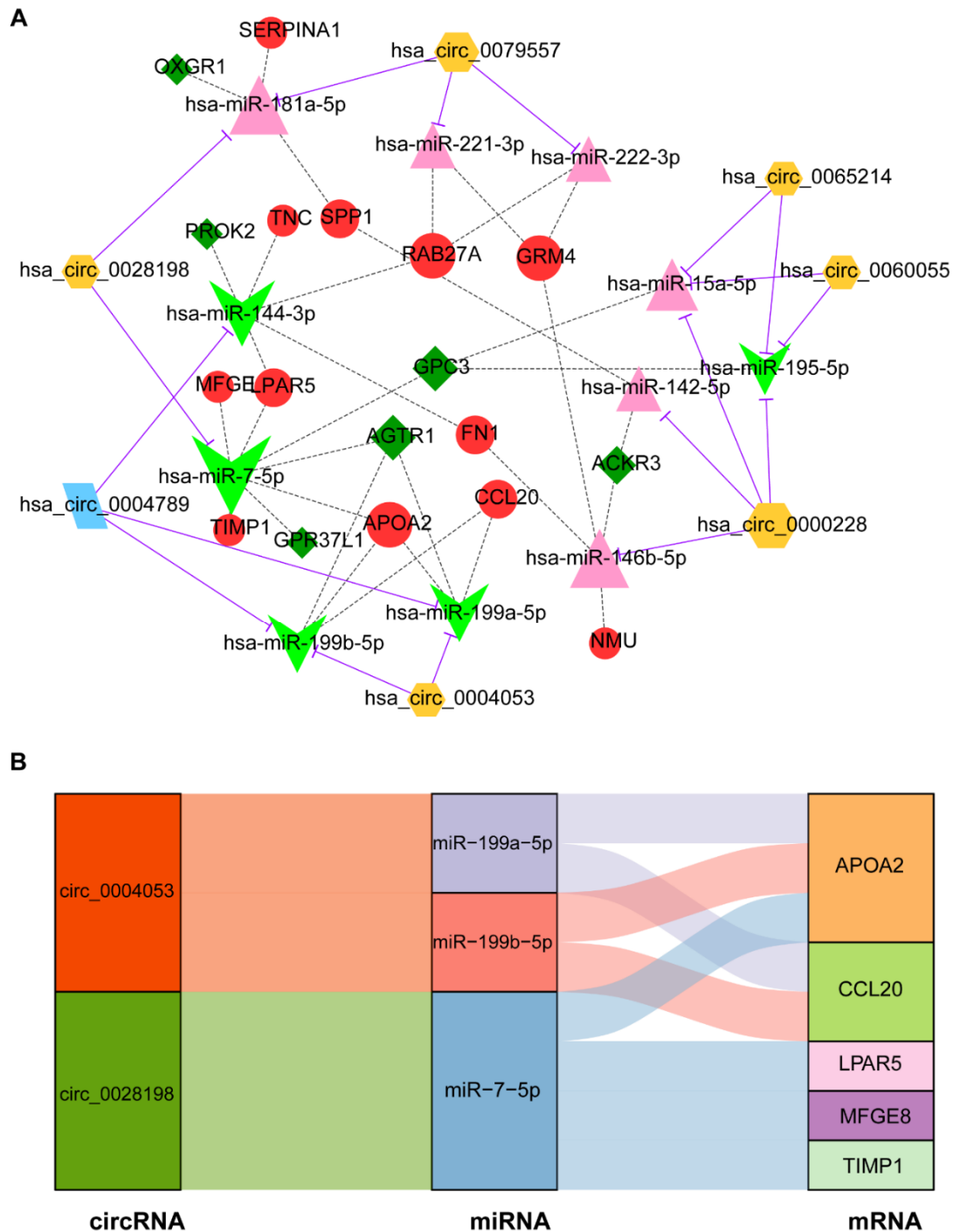


Figure 4. ceRNA regulatory network. (A) The mRNA-miRNA-circRNA regulatory network. Pink triangles represent upregulated miRNAs, pale green arrows represent downregulated miRNAs, yellow hexagons represent upregulated circRNAs, blue quadrilaterals represent downregulated circRNAs, red circles represent upregulated mRNAs, and green rhombuses represent downregulated mRNAs. (B) Sankey diagram showing the ceRNA regulatory pairs.

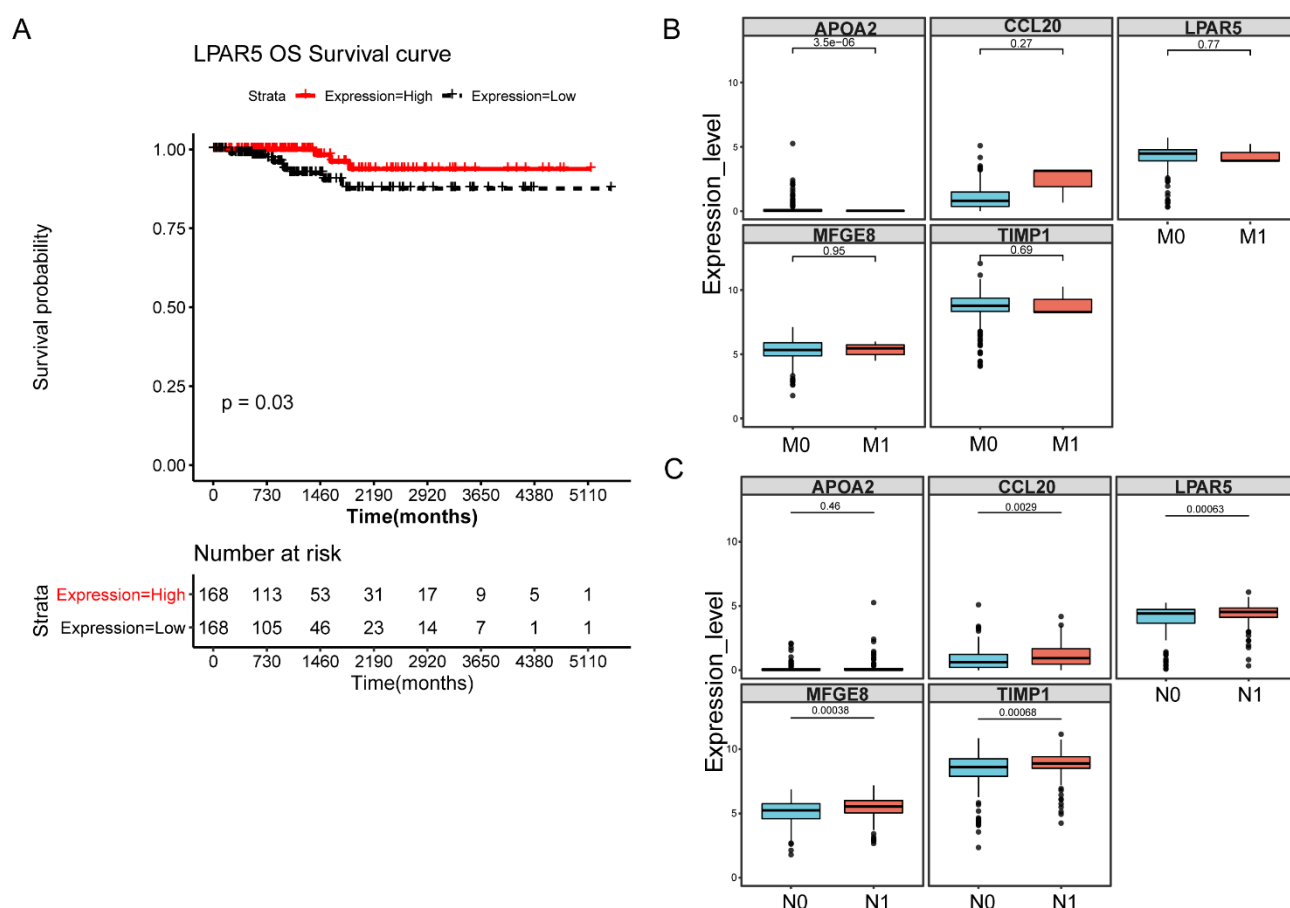


Figure 5. Associations of key mRNAs with survival and clinical features. (A) Kaplan–Meier survival curves showing the prognostic value of *LPAR5*. High *LPAR5* expression was associated with high survival probability ($P = 0.03$). (B) Box plot showing the expression differences of key mRNAs between M0 (non-metastasis) and M1 (metastasis), of which the expression of *APOA2* showed a significant difference ($P = 3.5E-06$), and the expression of *CCL20*, *LPAR5*, *MFGE8*, and *TIMP1* showed no differences (all $P > 0.05$). (C) Box plot showing the expression differences of key mRNAs between N0 (non-metastatic lymph node) and N1 (metastatic lymph node). Expression of *CCL20*, *LPAR5*, *MFGE8*, and *TIMP1* showed significant differences between metastatic and non-metastatic lymph nodes (all $P < 0.05$), while the expression of *APOA2* showed no differences ($P = 0.46$).

Furthermore, the ceRNA regulatory pairs were screened. Based on the ceRNA hypothesis, two RNA molecules regulated by the same miRNA are usually co-expressed, and miRNAs usually inhibit the expression of its target gene (mRNA). Therefore, positive circRNA-mRNA interaction pairs and negative miRNA-mRNA interaction pairs were selected to construct ceRNA regulatory pairs. A total of eight ceRNA regulatory pairs were obtained, including two circRNAs (circ_0004053 and circ_0028198), three miRNAs (miR-199a-5p, miR-199b-5p, and miR-7-5p), and five mRNAs, including *APOA2*, *CCL20*, *LPAR5*, *MFGE8* and *TIMP1* (Figure 4B). To investigate the correlation between miRNA and mRNA expression, two external datasets (GSE104006 and GSE151181) containing both miRNA and mRNA expression data were used. There were significant negative correlations between the expression of miR-7-5p and its targets *LPAR5*, *MFGE8*, and *TIMP1* in both

the GSE104006 and GSE151181 datasets (Figure S2, Table S2).

3.6. Association between the expression of key mRNAs with survival and clinical features

The five mRNAs in the ceRNA regulatory pairs were considered key mRNAs. We further investigated the association of these five key mRNAs with survival and clinical features (Figures 5 and S3). Survival analysis showed that only *LPAR5* expression was associated with patient survival; in particular, high *LPAR5* expression was associated with favorable patient survival (Figure 5A). In addition, *APOA2* expression showed significant differences between metastatic and non-metastatic tumors, while the remaining four mRNAs, *CCL20*, *LPAR5*, *MFGE8* and *TIMP1*, showed significant differences between metastatic and non-metastatic lymph nodes (Figure 5B).

3.7. Validation of the expression of key mRNAs

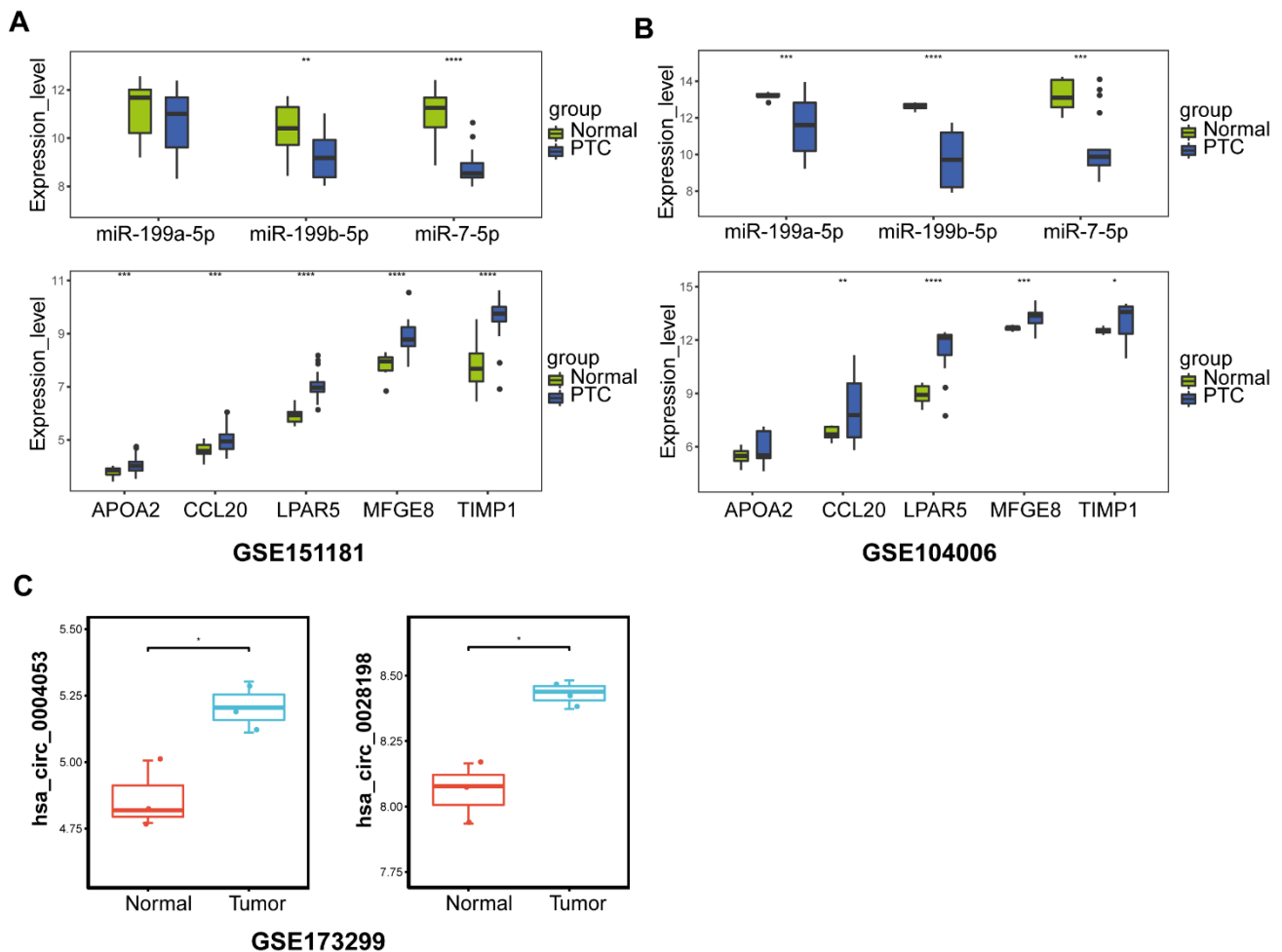


Figure 6. Validation of the expression of key circRNAs, miRNAs and mRNAs using external datasets.

The expression of key circRNAs (circ_0004053 and circ_0028198), miRNAs (miR-199a-5p, miR-199b-5p, and miR-7-5p) and mRNAs (*APOA2*, *CCL20*, *LPAR5*, *MFGE8* and *TIMP1*) in PTC and normal tissues was validated using external validation datasets. As shown in Figure 6A and

6B, the expression of the three miRNAs was confirmed to be downregulated in PTC compared to normal samples, and all the five mRNAs were upregulated in PTC compared to normal samples. The expression of miR-199a-5p in the GSE151181 dataset and expression of *APOA2* in the GSE104006 dataset showed no significant differences. The expression of circ_0004053 and circ_0028198 was upregulated in PTC compared to that in normal samples (Figure 6C). In addition, the expression of key miRNAs and mRNAs in PTC and normal tissues was validated using the UALCAN online tool. As shown in Figure 7, the expression of the three miRNAs was downregulated in tumor samples compared to that in normal samples, especially in classical thyroid papillary carcinoma and follicular thyroid papillary carcinoma (Figure 7A). The five mRNAs showed significantly higher expression in tumor samples than in normal samples; in particular, the expression of these five mRNAs was significantly higher in classical thyroid papillary carcinoma and tall thyroid papillary carcinoma (Figure 7B).

Box plots showing the expression of key miRNAs (upper panel) and mRNAs (lower panel) between PTC and normal samples using the external GSE151181 dataset (A) and the GSE104006 dataset (B). Box plots showing the expression of circRNAs in PTC and normal samples using the external GSE173299 dataset (C). The lines in the middle of each box represent the median expression values of the corresponding mRNA. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

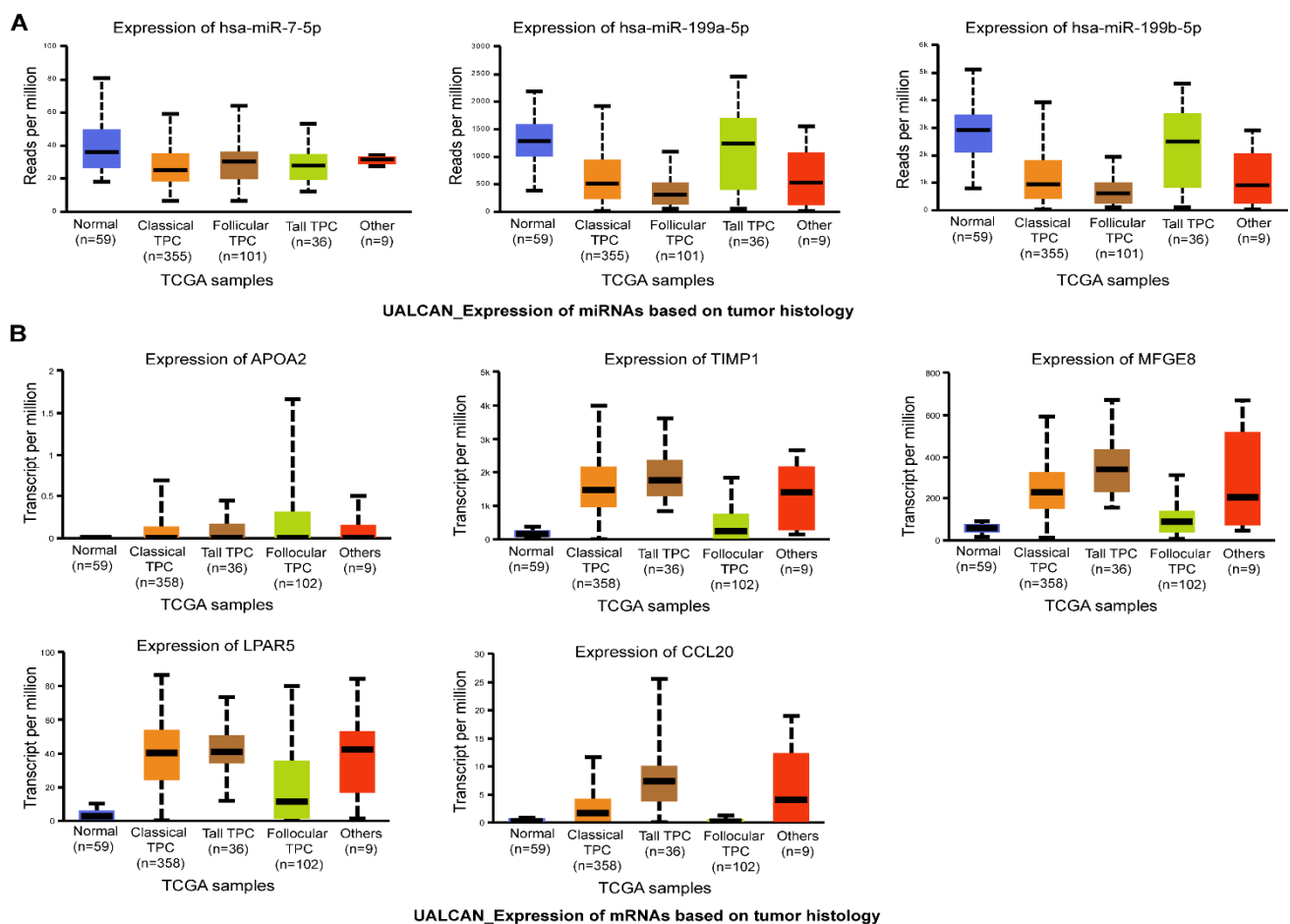


Figure 7. Validation of the expression of key miRNAs and mRNAs using the UALCAN online tool.

Box plots showing the expression differences of key miRNAs (A) and mRNAs (B) between normal and different types of tumor tissue samples, including classical thyroid papillary carcinoma, follicular thyroid papillary carcinoma, Tall thyroid papillary carcinoma and others. Blue blocks represent normal samples, and the other blocks represent tumor histology, including classical, follicular, and tall thyroid papillary carcinoma. The black line in each box represents the median expression value of the corresponding RNA.

3. Discussion

circRNAs are non-coding RNAs usually expressed in a specific tissue or developmental stage and have been considered promising alternative biomarkers owing to their high stability imparted by their covalently closed cyclic structures [20,21]. Various circRNAs have been identified in tumor tissues, functioning as ceRNAs by sponging miRNAs and regulating the transcription and translation of target genes. Through these mechanisms, circRNAs significantly influence several diseases and malignant tumors [22–24]. In this study, 137 circRNAs were found to be differentially expressed in PTC tissues, of which 37 key circRNAs were obtained using the SVM-RFE algorithm. In addition, 1039 dif-mRNAs were screened in PTC, of which 46 key mRNAs in PPI modules were identified. Among the 37 key dif-circRNAs, 46 key mRNAs, and 18 dif-miRNAs, 46 circRNA-miRNA-mRNA regulatory pairs were predicted. Finally, eight regulatory pairs in line with the ceRNA hypothesis were obtained, including two circRNAs (circ_0004053 and circ_0028198), three miRNAs (miR-199a-5p, miR-199b-5p, and miR-7-5p), and five mRNAs (*APOA2*, *CCL20*, *LPAR5*, *MFGE8*, and *TIMP1*).

circ_0028198 has been reported to promote the pathophysiology of atherosclerosis and lipid disorders through regulation of lipoprotein metabolism and cholesterol homeostasis by decreasing cholesterol efflux from macrophages [25–27]; however, the role of circ_0028198 in the development and progression of tumors has not yet been demonstrated. circ_0004053 was identified as a novel circRNA. In our study, circ_0004053 was found to regulate the expression of *APOA2* and *CCL20* by targeting miR-119a-5p and miR-119b-5p, while circ_0028198 targeted miR-7-5p to regulate the expression of *APOA2*, *LPAR5*, *MFGE8*, and *TIMP1*. miR-119a-5p and miR-119b-5p are all members of the miR-199 family, which has been found to be involved in different cancers. For example, Zhang et al. suggested that miR-199 could repress epithelial–mesenchymal transition (EMT) and the invasion of tumor cells by inhibiting the expression of Snail in hepatocellular carcinoma [28]. Koshizuka et al. revealed that the miR-199 family showed antitumor activity in head and neck squamous cell carcinoma, inhibiting the migration of tumor cells by decreasing the expression of integrin $\alpha 3$ [29]. miR-7 has been reported to regulate the expression of multiple oncogenes [30]. For example, Xiao et al. showed that miR-7-5p could inhibit tumor growth and metastasis by targeting neuro-oncological ventral antigen 2 in non-small cell lung cancer both *in vitro* and *in vivo* [31]. In addition, miR-7 has been reported to be involved in colorectal cancer growth and metastasis [32], reversing chemotherapy resistance in breast cancer [33] and inhibiting tumor progression in pancreatic cancer [34] and esophageal cancer [35]. miR-7-5p may be a diagnostic biomarker of PTC [36]; additionally, it has been reported to be involved in regulating the proliferation and migration of PTC cells [37]. These studies suggest the important role of miR-199 and miR-7-5p in tumors; moreover, circ_0004053 and circ_0028198 target these miRNAs and thus may play key roles in PTC.

The ceRNA regulatory network analysis suggested that circ_0004053 and circ_0028198 regulate the expression of *APOA2*, *CCL20*, *LPAR5*, *MFGE8*, and *TIMP1*. Further analysis revealed that the expression of these five mRNAs showed significant differences between metastasis/lymph

node metastasis and non-metastasis/non-lymph node metastasis. In addition, survival analysis showed that *LPAR5* expression was associated with patient survival. The *APOA2* gene encodes apolipoprotein A2, which has been suggested as a biomarker for pancreatic cancers [38,39]. *CCL20* encodes a C-C motif chemokine ligand for the receptor CCR6; thus, it promotes cell migration and invasion in thyroid cancer by activating CCR6 [40]. *LPAR5* encodes a lysophosphatidic acid receptor, which contributes to the metastasis and tumorigenesis of PTC by activating the PI3K/AKT pathway and EMT [41,42]. MFGE8 is a milk fat globule-epidermal growth factor that has been found to play crucial roles in regulating transforming growth factor- β -induced EMT in endometrial epithelial cells [43] and cardiac fibrosis [44]. Tissue inhibitors of metalloproteinases (TIMPs) have been considered to contribute to cancer hallmarks. TIMPs are always aberrantly expressed in tumors, and a high expression of TIMP1 is correlated with tumor progression and worse survival [45]. Zhang et al. suggested that TIMP1 was implicated in the lymphatic metastasis of PTC by regulating the invasive and migratory activities of PTC cells [46]. These data indicate that the five aforementioned mRNAs play significant roles in mediating EMT, metastasis, cell invasion, and migration, consistent with the findings that the expression of these five mRNAs exhibit significant differences between metastasis/lymph node metastasis and non-metastasis/non-lymph node metastasis.

Although we have validated the expression of key circRNAs, miRNAs, and mRNAs using external datasets, their expression should be further confirmed through biological experiments using clinical samples, and their clinical value should be further investigated based on clinical data with a large sample size. In addition, the ceRNA regulatory pairs predicted were based on positive expression trends between circRNAs and mRNAs and negative expression trends between miRNAs and mRNAs because expression correlation analysis could not be performed among the different datasets. We performed Pearson correlation analyses using external datasets to investigate the expression correlation between miRNAs and mRNAs, and no significant correlations were found between miR-199a/b-5p and their targets. Therefore, in the future, a dual luciferase assay should be used to verify the binding of an miRNA to its target, and the ceRNA regulatory pairs should be confirmed by conducting functional experiments.

4. Conclusions

In conclusion, we identified eight ceRNA regulatory pairs, involving two circRNAs (circ_0004053 and circ_0028198), three miRNAs (miR-119a-5p, miR-119b-5p, and miR-7-5p), and five mRNAs (*APOA2*, *CCL20*, *LPAR5*, *MFGE8*, and *TIMP1*), in PTC. These five mRNAs were correlated with metastatic and non-metastatic lymph nodes in PTC. Our study reveals several potential targets and the underlying regulatory mechanisms in PTC.

Acknowledgments

This work was supported by the Hainan Provincial Department of Science and Technology, under Grant No. HN01ENT07, and the Clinical Medicine Project of the Affiliated Hospital of Qingdao University, under Grant No. 3400.

Conflict of interest

The authors declare that they have no competing interests.

References

1. B. Jankovic, K. T. Le, J. M. Hershman, Hashimoto's thyroiditis and papillary thyroid carcinoma: is there a correlation?, *J. Clin. Endocrinol. Metab.*, **98** (2013), 474–482.
2. D. S. McLeod, A. M. Sawka, D. S. Cooper, Controversies in primary treatment of low-risk papillary thyroid cancer, *The Lancet*, **381** (2013), 1046–1057.
3. A. Cummings, M. Goldfarb, Thyroid carcinoma metastases to axillary lymph nodes: report of two rare cases of papillary and medullary thyroid carcinoma and literature review, *Endocr. Pract.*, **20** (2014), e34–e37.
4. G. H. Sakorafas, D. Sampanis, M. Safioleas, Cervical lymph node dissection in papillary thyroid cancer: current trends, persisting controversies, and unclarified uncertainties, *Surg. Oncol.*, **19** (2010), e57–e70.
5. L. L. Chen, The biogenesis and emerging roles of circular RNAs, *Nat. Rev. Mol. Cell Bio.*, **17** (2016), 205–211.
6. H. L. Sanger, G. Klotz, D. Riesner, H. J. Gross, A. K. Kleinschmidt, Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures, *Proc. Natl. Acad. Sci.*, **73** (1976), 3852–3856.
7. J. Li, J. Yang, P. Zhou, Y. Le, C. Zhou, S. Wang, et al., Circular RNAs in cancer: novel insights into origins, properties, functions and implications, *Am. J. Cancer Res.*, **5** (2015), 472–480.
8. Y. Zhang, X. O. Zhang, T. Chen, J. F. Xiang, Q. F. Yin, Y. H. Xing, et al., Circular intronic long noncoding RNAs, *Mol. Cell*, **51** (2013), 792–806.
9. W. R. Jeck, N. E. Sharpless, Detecting and characterizing circular RNAs, *Nat. Biotechnol.*, **32** (2014), 453–461.
10. Z. Zhong, M. Huang, M. Lv, Y. He, C. Duan, L. Zhang, et al., Circular RNA MYLK as a competing endogenous RNA promotes bladder cancer progression through modulating VEGFA/VEGFR2 signaling pathway, *Cancer Lett.*, **403** (2017), 305–317.
11. L. Lü, J. Sun, P. Shi, W. Kong, K. Xu, B. He, et al., Identification of circular RNAs as a promising new class of diagnostic biomarkers for human breast cancer, *Oncotarget*, **8** (2017), 44096–44107.
12. K. Y. Hsiao, Y. C. Lin, S. K. Gupta, N. Chang, L. Yen, H. S. Sun, et al., Noncoding effects of circular RNA CCDC66 promote colon cancer growth and metastasis, *Cancer Res.*, **77** (2017), 2339–2350.
13. D. Han, J. Li, H. Wang, X. Su, J. Hou, Y. Gu, et al., Circular RNA circMTO1 acts as the sponge of microRNA-9 to suppress hepatocellular carcinoma progression, *Hepatology*, **66** (2017), 1151–1164.
14. G. K. Smyth, LIMMA: linear models for microarray data, in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Springer, (2005), 397–420.
15. M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, et al., Gene ontology: tool for the unification of biology, *Nat. Genet.*, **25** (2000), 25–29.
16. M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Res.*, **28** (2000), 27–30.
17. D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, et al., STRING v10: protein–protein interaction networks, integrated over the tree of life, *Nucleic Acids Res.*, **43** (2015), D447–D452.

18. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Research*, **13** (2003), 2498–2504.
19. W. P. Bandettini, P. Kellman, C. Mancini, O. J. Booker, S. Vasu, S. W. Leung, et al., MultiContrast delayed enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study, *J. Cardiovasc. Magn. Reson.*, **14** (2012), 83.
20. H. D. Zhang, L. H. Jiang, D. W. Sun, J. C. Hou, Z. L. Ji, et al., CircRNA: a novel type of biomarker for cancer, *Breast Cancer*, **25** (2018), 1–7.
21. G. Wu, W. Zhou, X. Pan, Z. Sun, Y. Sun, H. Xu, et al., Circular RNA profiling reveals exosomal circ_0006156 as a novel biomarker in papillary thyroid cancer, *Mol. Ther. Nucleic Acids*, **19** (2020), 1134–1144.
22. Q. Liu, L. Z. Pan, M. Hu, J. Y. Ma, Molecular network-based identification of circular RNA-associated ceRNA network in papillary thyroid cancer, *Pathol. Oncol. Res.*, **26** (2020), 1293–1299.
23. S. Borran, G. Ahmadi, S. Rezaei, M. M. Anari, M. Modabberi, Z. Azarash, et al., Circular RNAs: new players in thyroid cancer, *Pathol. Res. Pract.*, **216** (2020), 153217.
24. X. Li, J. Ding, X. Wang, Z. Cheng, Q. Zhu, NUDT21 regulates circRNA cyclization and ceRNA crosstalk in hepatocellular carcinoma, *Oncogene*, **39** (2020), 891–904.
25. P. F. Tussy, I. R. Maldonado, C. F. Hernando, MicroRNAs and circular RNAs in lipoprotein metabolism, *Curr. Atheroscler. Rep.*, **23** (2021), 33.
26. G. Yu, Z. Yang, T. Peng, Y. Lv, Circular RNAs: rising stars in lipid metabolism and lipid disorders, *J. Cell Physiol.*, **236** (2021), 4797–4806.
27. L. Wang, Z. Zheng, X. Feng, X. Zang, W. Ding, F. Wu, et al., circRNA/lncRNA-miRNA-mRNA network in oxidized, low-density, lipoprotein-induced foam cells, *DNA Cell Biol.*, **38** (2019), 1499–1511.
28. H. Y. Zhang, C. H. Li, X. C. Wang, Y. Q. Luo, X. D. Cao, J. J. Chen, MiR-199 inhibits EMT and invasion of hepatoma cells through inhibition of Snail expression, *Eur. Rev. Med. Pharmacol. Sci.*, **23** (2019), 7884–7891.
29. K. Koshizuka, T. Hanazawa, N. Kikkawa, T. Arai, A. Okato, A. Kurozumi, et al., Regulation of ITGA3 by the anti-tumor miR-199 family inhibits cancer cell migration and invasion in head and neck cancer, *Cancer Sci.*, **108** (2017), 1681–1692.
30. T. B. Hansen, J. Kjems, C. K. Damgaard, Circular RNA and miR-7 in cancer, *Cancer Res.*, **73** (2013), 5609–5612.
31. H. Xiao, MiR-7-5p suppresses tumor metastasis of non-small cell lung cancer by targeting NOVA2, *Cell Mol. Biol. Lett.*, **24** (2019), 60.
32. K. Zeng, X. Chen, M. Xu, X. Liu, X. Hu, T. Xu, et al., CircHIPK3 promotes colorectal cancer growth and metastasis by sponging miR-7, *Cell Death Dis.*, **9** (2018), 417.
33. T. Hong, J. Ding, W. Li, miR-7 reverses breast cancer resistance to chemotherapy by targeting MRP1 and BCL2, *Onco. Targets Ther.*, **12** (2019), 11097–11105.
34. J. Xia, T. Cao, C. Ma, Y. Shi, Y. Sun, Z. P. Wang, et al., MiR-7 suppresses tumor progression by directly targeting MAP3K9 in pancreatic cancer, *Mol. Ther. Nucleic Acids*, **13** (2018), 121–132.
35. W. Shi, J. Song, Z. Gao, X. Liu, W. Wang, Downregulation of miR-7-5p Inhibits the Tumorigenesis of esophagus cancer via Targeting KLF4, *Onco. Targets Ther.*, **13** (2020), 9443–9453.

36. Y. Duan, Y. Zhang, W. Peng, P. Jiang, Z. Deng, C. Wu, MiR-7-5p and miR-451 as diagnostic biomarkers for papillary thyroid carcinoma in formalin-fixed paraffin-embedded tissues, *Pharmazie*, **75** (2020), 266–270.
37. J. Y. Han, S. Guo, N. Wei, R. Xue, W. Li, G. Dong, et al, CiRS-7 promotes the proliferation and migration of papillary thyroid cancer by negatively regulating the miR-7/epidermal growth factor receptor axis, *Biomed Res. Int.*, **2020** (2020), 9875636.
38. K. Honda, V. A. Katzke, A. Hüsing, S. Okaya, H. Shoji, K. Onidani, et al., CA19-9 and apolipoprotein-A2 isoforms as detection markers for pancreatic cancer: a prospective evaluation, *Int. J. Cancer*, **144** (2019), 1877–1887.
39. Y. Sato, T. Kobayashi, S. Nishiumi, A. Okada, T. Fujita, T. Sanuki, et al., Prospective study using plasma apolipoprotein a2-isoforms to screen for high-risk status of pancreatic cancer, *Cancers (Basel)*, **12** (2020), 2625.
40. W. Zeng, H. Chang, M. Ma, Y. Li, CCL20/CCR6 promotes the invasion and migration of thyroid cancer cells via NF-kappa B signaling-induced MMP-3 production, *Exp Mol. Pathol.*, **97** (2014), 184–190.
41. C. Y. Wu, C. Zheng, E. J. Xia, R. Quan, J. Hu, X. Zhang, et al., Lysophosphatidic acid receptor 5 (LPA5) plays a significance role in papillary thyroid cancer via phosphatidylinositol 3-Kinase/Akt/Mammalian target of rapamycin (mTOR) pathway, *Med. Sci. Monit.*, **26** (2020), e919820.
42. W. J. Zhao, L. L. Zhu, W. Q. Yang, S. J. Xu, J. Chen, X. F. Ding, et al., LPA5 promotes thyroid carcinoma cell proliferation and migration by activating class IA PI3K catalytic subunit p110 β , *Cancer Sci.*, **112** (2021), 1624–1632.
43. L. Yu, R. Hu, C. Sullivan, R. J. Swanson, S. Oehninger, Y. P. Sun, et al., MFGE8 regulates TGF- β -induced epithelial mesenchymal transition in endometrial epithelial cells in vitro, *Reproduction*, **152** (2016), 225–233.
44. B. Wang, Z. Ge, Y. Wu, Y. Zha, X. Zhang, Y. Yan, et al., MFGE8 is down-regulated in cardiac fibrosis and attenuates endothelial-mesenchymal transition through Smad2/3-Snail signalling pathway, *J. Cell. Mol. Med.*, **24** (2020), 12799–12812.
45. H. W. Jackson, V. Defamie, P. Waterhouse, R. Khokha, TIMPs: versatile extracellular regulators in cancer, *Nat. Rev. Cancer*, **17** (2017), 38–53.
46. W. Q. Zhang, W. Sun, Y. Qin, C. Wu, L. He, T. Zhang, et al., Knockdown of KDM1A suppresses tumour migration and invasion by epigenetically regulating the TIMP1/MMP9 pathway in papillary thyroid cancer, *J. Cell. Mol. Med.*, **23** (2019), 4933–4944.



AIMS Press

©2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)