Construction and analysis of an aberrant lncRNA-miRNA-mRNA network associated with papillary thyroid cancer

Yanxia Jiang, MD, PhD[a], Jiao Wang, MD, PhD[a], Jian Chen, MD, PhD[b], Jiancheng Wang, MD, PhD[a], Jixiong Xu, MD, PhD[a,*] ID

Abstract

Accumulating evidence has indicated that long noncoding RNAs (lncRNAs) are the main constituents of competing endogenous RNA (ceRNA) networks. Nonetheless, in the lncRNA-related ceRNA network of papillary thyroid cancer (PTC), the function of cancer-specific lncRNAs, as well as their use for the potential prediction of PTC prognosis, remains unclear. In this study, 384 RNA sequencing (RNA-seq) profiles of PTC patients were attained from The Cancer Genome Atlas (TCGA), an open-source database that offers vast amounts of RNA-seq data, and 75 miRNAs, 495 lncRNAs, and 1099 mRNAs ($P < .05$ and $|logFC| > 2$) were detected when compared with normal tissues. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were analyzed using the Cytoscape plug-in BinGo. An aberrant lncRNA-mRNA-miRNA ceRNA network consisting of 31 differentially expressed (DE)-lncRNAs, 13 DE-miRNAs, and 134 DE-mRNAs was built in TCGA. On the basis of overall survival (OS) analysis, 6 lncRNAs (CCAT1, SYNPR, SFTA1P, HOTAIR, HCG22, and CLDN10) were identified as prognostic biomarkers for patients in TCGA ($P < .05$). Through qRT-PCR, we designated 6 cancer-specific lncRNAs as having great significance for survival by verifying their expression in the 60 PTC patients who were diagnosed. The qRT-PCR and TCGA results were completely consistent. Our research provides data for further understanding the lncRNA-miRNA-mRNA ceRNA network and elucidating the molecular mechanisms of PTC.

Abbreviations: BP = biological processes, CC = cellular components, ceRNA = competing endogenous RNA, DE = differentially expressed, FC = fold change, GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, lncRNA = long noncoding RNAs, MF = molecular functions, OS = overall survival, PTC = papillary thyroid cancer, RNA-seq = RNA sequencing, ROC = receiver-operator characteristic.

Keywords: competitive endogenous RNA, long noncoding RNAs, papillary thyroid cancer

1. Introduction

The prevalence of thyroid cancer, which is a malignant endocrine tumor, continues to increase worldwide.[1] Among all kinds of thyroid cancer, papillary thyroid cancer (PTC) is identified as the most common pathological type and is characterized as a differentiated neoplasia.[2] The incidence of PTC has been increasing yearly and accounts for more than 90% of all thyroid cancers.[3] Although most PTCs are concealed and characterized by slow progression and a good overall prognosis, a small number of cases show increased invasiveness and recurrence after surgery.[4] Therefore, the need to investigate novel functional molecular and therapeutic targets for the treatment of PTC is increasing.

Long noncoding RNAs (lncRNAs) are greater than 200 nucleotides (nt) and do not code for proteins.[5] According to previous studies, lncRNAs are predominant in the human genome and may be used as diagnostic markers of cancers.[6,7] Aberrantly expressed lncRNAs, which can regulate gene expression at the level of transcription, posttranscription, and epigenetic levels, have been detected in PTC.[8–10] Moreover, as competing endogenous RNA (ceRNA) regulation in human cells was discovered, a vast number of studies have shown that mRNAs, lncRNAs, and other RNAs act as natural miRNA sponges and can inhibit miRNA function. Through competition for shared miRNAs, lncRNAs communicate with mRNA when
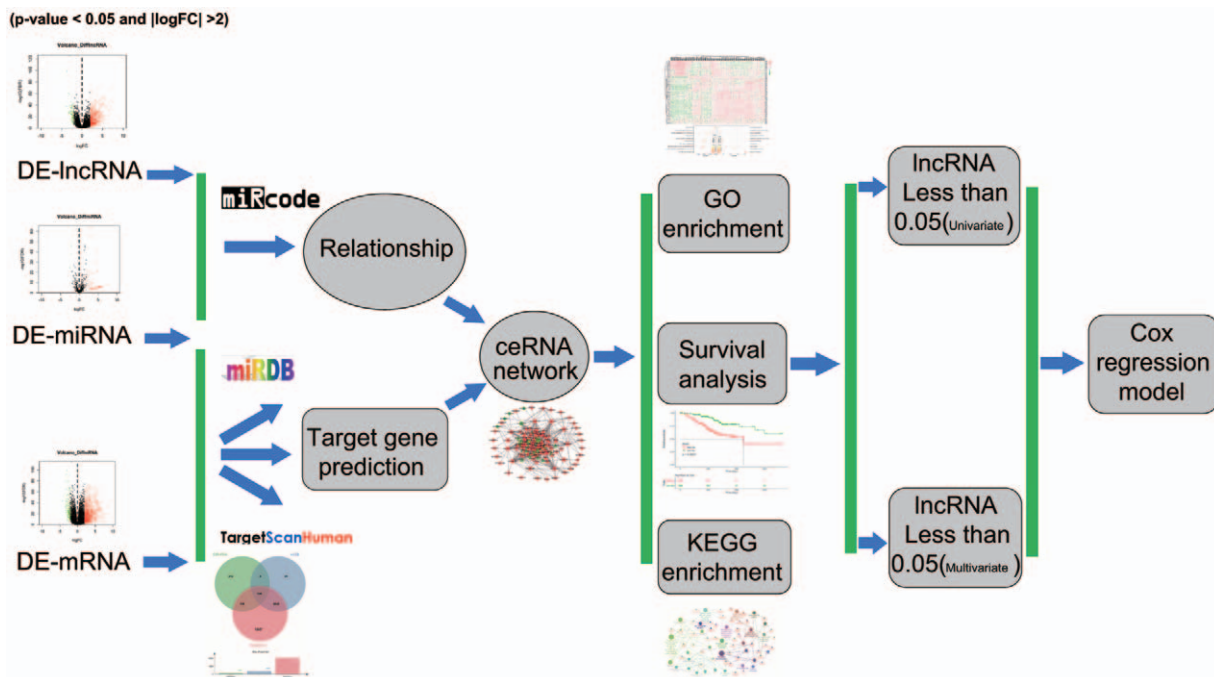
Figure 1. Flow chart of the bioinformatics analysis.

acting as ceRNAs.[11] In one study, Song et al[12] analyzed lncRNA profiles of 306 cervical squamous cancer (CESC) patients from The Cancer Genome Atlas (TCGA, http://cancer genome.nih.gov/) and then built a lncRNA-miRNA-mRNA network in CESC. According to the ceRNA hypothesis, Sui et al[13] examined 465 RNA sequencing (RNA-seq) profiles of lung adenocarcinoma (LUAD) patients from TCGA, ultimately revealing a correlation between the RNA data and clinical information of the patients, and a lncRNA-miRNA-mRNA ceRNA network of LUAD was generated.

In this research, abnormal expression profiles of mRNA, miRNA, and lncRNA in 384 PTC patients along with 58 nontumor samples were acquired. Next, an aberrant lncRNA-miRNA-mRNA ceRNA network was assembled for PTC. On the basis of overall survival (OS) analysis, 6 of 31 lncRNAs functioned as prognostic biomarkers in PTC patients. This lncRNA-miRNA-mRNA ceRNA network for PTC can be used to assist clinicians in explaining lncRNA function and for developing PTC prognosis and reducing its incidence. Diagnostic biomarkers that effectively help clinicians select the appropriate treatment are crucial for risk assessment and for early detection of PTC.

2. Materials and methods

2.1. Sample collection

In total, information for 509 PTC patients was obtained from TCGA. After ruling out PTC patients without complete follow-up information, 384 remained for inclusion in survival analysis. Moreover, we obtained mRNA, miRNA, and lncRNA expression profiles from tumors and relevant normal samples from the PTC data portal (https://tcga-data.nci.nih.gov/tcga/).

In addition, we acquired specimens from 60 patients diagnosed based on the National Comprehensive Cancer Network (NCCN)

clinical practice guidelines for PTC from May 2018 to April 2019 at the First Affiliated Hospital of Nanchang University. All patients who received targeted therapy or with other kinds of malignant tumors were excluded. A total of 20 plasma samples from healthy individuals were applied as controls for validation. Venous blood was collected in ethylenediaminetetraacetic acid tubes (BD Biosciences, Franklin Lakes, NJ). Subsequently, the plasma was transferred to a fresh test tube and stored at –80°C after being flash-frozen in liquid nitrogen. This research was authorized by Nanchang University's Research Ethics Committee, and written informed consent was acquired from all subjects.

2.2. RNA sequence data processing and computational analysis

For empirical analysis of digital gene expression data, the EdgeR package,[14] which is mainly applied for studying differential expression of repeated count data on the basis of Poisson mode, was employed to assess RNA-seq as well as miRNA sequencing (miRNA-seq) data from TCGA. We only considered significance at a P value for differentially expressed (DE) mRNA, miRNA, and lncRNA of less than .05 and |log2-fold change (FC)| $\geq$ 2. Figure 1 shows the bioinformatics analysis flow chart.

2.3. Construction of the lncRNA-miRNA-mRNA ceRNA network

The ceRNA network was built on the basis of the association among mRNA, miRNA and lncRNAs. First, aberrantly expressed mRNAs, miRNAs, and lncRNAs were identified, and then lncRNA-miRNA connections were predicted by utilizing miR-code[15] (http://www.mircode.org/), a database containing more than 10,000 lncRNAs that provides all of the transcriptome human miRNA target predictions based on comprehensive GENCODE

(http://www.gencodegenes.org/) gene annotation. Next, TargetScan[16] (http://www.targetscan.org/) and miRDB[17] (http://www.mirdb.org/), 2 popular online databases for the prediction of miRNA-targeted mRNA and functional annotation, were used to predict interaction between miRNA and mRNA. Then, every targeted mRNA in this study was chosen for interaction with DE-mRNAs. Ultimately, Cytoscape software[18] (version 3.5.1) was used to construct the ceRNA network.

2.4. Functional enrichment analysis of targeted DE-mRNAs

Database for Annotation, Visualization, and Integrated Discovery (DAVID), an online instrument, has been employed for functional enrichment analysis in numerous studies. In this study, DAVID was applied for Gene Ontology (GO) analysis of targeted genes' role from the aspects of cellular components (CC), biological processes (BP), and molecular functions (MF). In addition, Kyoto Encyclopedia of Genes and Genomes (KEGG) and an R Package, clusterProfiler,[19] were employed to differentiate enriched pathways of every targeted gene. The cutoff value was set as a *P* value less than .05.

2.5. RNA isolation from human plasma and RT-qPCR

Whole RNA was extracted from human plasma using mirVana PARIS RNA Isolation Kit (Ambion; Thermo Fisher Scientific, Inc., Waltham, MA) according to the manufacturer's liquid sample protocol. The RNA purity as well as the concentration extracted based on optical density at 260 and 280 nm. cDNA was synthesized from RNA via RT by utilizing gene-specific primers (Applied Biosystems; Thermo Fisher Scientific, Inc., Waltham, MA) and Moloney Murine Leukemia Virus RT Kit, according to the following manufacturer's protocol, as follows: 42°C for 2 to 5 minutes, 42°C for 15 minutes, and 70°C for 50 minutes. The primer sequences for PCR amplification are summarized in Table 1. qRT-PCR was utilized to assess candidate lncRNA expression levels with the StepOnePlus Real-Time PCR system. U6 was used as an internal control. The PCR conditions were as follows: 40 cycles at 95°C for 5 minutes, 95°C for 45 seconds, 55°C for 50 seconds, and 72°C for 15 seconds. The samples were analyzed in triplicate, and gene expression was quantified by normalizing expression of the target gene to the internal control utilizing the $2^{-\Delta\Delta Cq}$ formula.[20]

Table 1

The primer sequences for PCR amplification.

Gene	Primer sequence	Tm, °C,	Position	Product size, bp
CCAT1	F:GCAGAAAGGCCAGTGCT	65.3	1455	101
	R:GCAGGAGGGTGCTTGAC	65.3	1555	
SYNPR	F:GCTGGCATTGATTGGTG	61.4	559	123
	R:TCCCGGTATTTGTTCTGG	61.4	681	
SFTA1P	F:GCTTCCACGGATTTTCAC	61.4	556	69
	R:CATTCCAGGTGGGCTTT	62	624	
HOTAIR	F:AATTAGCGCCTCCCAGTC	64.3	1266	140
	R:GGGCTTCCTTGCTCTTCT	64.32	1405	
HCG22	F:GGACTGGGCTTCATTGTG	63.1	1680	148
	R:ACCCTGGTGGATGGATTT	63.1	1738	
CLDN10	F:CCACCAAAGCGTTCTGAC	63.7	197	101
	R:CTCCAAACCCTCCGACA	63.4	197	

2.6. Construction of lncRNA-based signatures

We first identified lncRNAs related to survival, with a *P* value <.05 in univariate Cox analysis. Then, multivariate Cox regression analysis was conducted based on the lncRNAs selected by univariate analysis and to construct a multi-lncRNA based signature for predicting survival in patients with PTC based on a linear combination of the regression coefficients derived from the multivariate Cox model coefficients. The risk score for each patient was calculated based on βi combined with corresponding expression data of the identified lncRNAs.[21,22]

Then, the PTC patients were classified into low- and high-risk groups according to the median risk score. Finally, Kaplan–Meier analysis was utilized between the low- and high-risk groups to assess OS. Receiver operating characteristic (ROC) analysis was performed to compare the specificity and sensitivity of the survival prediction using R software (version 3.4.3).[23]

2.7. Statistical analysis

All results are presented as the mean ± SD. Student *t* test was utilized to determine the significance of differences between normal and tumor groups, and a *P* value less than .05 was considered significant.

3. Results

### 3.1. Differentially expressed mRNAs, miRNAs, and lncRNAs

We calculated the expression profiles of miRNAs, mRNAs, and lncRNAs in 384 PTC samples and 58 nonmalignant samples and found that 1099 mRNAs, 75 miRNAs, and 495 lncRNAs were DE (|logFC| ≥ 2.0 and *P* < .05) (see Table, http://links.lww.com/MD/F143 Supplemental Content, which illustrates the DE mRNAs, miRNAs, and lncRNAs). Specifically, 399 lncRNAs, 50 miRNAs, and 866 mRNAs were overexpressed compared with the normal tissue, and 25 miRNAs, 96 lncRNAs, and 233 mRNAs were underexpressed. The heatmap of the analyzed RNA cluster analysis is displayed in Figures 2 to 4.

3.2. Prediction of miRNA targets and ceRNA network construction

A lncRNA-miRNA-mRNA ceRNA network for PTC was established to better comprehend the function of lncRNAs DE in this cancer. Initially, 75 miRNAs were recognized among PTC patient tissues and contiguous nontumor tissues. We found that these 75 DE-miRNAs target 11,650 mRNAs in the TargetScan database and 3069 mRNAs in the miRDB database. On the basis of a comparison of the 3 groups mentioned above, 134 mRNAs were chosen from the intersection of aberrantly expressed miRNAs (Fig. 5). We concentrated on the targeted relationship among the 13 specific miRNAs that interact with 31 specific lncRNAs (Table 2). In the next step, miRTarBase and TargetScan were utilized to predict miRNA-targeted mRNAs. Thirteen specific miRNAs associated with 134 intersecting mRNAs were identified (Table 3). The targets of a few of them are cancer-associated genes, including *EGR2, TRIM36, LAMP5, TGFA, RUNX2, BEAN1*, and *STAT3*. Finally, a ceRNA regulatory network for PTC was established by incorporating 31 DElncRNAs, 13 DEmiRNAs, and 134 DEmRNAs according to the
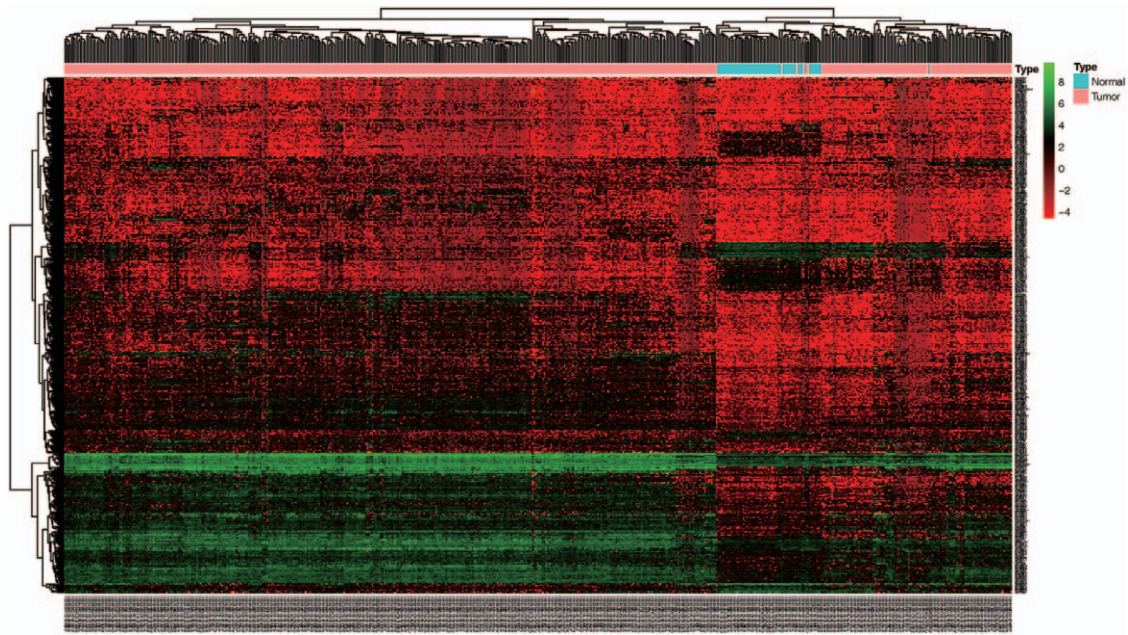
Figure 2. The basic heatmap of DE-lncRNAs showing expression in PTC. In total, 495 DE-lncRNAs were detected. In the sample, 399 DE-mRNAs were upregulated and 96 DE-lncRNAs downregulated. The color blue indicates low expression and red high expression during PTC progression.
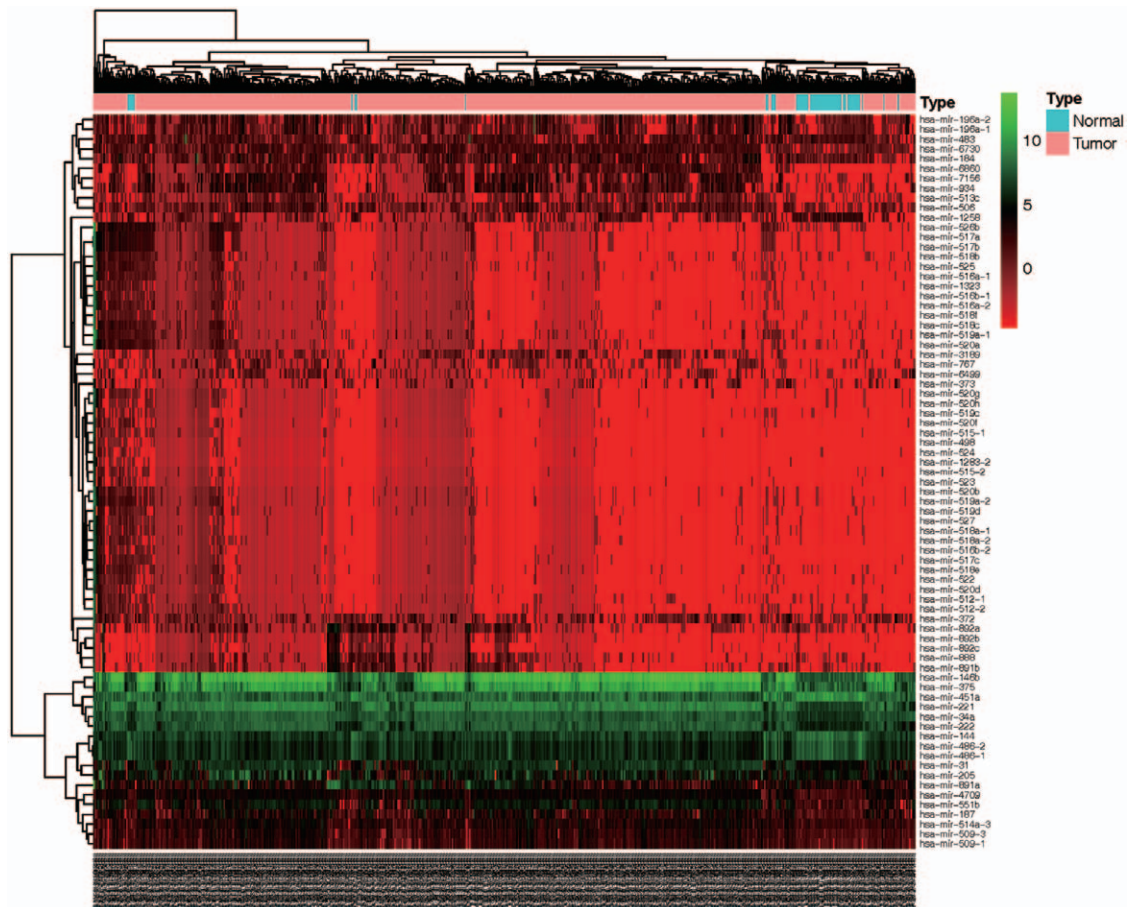
Figure 3. The basic heatmap of DE-miRNA expression in PTC. In total, 75 DE-miRNAs were detected. In the sample, 50 DEmRNAs were upregulated and 25 DE-lncRNAs downregulated. Low to high expression is shown by color varying from blue to red during PTC progression.
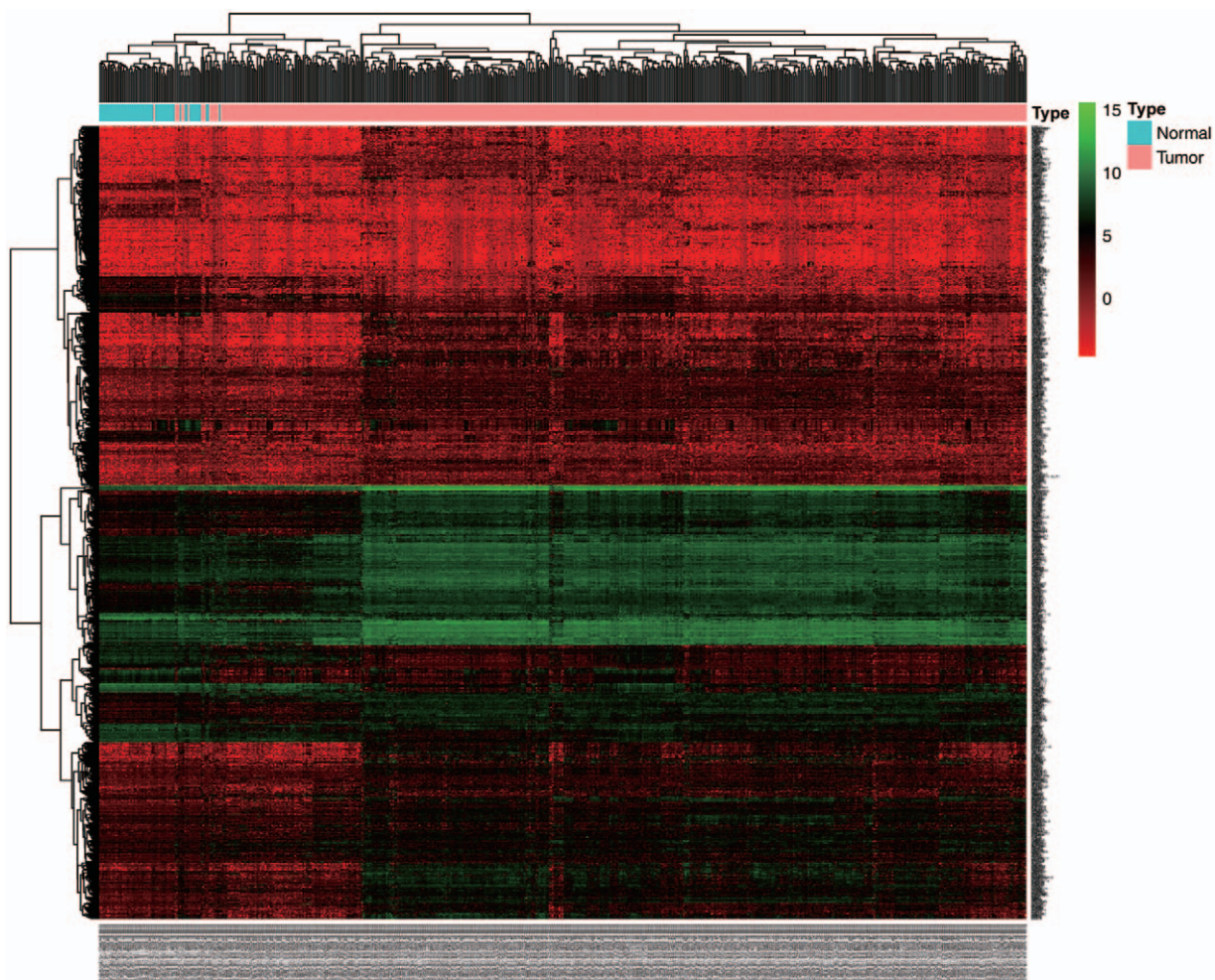
Figure 4. The basic heatmap of DE-mRNA showing expression in PTC. In total, 1099 DE-lncRNAs were detected. In the sample, 866 DE-lncRNAs were upregulated and 233 downregulated. Low to high expression is displayed by color varying from blue to red during PTC progression.

above data (Tables 2 and 3). Figure 6 displays the ceRNA network, which was plotted by utilizing Cytoscape 3.0.pt

3.3. Functional enrichment analysis

Functional enrichment analysis indicated enrichment of 120 GO (Fig. 7 and Table 4) and 87 KEGG (Fig. 8 and Table 5) categories. The top 5 GO terms were reactions to the plasma membrane, an integral constituent of the plasma membrane, cell junction, cell adhesion, and neuronal cell body. With an emphasis on biological pathways, KEGG analysis revealed that 20 pathways were expressively enriched, especially the PPAR signaling pathway, the HIF-1 signaling pathway, PI3K inhibitors blocking PI3K catalytic activity and lung fibrosis, participating in cancer development, invasion, and metastasis.

3.4. Prognostic assessment of lncRNAs

Cox regression analysis was performed to screen prognostic factors in PTC correlating with OS. The univariate Cox proportional hazards regression model revealed 6 of 31 DElncRNAs, such as *CCAT1, SYNPR, SFTA1P, HOTAIR, HCG22, and CLDN10*, to be associated with OS ($P < .05$); 5 of

these genes (*CCAT1, SYNPR, SFTA1P, HOTAIR*, and *HCG22*) correlated positively with OS (Fig. 9). Multivariable Cox regression analysis demonstrated 4 lncRNAs (*CCAT1, SYNPR, HCG22, CLDN10*) to be independent prognostic factors for OS in PTC patients ($P < .05$) (Table 6). The risk formula is as follows: $Risk\ Score=(0.261^*HCG22)-(0.924^*SYNPR.AS10) +(0.258^*UCA1)-(0.861^*CLDN10.AS1)$. Then, we employed ROC curves to examine diagnostic performance by comparing the area under the respective ROC curve (AUC). A KM plot of these lncRNAs associated with OS is displayed in Figure 10A. For the ROC curve displayed in Figure 10B, the AUC of the four lncRNAs was 0.941.

3.5. qRT-PCR verification

Six specific lncRNAs (*CCAT1, SYNPR, SFTA1P, HOTAIR, HCG22, and CLDN10*) were selected to verify the accuracy and consistency of the results mentioned above. Paired *t* tests were used to estimate differences among PTC tumor tissues and nontumor tissues. *CCAT1, SYNPR, SFTA1P, HOTAIR*, and *HCG22* were upregulated in PTC tumor tissues when compared with nontumor tissues, whereas *CLDN10* was downregulated (Fig. 11).
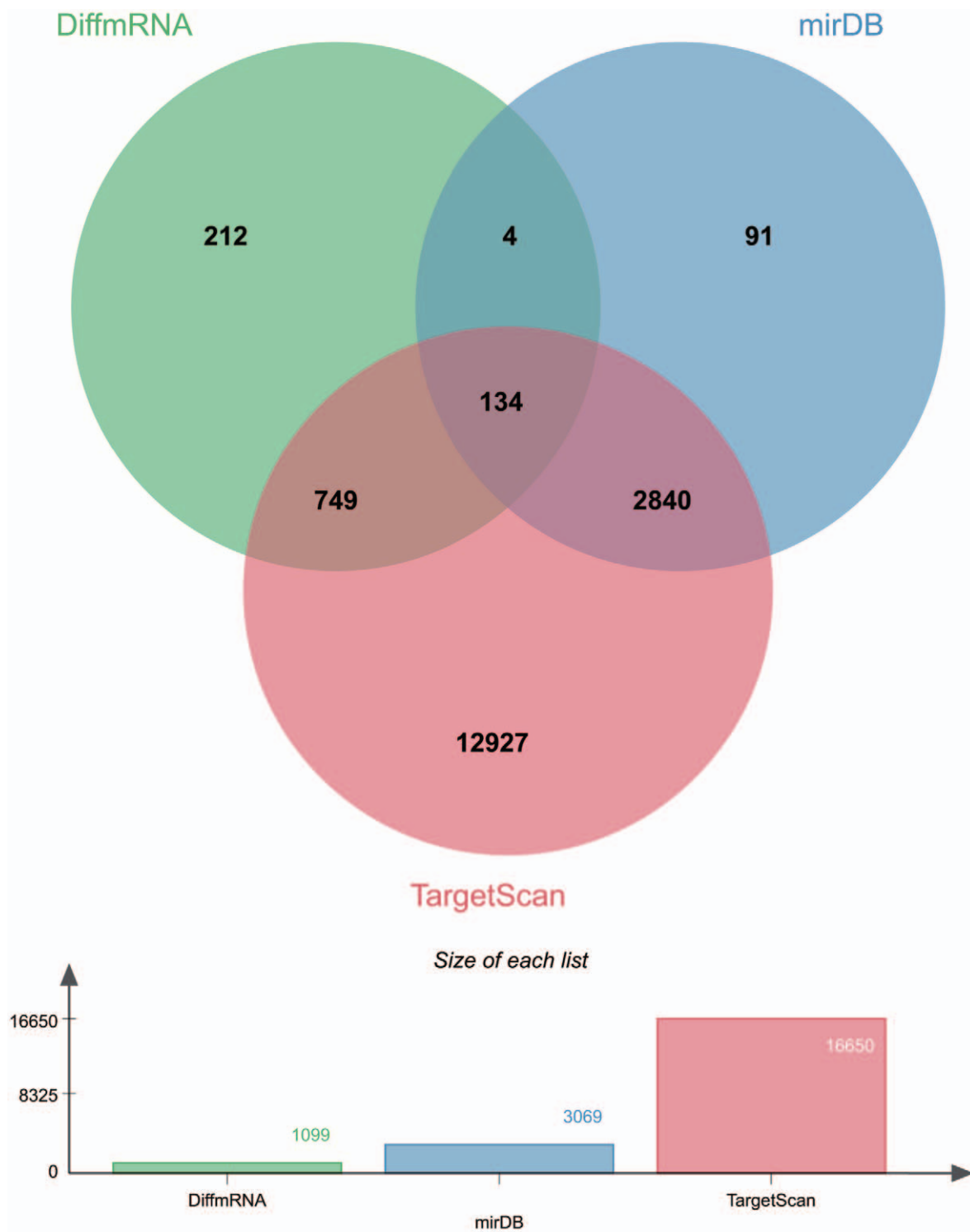
Figure 5. Venn diagram of mRNAs involved in the ceRNA regulation network. The green area shows the differential expression of mRNAs. The blue area demonstrates only the target number of 3069 mRNAs in the miRDB database. Red represents only the target number of 11,650 mRNAs in the TargetScan database, with 75 DEmiRNAs targeting 11,650 mRNAs in the TargetScan database and 3069 in the miRDB database.

4. Discussion

PTC is the most frequently occurring histological type of thyroid carcinoma. The incidence rate of PTC has suddenly increased recently, with 5% to 10% of PTC patients experiencing particularly aggressive disease progression.[3] As we cannot identify the fundamental molecular mechanisms of aggressiveness and mortality in these patients, increasing attention has been paid to identifying cancer-related genes as well as the exact regulatory mechanism of PTC initiation.

## Table 2

miRNAs that may target PTC-specific lncRNAs.

lncRNAs	miRNAs
IGF2-AS	hsa-mir-519d
LINC00302	hsa-mir-31, hsa-mir-506
AC022148.1	hsa-mir-372, hsa-mir-373, hsa-mir-144, hsa-mir-519d, hsa-mir-205, hsa-mir-221, hsa-mir-222, hsa-mir-31
LINC00313	hsa-mir-372, hsa-mir-373, hsa-mir-187, hsa-mir-205, hsa-mir-31, hsa-mir-375
AC004832.1	hsa-mir-519d, hsa-mir-31
AC002511.1	hsa-mir-519d
AC006305.1	hsa-mir-519d, hsa-mir-221, hsa-mir-222, hsa-mir-506, hsa-mir-375
AP000525.1	hsa-mir-31
CCAT1	hsa-mir-486-5P, hsa-mir-506
AC010336.2	hsa-mir-372, hsa-mir-373, hsa-mir-144, hsa-mir-519d, hsa-mir-205, hsa-mir-31
CLDN10-AS1	hsa-mir-221, hsa-mir-222
MIR181A2HG	hsa-mir-205
LINC00365	hsa-mir-519d
LINC00457	hsa-mir-144
SFTA1P	hsa-mir-221, hsa-mir-222
LINC00475	hsa-mir-205, hsa-mir-506
LINC00423	hsa-mir-31
HOTAIR	hsa-mir-519d, hsa-mir-221, hsa-mir-222, hsa-mir-506, hsa-mir-375
HCG22	hsa-mir-31, hsa-mir-506
MIR4500HG	hsa-mir-144, hsa-mir-31
MIR205HG	hsa-mir-205, hsa-mir-221, hsa-mir-222, hsa-mir-31, hsa-mir-506
CYP1B1-AS1	hsa-mir-205
LINC00460	hsa-mir-221, hsa-mir-222
LINC00284	hsa-mir-519d, hsa-mir-205, hsa-mir-506
AL158206.1	hsa-mir-372, hsa-mir-373, hsa-mir-221, hsa-mir-222
AC068594.1	hsa-mir-221, hsa-mir-222
AC011383.1	hsa-mir-31
SYNPR-AS1	hsa-mir-375, hsa-mir-187
OPCML-IT1	hsa-mir-372, hsa-mir-373, hsa-mir-519d, hsa-mir-184, hsa-mir-205, hsa-mir-506, hsa-mir-375
AP001029.2	hsa-mir-31

Dysregulation of lncRNA expression is a primary factor in cancer development, and lncRNAs are becoming potential prognostic biomarkers in cancer.[24] Indeed, a growing number of studies have provided evidence to suggest that lncRNA dysregulation may contribute to the biological functions of cancers.[25,26] Furthermore, there is increasing evidence that by binding with miRNAs, lncRNAs act as endogenous miRNA sponges.[27,28] Zhang et al[29] demonstrated that the lncRNA

Table 3

miRNAs targeting PTC-specific mRNAs.

miRNAs	mRNAs
hsa-mir-222	PCDHAC2, NRXN1, GPM6A, C6, KCNQ3, KIAA1549L, PHEX, PCDHA1, KCNK2, PLXNC1, CYP1B1, GABRA1, C6orf118, CLVS2, NRK
hsa-mir-31	CD109, CAPN8, ST8SIA3, KIAA1549L, RPH3A, HOXC13, MUM1L1, PLAG1, CSMD1, RBFOX1, KRT6C, PAX9, CLVS2, MUC21, SALL3, DPP6, KRT6A, UCN2, LRP4
hsa-mir-519d	NR4A3, IL1RAP, PCDHA1, PCDHAC2, PRR15, SMOC2, OSR1, SLC4A4, TBC1D2, GABBR2, SRCIN1, LRP1B, TIAM1, FOXQ1, DPYSL5, EPHA5, EGR2, TRIM36, LAMP5, IGSF10, SALL3, COL19A1, MAPK4
hsa-mir-373	EPHA5, TRIM36, HS3ST4, DPYSL5, PCDHA1, TBC1D2, PLAG1, TMEM100, C6orf15, ANK2, RAB27B, CACNA1E, PCDHAC2, GRIA2, SLC45A2, RYR2, PLEKHS1, C2CD4A, GRM5, NR4A3, TIAM1, PRSS23, CXCL14, GPM6A
hsa-mir-506	DPP4, PCDHAC2, SLITRK5, EPS8, LHX2, C20orf85, GLRB, FOXQ1, IL17RD, NYAP2, MYRF, MKX, SDK1, TNFRSF11B, GRIA2, RYR2, NRCAM, CDH2, CPA3, CTSH, SLCO4C1, SLITRK4, EYA1, PCDHA1, TFCP2L1, PAPSS2, PTPRQ, KCNK2, EPHA10, SPOCK3, ULBP1, AT8L, RYR1, SDC4, GMNC, SLC5A8
hsa-mir-205	LRRK2, GLRB, MGAT3, TGFA, RUNX2, BEAN1, ADAMTS9, GPM6A, C11orf86, EPPK1, C6orf222, PTGFR, PAX9, SHISA6, DLG2
hsa-mir-144	GRM5, WIF1, EPHA3, EPHA5, NRP2, GRIK3, PLXNC1, GDF10, SLITRK4, ACBD7, GABRA1, GRHL3, CDKN2B, NYAP2, ALDH1A3, FN1, CDH6
hsa-mir-221	PCDHA1, NRK, C6, C6orf118, PHEX, CLVS2, NRXN1, PLXNC1, KIAA1549L, PCDHAC2, GPM6A, KCNK2, KCNQ3, CYP1B1, GABRA1
hsa-mir-372	GRIA2, PLAG1, KCNA1, TIAM1, RYR2, TMEM100, RAB27B, GRM5, PLEKHS1, EPHA5, PRSS23, CXCL14, CACNA1E, NR4A3, ANK2, TBC1D2, HS3ST4, TRIM36, C2CD4A, SLC45A2, GPM6A, PCDHA1, PCDHAC2, DPYSL5
hsa-mir-486-5P	DCSTAMP, STAT3, SIRT1, COX-2
hsa-mir-375	HNF1B, CLCA2
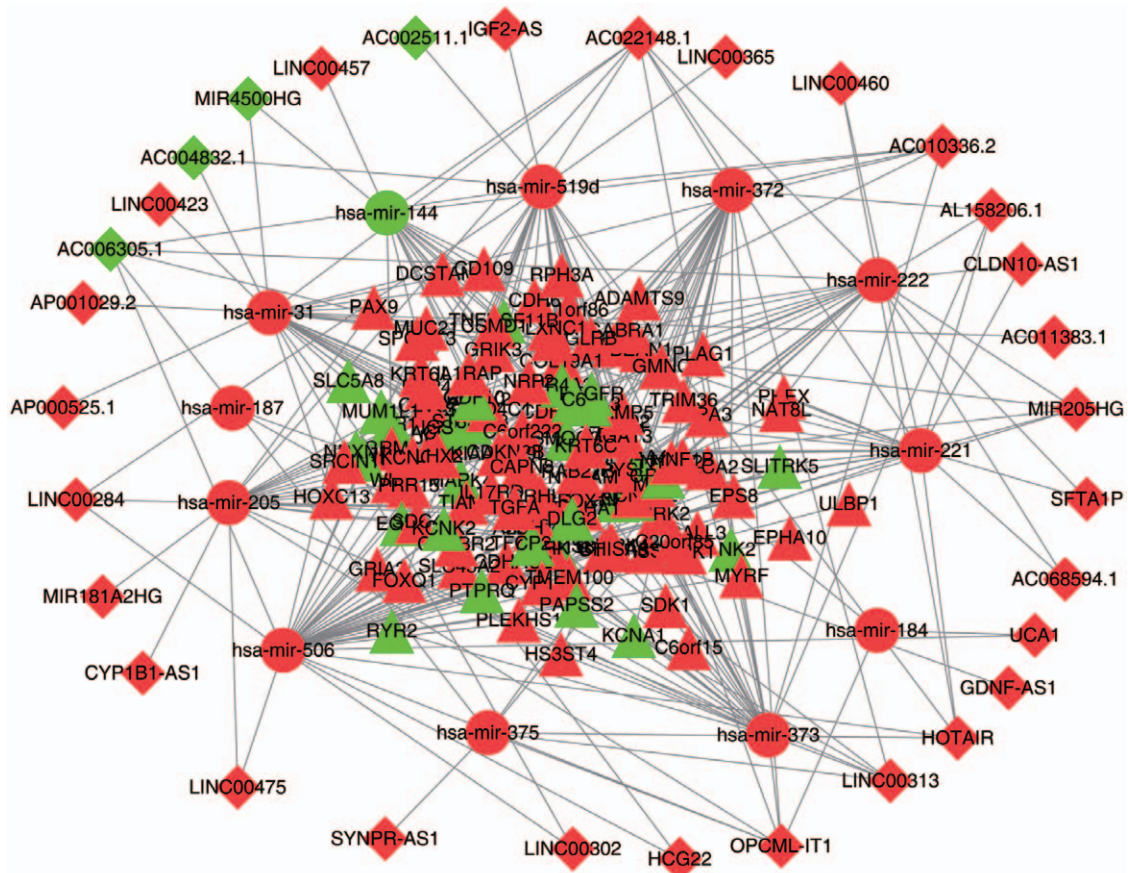
Figure 6. The dysregulated network of lncRNA-mRNA-miRNA ceRNA. lncRNA is denoted by a diamond shape, mRNA is represented by a triangle shape, and miRNA is represented by a round rectangle. Red and green shapes show upregulation and downregulation, respectively.
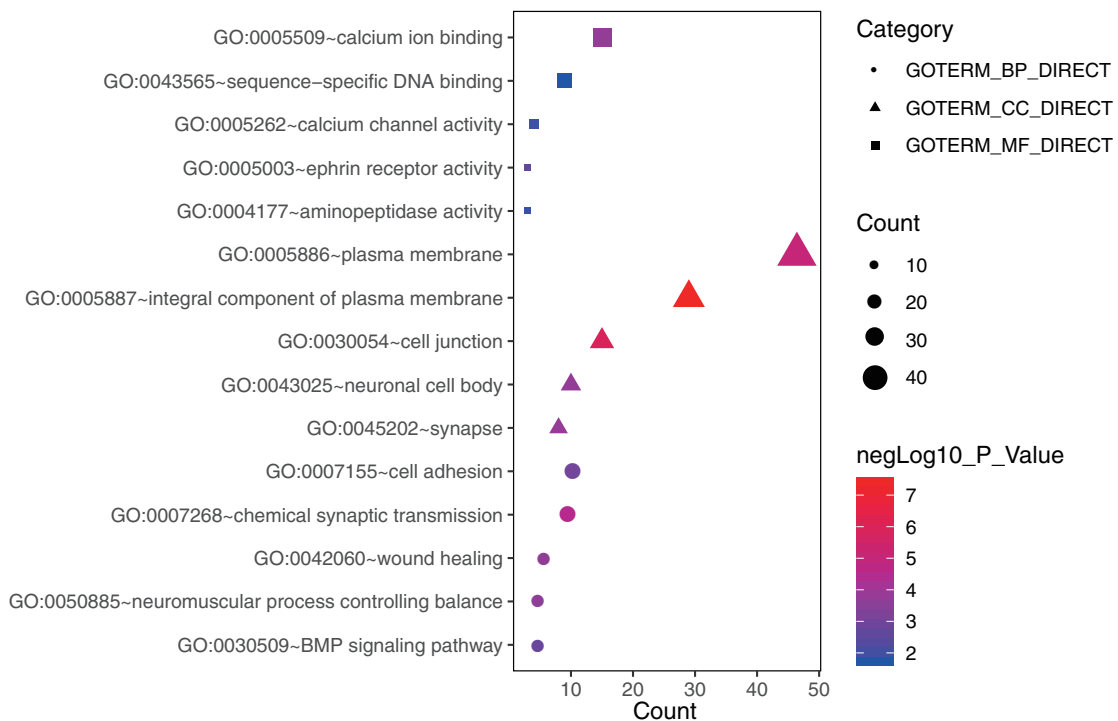
**Figure 7.** Top 5 enriched GO terms for genes of the ceRNA network.

Table 4

Top 5 Gene ontology enriched terms of targeted genes in competitive endogenous RNA crosstalk associated with PTC.

Category	Term	Count	P
GOTERM_BP_DIRECT	GO:0007268~chemical synaptic transmission	10	3.05E-05
GOTERM_BP_DIRECT	GO:0042060~wound healing	6	1.87E-04
GOTERM_BP_DIRECT	GO:0050885~neuromuscular process controlling balance	5	2.38E-04
GOTERM_BP_DIRECT	GO:0007155~cell adhesion	11	9.28E-04
GOTERM_BP_DIRECT	GO:0030509~BMP signaling pathway	5	.001609
GOTERM_CC_DIRECT	GO:0005887~integral component of plasma membrane	29	3.91E-08
GOTERM_CC_DIRECT	GO:0030054~cell junction	15	1.19E-06
GOTERM_CC_DIRECT	GO:0005886~plasma membrane	48	8.68E-06
GOTERM_CC_DIRECT	GO:0045202~synapse	8	1.51E-04
GOTERM_CC_DIRECT	GO:0043025~neuronal cell body	10	1.74E-04
GOTERM_MF_DIRECT	GO:0005509~calcium ion binding	15	1.72E-04
GOTERM_MF_DIRECT	GO:0005003~ephrin receptor activity	3	.002109
GOTERM_MF_DIRECT	GO:0005262~calcium channel activity	4	.009965
GOTERM_MF_DIRECT	GO:0004177~aminopeptidase activity	3	.0126
GOTERM_MF_DIRECT	GO:0043565~sequence-specific DNA binding	9	.017185

UCA1 promotes cancer progression by functioning as a ceRNA of activating transcription factor 2 in prostate cancer. In addition, Xia et al[30] demonstrated that the lncRNA Fer-1-like family member 4 suppresses the growth of cancer cell by serving as a ceRNA and modulating phosphatase and tensin homolog expression. However, only limited research has described the interaction among mRNAs and lncRNAs in PTC, with the results indicating that lncRNAs are a vital component of the ceRNA network, and such networks have been inadequately described.[31,32]

In this study, utilizing RNA-seq data downloaded from TCGA, DEmiRNAs, DEmRNAs, and DElncRNAs were identified in PTC samples, with 75 miRNAs, 494 lncRNAs, and 1099 mRNAs DE. A lncRNA-mRNA-miRNA ceRNA network of aberrant molecules was established for PTC; this network consists of 13 DEmiRNAs, 31 DElncRNAs, and 134 DEmRNAs. GO analysis as well as KEGG pathway analysis were employed to assess biological functions enriched among the DE coding genes. RNA-related GO analysis results showed significant ($P < .05$) enrichment of 120 GO terms in the BP category. These important
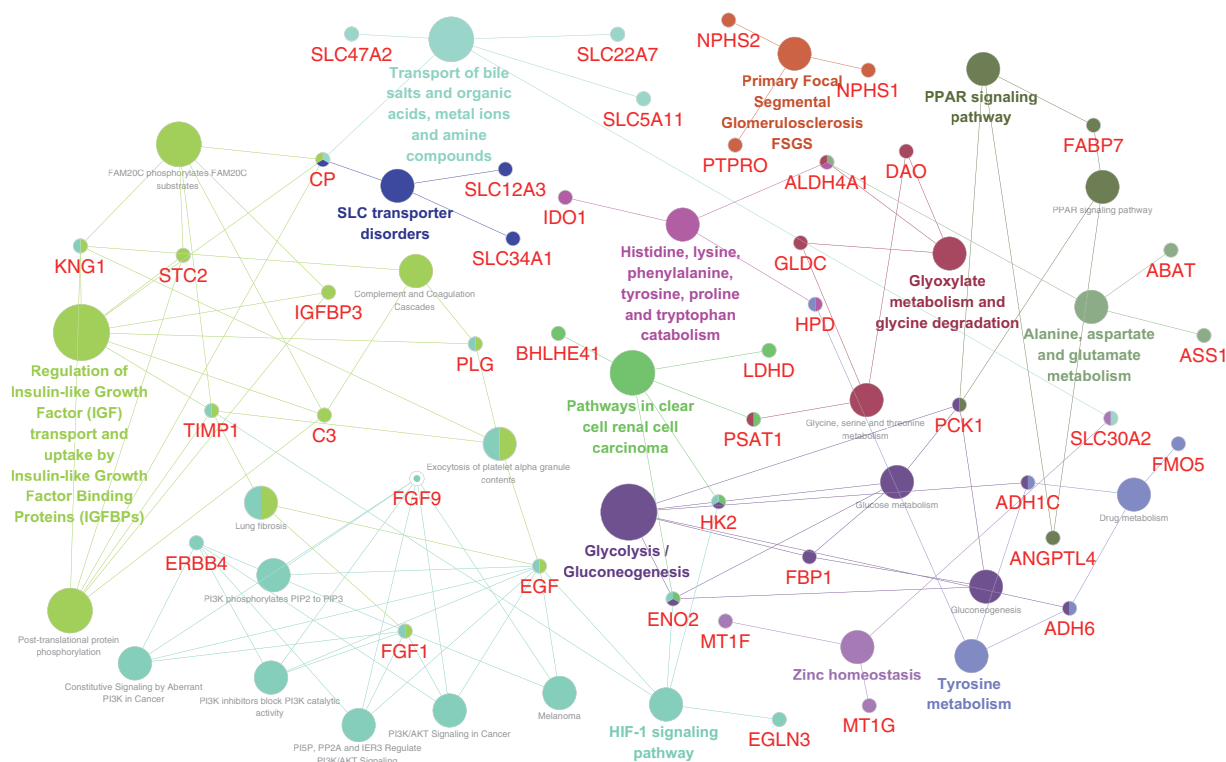

Figure 8. KEGG pathways plug-in ClueGO, with Cytoscape used to visualize the interaction network. (Circles of the same color represent enriched pathways and their corresponding genes. The lines also show the interaction between the enriched pathways and their corresponding genes).

Table 5

KEGG analysis of targeted genes in competitive endogenous RNA crosstalk associated with THCA.

Ontology source	GO term	P	Genes
KEGG	Alanine, aspartate, and glutamate metabolism	.00	ABAT, ALDH4A1, ASS1
KEGG	Drug metabolism	.01	ADH1C, ADH6, FMO5
KEGG	Glycine, serine, and threonine metabolism	.00	DAO, GLDC, PSAT1
KEGG	Glycolysis/Gluconeogenesis	.00	ADH1C, ADH6, ENO2, FBP1, HK2, PCK1
KEGG	HIF-1 signaling pathway	.00	EGF, EGLN3, ENO2, HK2, TIMP1
KEGG	Melanoma	.01	EGF, FGF1, FGF9
KEGG	PPAR signaling pathway	.01	ANGPTL4, FABP7, PCK1
KEGG	Tyrosine metabolism	.00	ADH1C, ADH6, HPD
REACTOME	Constitutive Signaling by Aberrant PI3K in Cancer	.00	EGF, ERBB4, FGF1, FGF9
REACTOME	Glyoxylate metabolism and glycine degradation	.00	ALDH4A1, DAO, GLDC
REACTOME	Histidine, lysine, phenylalanine, tyrosine, proline, and tryptophan catabolism	.00	ALDH4A1, HPD, IDO1
REACTOME	PI3K inhibitors block PI3K catalytic activity	.00	EGF, ERBB4, FGF1, FGF9
REACTOME	SLC transporter disorders	.01	CP, SLC12A3, SLC34A1
REACTOME	Transport of bile salts and organic acids, metal ions, and amine compounds	.00	CP, SLC22A7, SLC30A2, SLC47A2, SLC5A11
REACTOME	Exocytosis of platelet alpha granule contents	.00	EGF, KNG1, PLG, TIMP1
Wiki Pathways	Complement and Coagulation Cascades	.01	C3, KNG1, PLG
Wiki Pathways	Lung fibrosis	.01	EGF, FGF1, TIMP1
Wiki Pathways	Pathways in clear cell renal cell carcinoma	.00	BHLHE41, ENO2, HK2, LDHD, PSAT1
Wiki Pathways	Primary Focal Segmental Glomerulosclerosis FSGS	.01	NPHS1, NPHS2, PTPRO
Wiki Pathways	Zinc homeostasis	.00	MT1F, MT1G, SLC30A2

GO terms are related to reactions to the plasma membrane, an essential constituent of the plasma membrane, cell junction, cell adhesion, and neuronal cell body. Pathway analysis further showed enrichment of 87 pathways, mainly involving the PPAR signaling pathway, the HIF-1 signaling pathway, PI3K inhibitors blocking PI3K catalytic activity, and lung fibrosis.

In addition, survival analysis revealed that *CCAT1, SYNPR, SFTA1P, HOTAIR, HCG22,* and *CLDN10* correlate with OS. These genes are involved in the progression of various cancers. CCAT1 is a prognostic biomarker for thyroid cancer.[33] Studies have shown that in tumorigenesis, HOTAIR serves as an oncogene that can facilitate the invasion and metastasis of
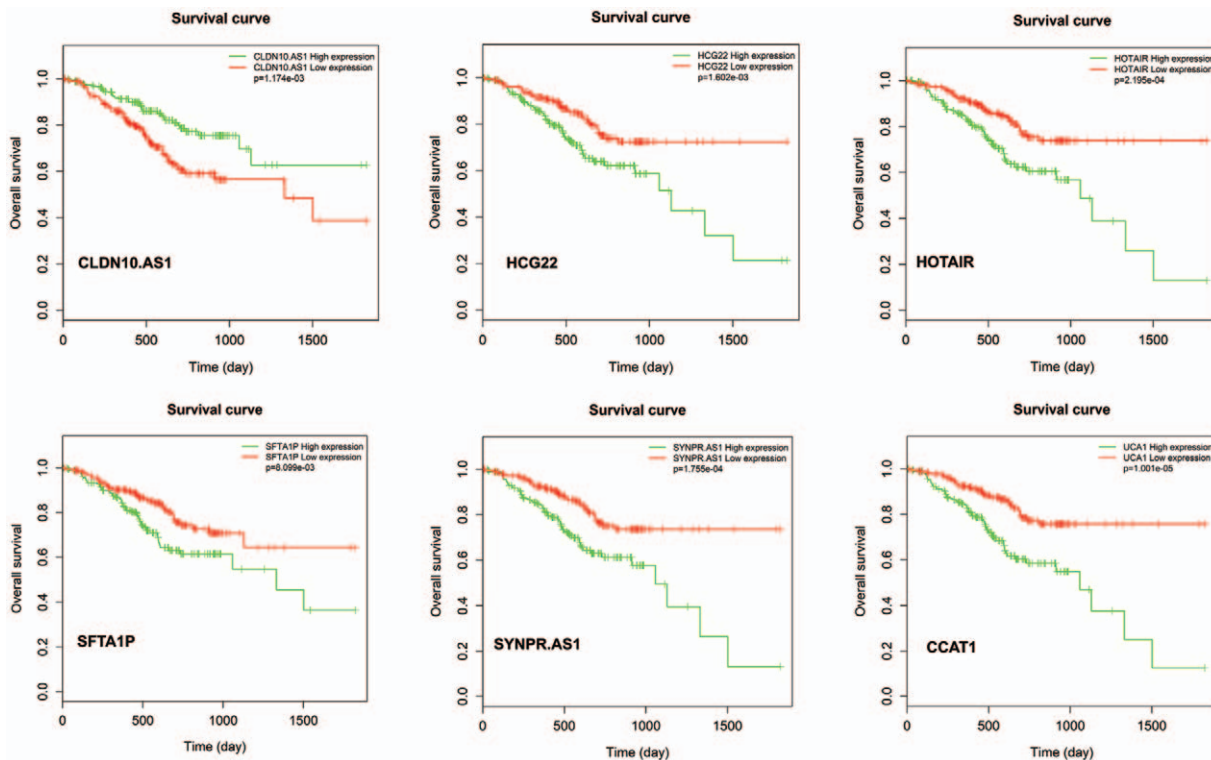

Figure 9. Kaplan–Meier survival curves for 6 lncRNAs associated with overall survival. The horizontal axis represents overall survival time, days; the vertical axis represents the survival function.

Table 6

Univariate and multivariate Cox regression analysis of 6 significantly lncRNAs associated with overall survival in TCGA THCA dataset.

	Univariate regression model			Multivariate regression model		
	HR	P	95% CI	HR	P	95% CI
Test set (n = 379)						
HCG22	1.395253	.011[*]	1.082–1.799	1.356911	.02[*]	1.048–1.756
SYNPR.AS1	0.342737	.038[*]	0.121–0.973	0.329204	.044[*]	0.115–0.941
SFTA1P	1.395908	.028[*]	1.037–1.879	1.341509	.053	0.996–1.807
HOTAIR	1.657563	.021[*]	1.077–2.552	1.510335	.050	1.2282–1.817
UCA1	0.675167	.025[*]	0.479–0.952	0.673215	.027[*]	0.473–0.958
CLDN10.AS1	1.427242	.003[*]	1.123–1.813	1.428991	.004[*]	1.124–1.817

[*] P value less than .05.

A

B
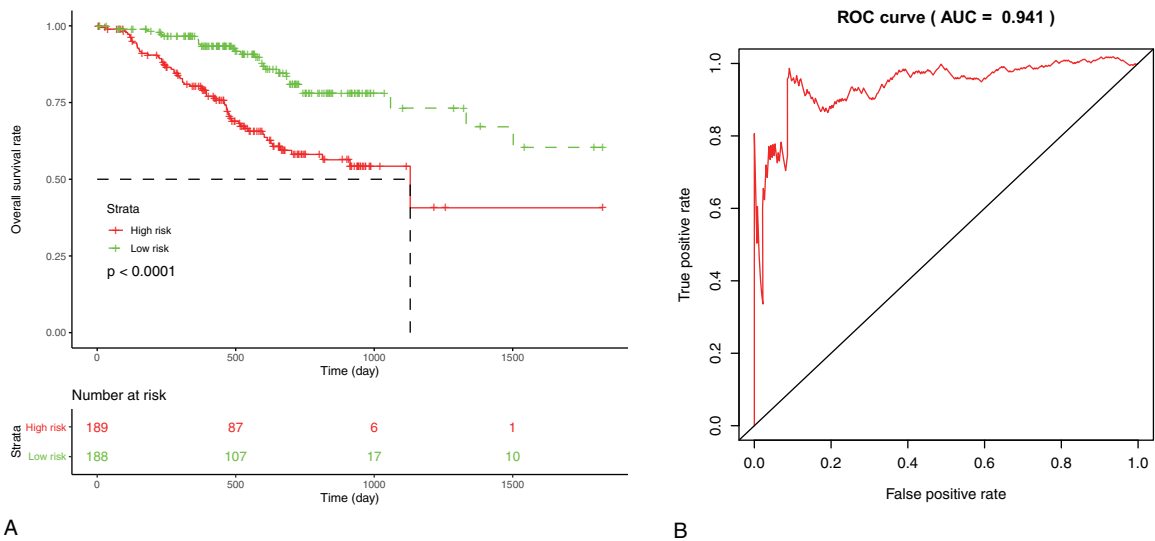
Figure 10. The 6-lncRNA signature demonstrates the prognostic efficiency in the PTC training cohort from TCGA. (A) Survival analysis including the high-risk group and low-risk group using Kaplan–Meier curves. (B) Receiver operating characteristic analysis of sensitivity and specificity by risk score for predicting OS.
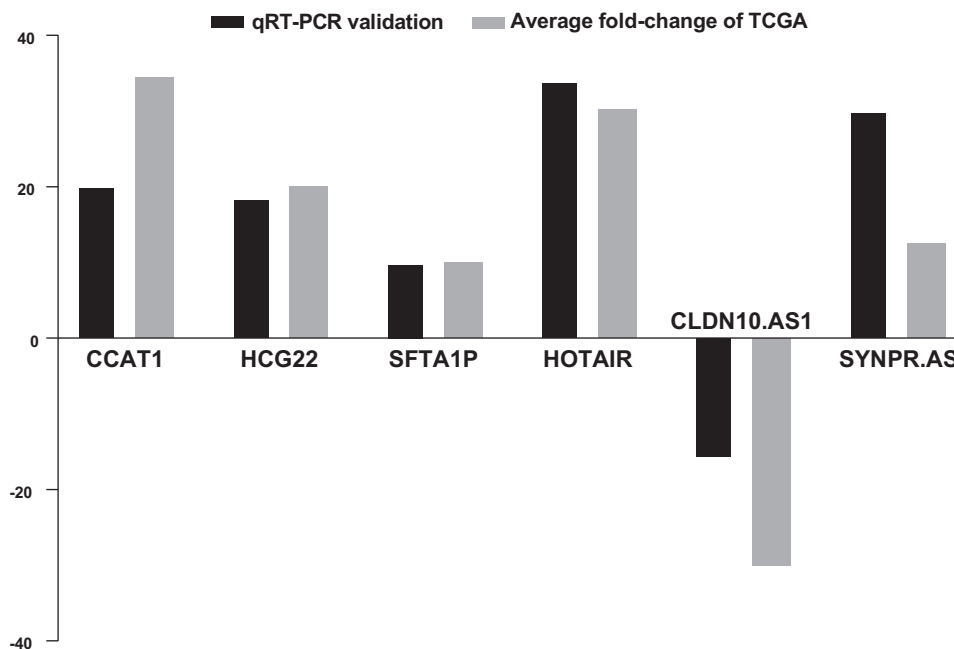
Figure 11. qRT-PCR validation of 6 differentially expressed key lncRNAs. Fold change (2-ΔΔCt) of lncRNAs between TCGA and qRT-PCR results is shown.

PTC.[34,35] HOTAIR is also a prognostic factor for survival in glioma, breast cancer, and colon cancer patients, and increased expression of HOTAIR is associated with enhanced metastasis.[36,37] The function of other lncRNAs remains unclear, though studies have illustrated that SFTA1P is overexpressed in gastric cancer,[38] lung squamous cell carcinoma,[39] and colorectal cancer.[40] Moreover, CLDN10 is involved in the malignant phenotype of osteosarcoma cells,[41] correlates with breast cancer patient survival,[42] and is DE in ovarian cancer.[43] In the current study, these genes were discovered to be DE in PTC. By using qRT-PCR, we selected 6 cancer-specific lncRNAs crucial for survival to verify expression profiles in 60 PTC recently diagnosed patients. The expression results between qRT-PCR and TCGA were 100% in agreement. This study offers data for a better understanding of the lncRNA-miRNA-mRNA ceRNA network and for elucidating the molecular mechanisms of PTC.

However, there are some limitations in this study. First, some important clinical variables were not available from public datasets at the time of the prognostic model construction. Future studies should include more clinical variables. Second, the potential mechanism underlying the lncRNA prognostic capacity in PTC remains to be explored.

In summary, we generated a lncRNA-miRNA-mRNA ceRNA network to clarify the unknown ceRNA regulatory network in PTC. The results of our experiments illustrate that lncRNAs play a crucial role in PTC development; six lncRNAs may be related to OS in PTC. This network is a valuable tool for identifying lncRNAs related to the OS of PTC patients. In addition, we utilized qRT-PCR verification as well as bioinformatics analysis to confirm the validity and reliability of the expression level of key lncRNAs. Further studies are needed to investigate the biological mechanisms of the six lncRNAs in PTC. Our results offer a new perspective for a better understanding of the lncRNA-related ceRNA network and provide information for the development of a promising prognostic tool for determining OS in PTC.

Author contributions

Data curation: Yanxia Jiang, Jiao Wang;
Methodology: Yanxia Jiang, Jiangchen Wang;
Supervision: Yanxia Jiang;
Writing – original draft: Yanxia Jiang, Jiang Li;
Writing – review & editing: Jixiong Xu.

References

[1] Haugen BR, Sawka AM, Alexander EK, et al. American Thyroid Association Guidelines on the Management of Thyroid Nodules and Differentiated Thyroid Cancer Task Force Review and Recommendation on the Proposed Renaming of Encapsulated Follicular Variant Papillary Thyroid Carcinoma Without Invasion to Noninvasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features. Thyroid 2017;27: 481–3.

[2] Ye W, Deng X, Fan Y. Exosomal miRNA423-5p mediated oncogene activity in papillary thyroid carcinoma: a potential diagnostic and biological target for cancer therapy. Neoplasma 2019;66:516–23.

[3] Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 2016;26:1–33.

[4] Guilmette J, Nose V. Hereditary and familial thyroid tumours. Histopathology 2018;72:70–81.

[5] Liang WQ, Zeng , Chen CF, et al. Long noncoding RNA H19 is a critical oncogenic driver and contributes to epithelial-mesenchymal transition in papillary thyroid carcinoma. Cancer Manag Res 2019;11:2059–72.

[6] Kaiser MI. ENCODE and the parts of the human genome. Stud Hist Philos Biol Biomed Sci 2018;72:28–37.

[7] Achour C, Aguilo F. Long non-coding RNA and Polycomb: an intricate partnership in cancer biology. Front Biosci (Landmark Ed) 2018;23: 2106–32.

[8] Chen F, Yin S, Zhu J, et al. lncRNA DGCR5 acts as a tumor suppressor in papillary thyroid carcinoma via sequestering miR-2861. Exp Ther Med 2019;17:895–900.

[9] Li X, Zhong W, Xu Y, et al. Silencing of lncRNA LINC00514 inhibits the malignant behaviors of papillary thyroid cancer through miR-204-3p/CDC23 axis. Biochem Biophys Res Commun 2019;508:1145–8.

[10] Feng L, Yang B, Tang XD. Long noncoding RNA LINC00460 promotes carcinogenesis via sponging miR-613 in papillary thyroid carcinoma. J Cell Physiol 2019;234:11431–9.

[11] Jiang W, Zhan H, Jiao Y, et al. A novel lncRNA-miRNA-mRNA network analysis identified the hub lncRNA RP11-159F24.1 in the pathogenesis of papillary thyroid cancer. Cancer Med 2018;7:6290–8.

[12] Song J, Ye A, Jiang E, et al. Reconstruction and analysis of the aberrant lncRNA-miRNA-mRNA network based on competitive endogenous RNA in CESC. J Cell Biochem 2018;119:6665–73.

[13] Sui J, Li YH, Zhang YQ, et al. Integrated analysis of long non-coding RNAassociated ceRNA network reveals potential lncRNA biomarkers in human lung adenocarcinoma. Int J Oncol 2016;49:2023–36.

[14] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–40.

[15] Jeggari A, Marks DS, Larsson E. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. Bioinformatics 2012;28:2062–3.

[16] Agarwal V, Bell GW, Nam JW, et al. Predicting effective microRNA target sites in mammalian mRNAs. Elife 2015;4:

[17] Wang X. miRDB: a microRNA target prediction and functional annotation database with a wiki interface. RNA 2008;14:1012–7.

[18] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.

[19] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16: 284–7.

[20] Wang G, Chen H, Liu J. The long noncoding RNA LINC01207 promotes proliferation of lung adenocarcinoma. Am J Cancer Res 2015;5:3162–73.

[21] Meng J, Li P, Zhang Q, et al. A four-long non-coding RNA signature in predicting breast cancer survival. J Exp Clin Cancer Res 2014;33:84.

[22] Zhu X, Tian X, Yu C, et al. A long non-coding RNA signature to improve prognosis prediction of gastric cancer. Mol Cancer 2016;15:60.

[23] Mihaly Z, Kormos M, Lanczky A, et al. A meta-analysis of gene expression-based biomarkers predicting outcome after tamoxifen treatment in breast cancer. Breast Cancer Res Treat 2013;140: 219–32.

[24] Chen J, Hu L, Chen J, et al. Detection and analysis of Wnt pathway related lncRNAs expression profile in lung adenocarcinoma. Pathol Oncol Res 2016;22:609–15.

[25] Shen X, Zhang Y, Wu X, et al. Upregulated lncRNA-PCAT1 is closely related to clinical diagnosis of multiple myeloma as a predictive biomarker in serum. Cancer Biomark 2017;18:257–63.

[26] Sedlarikova L, Gromesova B, Kubaczkova V, et al. Deregulated expression of long non-coding RNA UCA1 in multiple myeloma. Eur J Haematol 2017;99:223–33.

[27] Li F, Huang C, Li Q, et al. Construction and comprehensive analysis for dysregulated long non-coding RNA (lncRNA)-associated competing endogenous RNA (ceRNA) network in gastric cancer. Med Sci Monit 2018;24:37–49.

[28] Zhao Y, Wang H, Wu C, et al. Construction and investigation of lncRNA-associated ceRNA regulatory network in papillary thyroid cancer. Oncol Rep 2018;39:1197–206.

[29] Zhang S, Dong X, Ji T, et al. Long non-coding RNA UCA1 promotes cell progression by acting as a competing endogenous RNA of ATF2 in prostate cancer. Am J Transl Res 2017;9:366–75.

[30] Xia T, Liao Q, Jiang X, et al. Long noncoding RNA associated-competing endogenous RNAs in gastric cancer. Sci Rep 2014;4:6088.

[31] Chen S, Fan X, Gu H, et al. Competing endogenous RNA regulatory network in papillary thyroid carcinoma. Mol Med Rep 2018;18: 695–704.

[32] Chu A, Liu J, Yuan Y, et al. Comprehensive analysis of aberrantly expressed ceRNA network in gastric cancer with and without H.pylori infection. J Cancer 2019;10:853–63.

[33] Yang T, Zhai H, Yan R, et al. lncRNA CCAT1 promotes cell proliferation, migration, and invasion by down-regulation of miR-143 in FTC-133 thyroid carcinoma cell line. Braz J Med Biol Res 2018;51:e7046.

[34] Li HM, Yang H, Wen DY, et al. Overexpression of LncRNA HOTAIR is associated with poor prognosis in thyroid carcinoma: a study based on TCGA and GEO data. Horm Metab Res 2017;49:388–99.

[35] Zhu H, Lv Z, An C, et al. Onco-lncRNA HOTAIR and its functional genetic variants in papillary thyroid carcinoma. Sci Rep 2016;6:31969.

[36] Mahmoudian-Sani MR, Jalali A, Jamshidi M, et al. Long non-coding RNAs in thyroid cancer: implications for pathogenesis, diagnosis, and therapy. Oncol Res Treat 2019;42:136–42.

[37] Zhang Y, Yu S, Jiang L, et al. HOTAIR is a promising novel biomarker in patients with thyroid cancer. Exp Ther Med 2017;13:2274–8.

[38] Ma H, Ma T, Chen M, et al. The pseudogene-derived long non-coding RNA SFTA1P suppresses cell proliferation, migration, and invasion in gastric cancer. Biosci Rep 2018;38:

[39] Li L, Yin JY, He FZ, et al. Long noncoding RNA SFTA1P promoted apoptosis and increased cisplatin chemosensitivity via regulating the hnRNP-U-GADD45A axis in lung squamous cell carcinoma. Oncotarget 2017;8:97476–89.

[40] Zhao J, Xu J, Shang AQ, et al. A six-LncRNA expression signature associated with prognosis of colorectal cancer patients. Cell Physiol Biochem 2018;50:1882–90.

[41] Zhang X, Wang X, Wang A, et al. CLDN10 promotes a malignant phenotype of osteosarcoma cells via JAK1/Stat1 signaling. J Cell Commun Signal 2019;13:395–405.

[42] Liao J, Li J, Cheng H, et al. CLDN10 single nucleotide polymorphism rs1325774 alters the risk of breast cancer in south chinese women. Medicine (Baltimore) 2018;97:e13187.

[43] Gao Y, Liu X, Li T, et al. Cross-validation of genes potentially associated with overall survival and drug resistance in ovarian cancer. Oncol Rep 2017;37:3084–92.