Identification of RNA Expression Profiles in Thyroid Cancer to Construct a Competing Endogenous RNA (ceRNA) Network of mRNAs, Long Noncoding RNAs (lncRNAs), and microRNAs (miRNAs)

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

ABC 1 Yuanxin Xu
DE 2 Jiuwei Chen
F 3 Zhihui Yang
G 4 Lihua Xu

1 Department of Endocrinology, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, P.R. China
2 Harbin Medical University, Harbin, Heilongjiang, P.R. China
3 Department of Public Health Sciences, Stockholm University, Stockholm, Sweden
4 Department of Nursing, The Hospital of Heilongjiang Province, Harbin, Heilongjiang, P.R. China

Corresponding Author: Lihua Xu, e-mail: Lihua_xu63@outlook.com
Source of support: Departmental sources

Background: The aims of this study were to use RNA expression profile bioinformatics data from cases of thyroid cancer from the Cancer Genome Atlas (TCGA), the Kyoto Encyclopedia of Genes and Genomes (KEGG), and the Gene Ontology (GO) databases to construct a competing endogenous RNA (ceRNA) network of mRNAs, long noncoding RNAs (lncRNAs), and microRNAs (miRNAs).

Material/Methods: TCGA provided RNA profiles from 515 thyroid cancer tissues and 56 normal thyroid tissues. The DESeq R package analyzed high-throughput sequencing data on differentially expressed RNAs. GO and KEGG pathway analysis used the DAVID 6.8 and the ClusterProfile R package. Kaplan-Meier survival statistics and Cox regression analysis were performed. The thyroid cancer ceRNA network was constructed based on the miRDB, miRTarBase, and TargetScan databases.

Results: There were 1,098 mRNAs associated with thyroid cancer; 101 mRNAs were associated with overall survival (OS). Multivariate analysis developed a risk scoring system that identified seven signature mRNAs, with a discriminative value of 0.88, determined by receiver operating characteristic (ROC) curve analysis. A ceRNA network included 13 mRNAs, 31 lncRNAs, and seven miRNAs. Four out of the 31 lncRNAs and all miRNAs were down-regulated, and the remaining RNAs were upregulated. Two lncRNAs (MIR1281A2HG and OPCML-IT1) and one miRNA (miR-184) were significantly associated with OS in patients with thyroid cancer.

Conclusions: Differential RNA expression profiling in thyroid cancer was used to construct a ceRNA network of mRNAs, lncRNAs, and miRNAs that showed potential in evaluating prognosis.

MeSH Keywords: Carcinogenesis • Gene Expression Profiling • MicroRNAs • RNA, Long Noncoding • Thyroid Neoplasms

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/912450
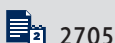
📄 2705   🗎 8   📊 7   📖 26

Background

Thyroid cancer is the most common primary endocrine malignancy. During the past 30 years, the reported incidence of thyroid cancer has increased, which is partly due to improvements in cancer screening and detection, but there has also been an increasing incidence [1,2]. In 2017, thyroid cancer represented approximately 3.4% of new cases of cancer in the USA [3].

Primary malignancy of the thyroid gland originates from para-follicular C-cells and follicular cells, with follicular-cell derived cancer being more common. Medullary thyroid cancer (MTC) is usually a low-grade malignancy that accounts for 5–10% of primary thyroid cancer and is the only thyroid malignancy that originates from parafollicular C-cells. Follicular cell-derived thyroid cancer can be divided into differentiated thyroid cancer (DTC), poorly differentiated thyroid cancer (PDTC) and anaplastic thyroid cancer (ATC). ATC and PDTC are uncommon forms of thyroid cancer, representing less than 5% and 6.7% of thyroid cancer respectively, but typically exhibit aggressive clinical behavior with poor prognosis. Of the DTCs, papillary thyroid cancer (PTC) is the most common form, representing more than 80% of all cases of thyroid cancer, while follicular thyroid cancer (FTC) is the second most common thyroid cancer accounting for 10% of cases. DTC generally exhibits far less aggressive behavior than PDTC and ATC, and patients usually have a good prognosis although rapid progression and the poor clinical outcome can occur in some cases. Clinical diagnostic challenges also exist for the DTCs as fine-needle aspiration biopsy (FNAB) and ultrasonography are not sufficient to distinguish this tumor from benign thyroid nodules, which may lead to overdiagnosis and overtreatment. Therefore, the identification of new molecular biomarkers for the diagnosis and prognosis of thyroid cancer is important to improve treatment strategies, including for patients with DTC [4].

Although thyroid cancer is a multifactorial disease, inherited and acquired genetic alterations play a critical role in the in the development of this cancer and have been extensively studied [5]. Some of the most widely studied gene changes include mutations in the BRAF gene, RET/PTC gene rearrangements, RAS gene mutations, and PAX8-PPARγ gene rearrangement [6–8].

With the emergence of high-throughput sequencing methods, approximately 20,000 pseudogenes have been identified [8,9]. Long non-coding RNA (lncRNA) and microRNA (miRNA) may act as pseudogenes, by modifying and regulating gene expression. Recently, lncRNAs and miRNAs have been shown to have a role in many types of cancer, including thyroid cancer. For example, lncRNA AB074169 was reported to exhibit tumor suppressive properties by targeting KHSRP-mediated p21 expression [10] and lncRNA GAS5 was found to act as a competing endogenous RNA (ceRNA) to modulate PTEN expression by targeting

miR-222-3p in papillary thyroid cancer [11]. Therefore, the hypothesis that drove the present study was that a network consisting of mRNA, lncRNA, and miRNA might have a role in the pathogenesis of thyroid cancer.

Therefore, the aims of this study were to use RNA expression profile bioinformatics data from cases of thyroid cancer from the Cancer Genome Atlas (TCGA), the Kyoto Encyclopedia of Genes and Genomes (KEGG), and the Gene Ontology (GO) databases to construct a ceRNA network of mRNAs, lncRNAs, and miRNAs. The study also included survival analysis and KEGG functional analysis to provide new understanding of the role of the ceRNA network in the pathogenesis and progression of thyroid cancer.

Material and Methods

The Cancer Genome Atlas (TCGA) patient and sample information

RNA sequencing data of 515 cases of thyroid cancer and 56 normal thyroid tissues were downloaded from The Cancer Genome Atlas (TCGA) on April 2018. The Illumina HiSeqRNASeq and Illumina HiSeqmiRNASeq platforms were used to generate the mRNA and microRNA (miRNA) expression profiles of each sample. The National Cancer Institute (NCI) Genomic Data Commons (GDC) Data Transfer Tool (*https://gdc.cancer.gov/access-data/gdc-data-transfer-tool*) was used to download clinical information and the gene expression profiles from level 3 mRNA-Seq and miRNA-Seq. The study met the publication guidelines provided by TCGA (*https://cancergenome.nih.gov/publications/publicationguidelines*). No ethical approval was required for the study as all patient data were acquired from TCGA.

Identification of differentially expressed RNA

To identify the differentially expressed mRNAs, long non-coding RNAs (lncRNAs), and miRNAs in thyroid cancer compared with normal tissues, the DESeq R package analyzed high-throughput sequencing data on differentially expressed RNAs [12], R software was utilized with thresholds set at log^2 fold change >2, and a false discovery rate (FDR) or adjusted P-value <0.01. ENSEMBL (*https://www.ensembl.org/*) was used to annotate the differentially expressed mRNAs and lncRNAs.

Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis

Analysis of the Gene Ontology (GO) database was performed to further study the potential biological processes and pathways of the differentially expressed RNAs. The Annotate, Visualize and Integrate Discovery Database (DAVID 6.8) (*http://david.abcc.*

ncifcrf.gov/) [12] was used for GO analysis, with the significance level set at a FDR <0.05. KEGG pathway analysis was performed with the KEGG Oncology-Based Annotation System 3.0 (KOBAS 3.0) (*kobas.cbi.pku.edu.cn/*) with the significance level set at an adjusted P-value <0.05. To generate the data plot, the GOplot package in R was used. Cytoscape version 3.6.1 [13] was used for assembly and visualization of the network.

Construction of the lncRNA, miRNA, and mRNA ceRNA network

A ceRNA network was constructed, as lncRNAs are believed to be able to adhere to miRNAs and prevent them from binding their target genes [11]. To change the miRNA sequences, the starBase version 2.0 database (*http://starbase.sysu.edu.cn/*) was used. The MIRanda database (*http://www.microrna.org/*) was used to predict the interactions between lncRNAs and miRNAs. The target genes of miRNAs were predicted with miRDB (*http://www.mirdb.org/*), miRTarBase (*http://mirtarbase.mbc.nctu.edu.tw/*) [14], and TargetScan (*http://www.targetscan.org/*) databases. Also, differentially expressed miRNA were used to select the intersecting lncRNAs and mRNAs. The ceRNA network consisting of lncRNA, miRNA, and mRNA was constructed and visualized with Cytoscape version 3.6.1. The significance level was set at log2 fold change >2, and P<0.05.

Survival analysis and risk scoring

The differentially expressed mRNAs underwent Kaplan-Meier analysis and the log-rank test to identify the mRNAs associated with to overall survival (OS). P<0.05 was considered as statically significant. Univariate Cox regression analysis was applied, where P<0.01 was considered as statistically significant. The mRNAs that were significantly associated with OS in univariate Cox regression were further analyzed using a multivariate Cox regression model. A risk scoring system was established following the results of the multivariate Cox regression model where patients were grouped into a low-risk and a high-risk group, based on the median risk score calculation. Kaplan-Meier survival analysis with the log-rank test and receiver operating characteristic (ROC) analysis was performed to assess the results of the risk scoring system. All survival analysis was performed in R using the R package ('survival').

Results

Characteristics of the patients with thyroid cancer from the Cancer Genome Atlas (TCGA) database

The Cancer Genome Atlas (TCGA) provided clinical information for 507 of 515 patients with thyroid cancer. The clinical features, including the histopathology, tumor grade and stage, patients age and gender are summarized in Table 1. The median patient age was 46 years (range, 15–89 years). The histological subtypes of thyroid cancer including follicular carcinoma, classical papillary carcinoma, follicular variant of papillary carcinoma, tall cell papillary carcinoma and other not specified subtypes, which were included in the TCGA dataset, although the majority of cases were classical papillary thyroid carcinoma. The median follow-up time was 9.46 years (range, 0–54.23 years).

Differentially expressed mRNAs from the Gene Ontology (GO) database

A total of 1,098 mRNAs were considered to be differentially expressed, of which 234 (21.3%) were down-regulated, while 864 (78.7%) were upregulated. The mRNA expression profiles are visualized using volcano plots (Supplementary Figure 1). To investigate the biological mechanisms of these differentially expressed genes, the Gene Ontology (GO) database analysis was conducted. GO enrichment analysis showed that 19 significant functions were involved through these mRNAs (P<0.01) (Figure 1A). Thirty mRNAs were enriched in the plasma membrane, which represented the majority in the analysis. Based on the z-scores, the extracellular region was dominated by upregulated mRNAs, while the external side of the plasma membrane was dominated by downregulated mRNAs (Figure 1B). All enriched mRNA expression relationships are demonstrated in the plot shown in Figure 1C.

Prognosis-related mRNAs

Patients were divided into low-expression and high-expression clusters based on the median differentially expressed mRNAs. As a result, a total of 101 mRNAs were identified as related to overall survival by Kaplan-Meier survival analysis (P<0.05) (Table 2). Sixteen of these mRNAs were verified with univariate Cox regression (Table 3). Multivariate Cox regression analysis identified a seven-signature mRNA expression profile predictive of overall survival (OS) (Table 4, Figure 2A). Based on these results, a risk model was established, which showed that the seven signature mRNAs were significantly associated with OS (P<0.05) (Figure 2B). The discriminative value of the risk scoring system was determined with receiver operating characteristic (ROC) analysis, with an area under the curve (AUC) discriminative value of 0.88 (Figure 2C). This risk scoring system had a high sensitivity and specificity (93% and 100%, respectively) in predicting OS. The high-risk scoring group represented a promising finding in the use of this approach to predict OS in patients with thyroid cancer.

Table 1. Clinicopathological characteristics of 507 patients with thyroid cancer.

Characteristics	Subtype	Patinets n (%)	
Age	>47	238	(44.9)
	≤47	269	(55.1)
Gender	Male	136	(26.8%)
	Female	371	(73.2%)
Histological type	Thyroid papillary carcinoma classical/usual	359	(70.8%)
	Thyroid papillary carcinoma follicular	102	(20.1%)
	Thyroid papillary carcinoma – tall cell	37	(7.3%)
	Other, specify	9	(1.8%)
Focus type	Multifocal	228	(45.0%)
	Unifocal	269	(53.1%)
	Unknown	10	(1.9%)
Tumor stage	T1	44	(8.7%)
	T1a	20	(4.0%)
	T1b	80	(15.8%)
	T2	167	(32.9%)
	T3	171	(33.7%)
	T4	9	(1.8%)
	T4a	14	(2.8%)
	Unknown	2	(0.4%)
Lymph node	N0	231	(45.6%)
	N1	58	(11.4%)
	N1a	93	(18.3%)
	N1b	75	(14.8%)
	Unknown	50	(9.9%)
Metastasis	M0	283	(55.9%)
	M1	9	(1.8%)
	Unknown	215	(42.4%)
Survival status	Alive	491	(96.8%)
	Dead	16	(3.2%)

N=507.

The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis

A total of nine KEGG pathways were found to be enriched based on the 101 survival-related mRNAs (Figure 3A). As shown in Figure 3B, the neuroactive ligand-receptor interaction represents the highest cancer-related pathway, which contains 39 mRNAs, followed by transcriptional misregulation in cancer mRNAs. In the network diagram, all the KEGG related mRNAs and expression levels are visualized, the blue genes represent upregulation while the green genes represent down-regulation and the pathways are represented in red (Figure 3C). The pathway identities are described in Table 5.

Construction of the competing endogenous RNA (ceRNA) network of mRNAs in thyroid cancer

To further understand the function of the differentially expressed mRNAs in thyroid cancer, a ceRNA network composed of mRNA, lncRNA, and miRNA was created that predicted the lncRNA and miRNA interactions. As a result, 492 lncRNAs (397 upregulated and 95 down-regulated) and 72 miRNAs (67 up-regulated and five down-regulated) were identified as differentially expressed in thyroid cancer. The lncRNA and miRNA expression levels are shown in the volcano plot and heat map (Supplementary Figure 1B, 1C). A total of 31 lncRNAs were targeted by 12 key miRNAs described in the ceRNAs
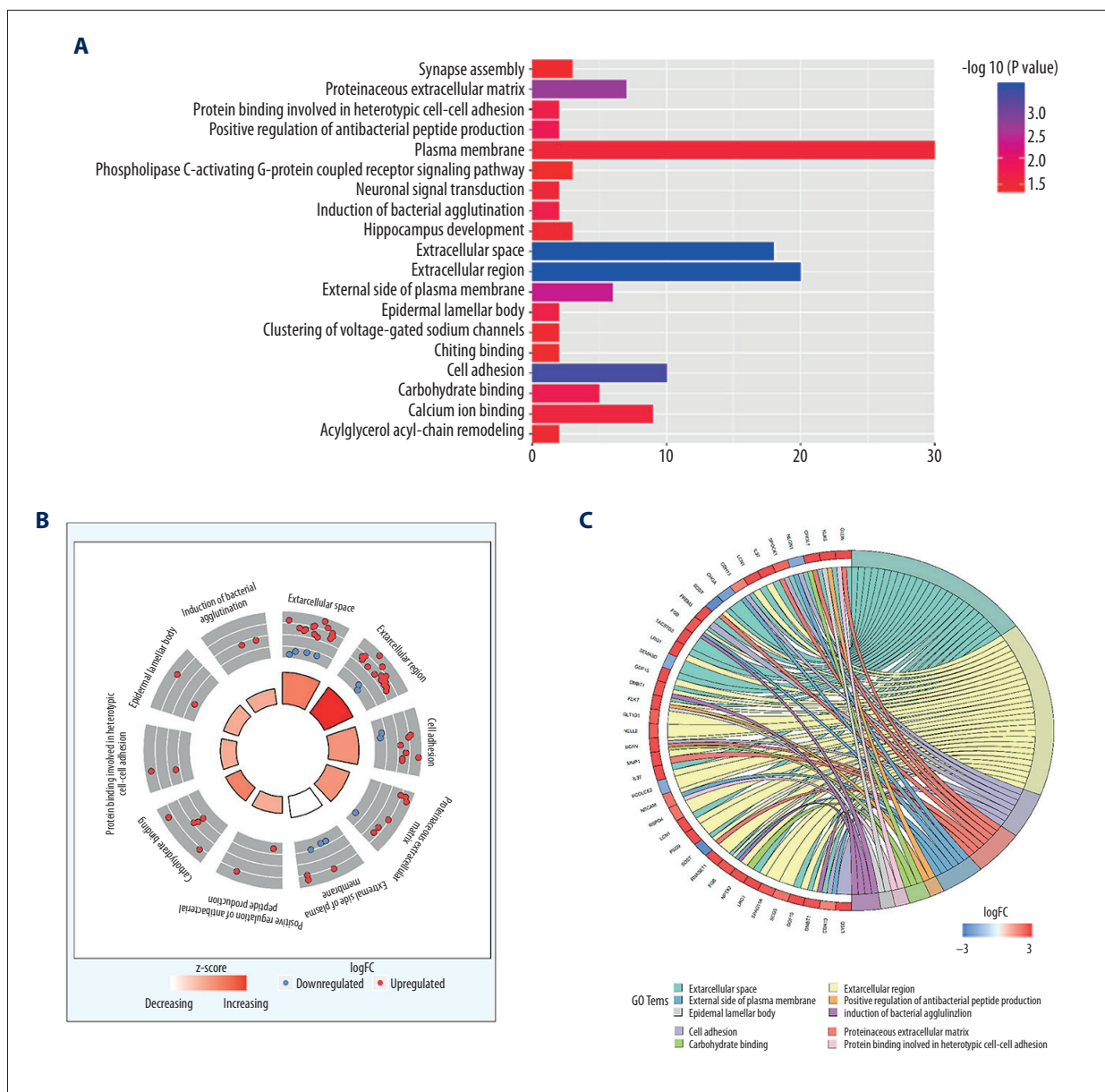
Figure 1. Differentially expressed mRNAs and Gene Ontology (GO) analysis. (**A**) Nineteen significant biological processes in Gene Ontology (GO) analysis of differentially expressed mRNAs. (**B**) The circle plot shows the expression level and z-score of differentially expressed mRNAs. (**C**) All differentially expressed mRNAs related to Gene Ontology (GO) analysis are shown diagrammatically.

network (Table 6). Also, 13 mRNAs were predicted to be targeted by seven key miRNAs (Table 7). In this ceRNA network, four lncRNAs and all miRNAs were down-regulated, while 27 lncRNAs and mRNAs were upregulated (Table 8, Figure 4).

Then, the dysregulated ceRNA network was constructed with differentially expressed RNAs, including 31 lncRNAs, seven miRNAs, and 13 mRNAs. The results indicated that differentially expressed lncRNAs could indirectly interact with mRNAs through miRNAs in thyroid cancer. The ceRNA network is

illustrated in Figure 5, where the red genes represent upregulation and the green genes represent down-regulation. To further identify the differentially expressed RNAs with prognostic significance, Kaplan-Meier survival analysis was used. As a result, two of 31 differentially expressed lncRNAs (MIR181A2HG and OPCML-IT1) and one of seven differentially expressed miRNAs (miR-184) were significantly associated with OS (log-rank test, P <0.05) (Figure 6).

Table 2. Overall survival (OS) in 101 mRNAs identified with Kaplan-Meier survival analysis.

mRNA	P-value
DMBT1, FGB, GPRIN1, LRG1, CLT1D1, MMP1, FNDC4, EYA1, ZSCAN4, DTX4, NELL2, ENTPD2, EGR2, E2F1, NLGN1, ADRA1B, TUSC3, NPC2, QRFPR, RAB27B, LMOD1	P<0.01
VAX2, GMNC, DIRAS3, FREM3, B3GN7, CDH3, CHI3L1,	0.01≤P<0.03
MS4A15, CLDN10, SPOCK2, GALNT7, XKRX, DGAT2L6, MYF5, NOD1, TACSTD2, PSG9, NLRP4, LY6D, PRR15, KCNA1, SPOCK1, BCAN, FIBCD1, PAPSS2, LRP1B, LCN1, RERGL, PCDHGC5, LHFPL3, TM4SF4, RNASE11, VRTN	
FAM83A, IL37, SCG5, KIAA1549L, HLA-G, STAC2, CLDN16, GPX2, LGALS13, DUSP5, HTR1E, FOXJ1, SEMA3D, GLDN, ENTPD1, KRT6C, GGCT, RXRG, FAM43A, NPTX2, TFCP2L1, SPAG11A, TCIM, ABCC11, GDF15, GALR2, PCOLCE2, PLCD3, KLK7, TEKT1, NAT8L, IRS4, CALHM4, OR4D6, SOST, SKOR2, EPHA5, CHGA, NRCAM, KLK5, STAB2, TMEM215, RSPO4	0.03≤P<0.05

Table 3. RNA expression was verified with univariate Cox regression analysis.

Gene	HR	z	P-value
NLGN1	1.829110700	3.774084277	0.000160596
PAPSS2	2.115458744	3.680176199	0.000233073
PCOLCE2	1.593571036	3.403752823	0.000664669
TEKT1	1.584227855	3.399998286	0.000673863
ZSCAN4	0.616746419	−3.32807659	0.000874478
ADRA1B	0.527596332	−3.29872768	0.000971241
KCNA1	1.364999932	3.198282015	0.00138249
NPC2	0.522251501	−3.16936368	0.001527731
EYA1	1.643738302	3.122550636	0.001792913
DTX4	0.716121906	−2.97322480	0.002946885
E2F1	0.488174682	−2.89423600	0.003800824
OR4D10	0.716299400	−2.89302393	0.003815522
GGCT	0.522082026	−2.70204197	0.006891506
FNDC4	0.698805379	−2.67016717	0.007581349
PRR15	0.810414585	−2.64626221	0.008138670
RXRG	0.827274307	−2.58362407	0.009776839

Table 4. Prediction of overall survival (OS) by mRNA expression profiles using multivariate Cox regression analysis.

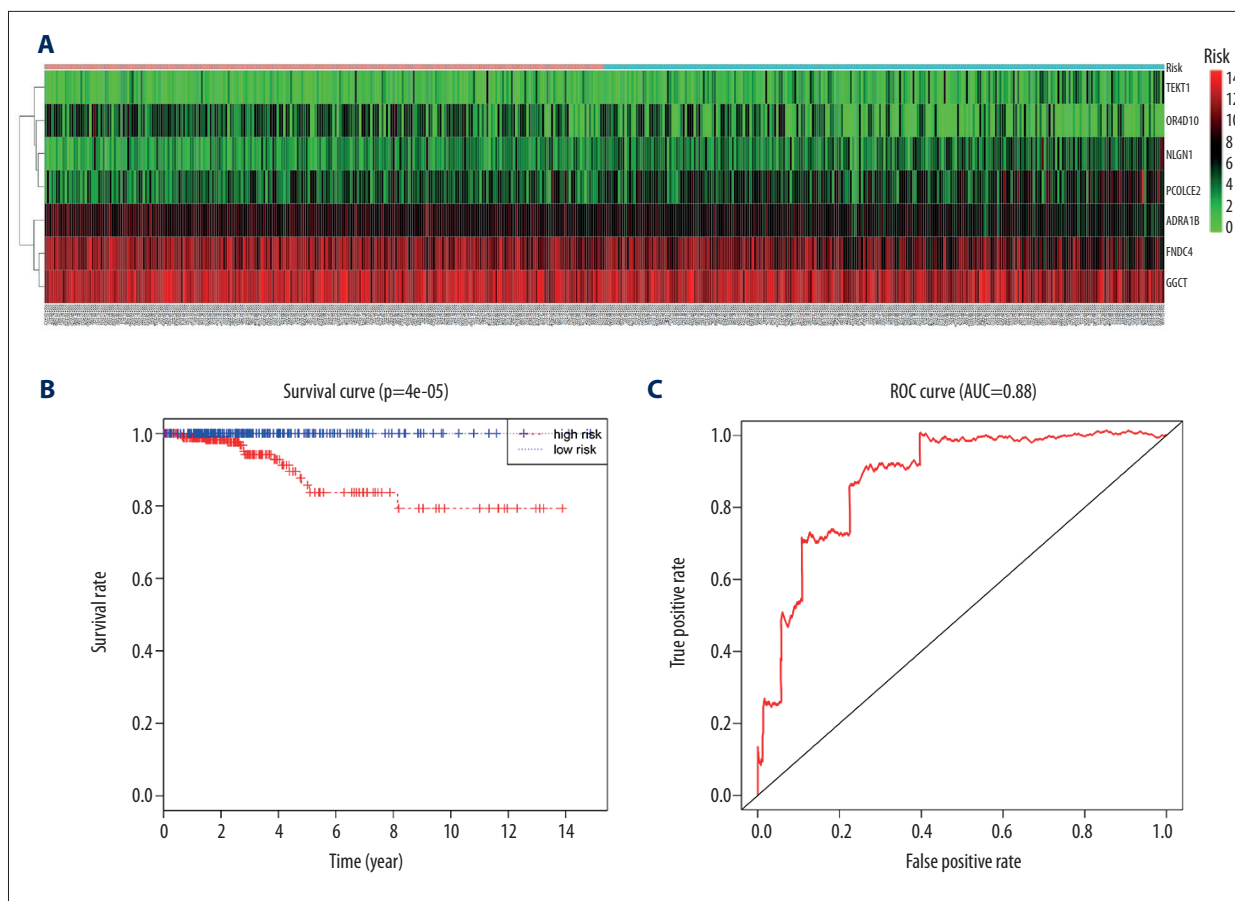| Gene	Coef	Exp(coef)	Se(coef)	z	P
ADRA1B	−0.516	0.597	0.253	−2.04	0.0417
FNDC4	−0.381	0.683	0.203	−1.87	0.0610
GGCT	0.567	1.764	0.362	1.57	0.1174
NLGN1	0.319	1.376	0.196	1.63	0.1029
OR4D10	−0.236	0.790	0.163	−1.44	0.1491
PCOLCE2	0.414	1.512	0.130	3.18	0.0015
TEKT1	0.236	1.267	0.160	1.48	0.1399

Figure 2. Survival analysis and Cox regression analysis based on risk scoring. (**A**) A heat map of seven signature mRNA expression profiles predictive of overall survival (OS) by multivariate Cox regression analysis. (**B**) Seven signature mRNAs were significantly associated with OS by risk model analysis (P<0.05). (**C**) Receiver operating characteristic (ROC) curve analysis of the risk scoring system. Area under the curve (AUC)=0.88.

Discussion

Thyroid cancer is a multifactorial disease, and although the molecular changes associated with this cancer have been extensively studied in the past, clinical challenges still exist. In this study, RNA expression profile bioinformatics data from cases of thyroid cancer from the Cancer Genome Atlas (TCGA), the Kyoto Encyclopedia of Genes and Genomes (KEGG), and the Gene Ontology (GO) databases were used to construct a competing endogenous RNA (ceRNA) network of mRNAs, long non-coding RNAs (lncRNAs), and microRNAs (miRNAs). The study identified 1098 mRNAs, 492 lncRNAs, and 72 miRNAs associated with differentiated thyroid cancer compared with normal thyroid tissue, 101 of the mRNAs were found to be associated with overall survival (OS), and a risk scoring system of seven signature expression profiles was established. A ceRNA network was then constructed consisting of 13 mRNAs, 31 lncRNAs and seven miRNAs, where two lncRNAs and one miRNA were identified as associated with patient prognosis in thyroid cancer.

Of the 1,098 mRNAs that were significantly differentially expressed in thyroid cancer, GO functional enrichment analysis showed that these genes were mainly enriched in the plasma membrane. Interestingly, several previous studies have shown that a large number of genes located in the plasma membrane have roles in thyroid cancer [15,16]. In this analysis, 101 differentially expressed mRNAs were shown to be related with OS, and these results were consistent with several previous reports, confirming the associations of some of these genes with prognosis. Luo et al. demonstrated that overexpression of the CHI3L1 gene was significantly associated with tumor metastasis and reduced OS in patients with thyroid cancer [17]. Also, CLDN10, E2F1, and HLA-G gene expression have been previously reported to be associated with poor survival in thyroid cancer [18–20], which are findings that support those of the present study.

Based on the results acquired by multivariate Cox regression analysis, this study identified a seven-signature mRNA expression profile predictive of OS. The discriminative value of the
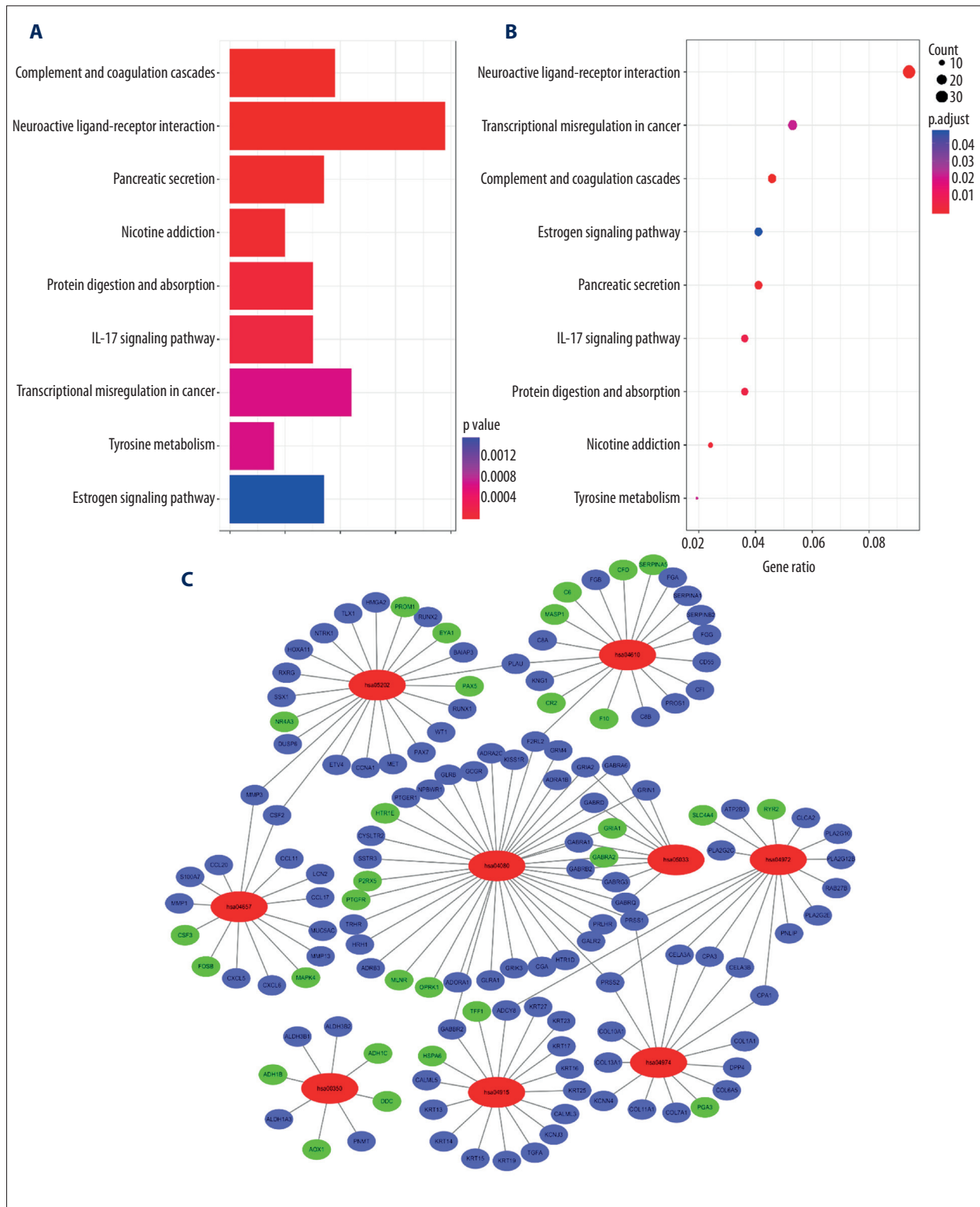
Figure 3. Prognosis-related mRNA and related the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. (**A**) Nine enrichment of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for differentially expressed mRNAs related to overall survival (OS) in thyroid carcinoma. (**B**) The scattergram shows the enrichment of the KEGG pathways. (**C**) The network of differentially expressed mRNAs involved in the KEGG pathways in thyroid cancer. Upregulated genes are represented by a blue ellipse, while downregulated genes are represented by a green ellipse.

Table 5. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways identification and corresponding description.

Pathway ID	Description	Adj.P-value	Number of DERNAs
hsa04610	Complement and coagulation cascades	9.67E-06	14
hsa04080	Neuroactive ligand-receptor interaction	9.67E-06	39
hsa04972	Pancreatic secretion	0.001724653	17
hsa05033	Nicotine addiction	0.003338851	10
hsa04974	Protein digestion and absorption	0.00652692	15
hsa04657	IL-17 signaling pathway	0.007949887	15
hsa05202	Transcriptional misregulation in cancer	0.02251093	22
hsa00350	Tyrosine metabolism	0.02251093	8
hsa04915	Estrogen signaling pathway	0.046699816	17

Table 6. Differentially expressed microRNAs (miRNAs) targeting long noncoding RNAs (lncRNAs) in the competing endogenous RNA (ceRNA) network.

LncRNA	miRNA
IGF2-AS	miR-519d
LINC00302	miR-31
AC022148.1	miR-372; miR-373; miR-144; miR-519d; miR-205; miR-221; miR-222; miR-31
LINC00313	miR-372; miR-373; miR-187; miR205; miR-31; miR375
AC004832.1	miR-519d; miR-31
AC002511.1	miR-519d
AC006305.1	miR-519d; miR-221; miR-222; miR-375
AP000525.1	miR-31
UCA1	miR-184
AC010336.2	miR-372; miR-373; miR-144; miR-519d; miR-205; miR-31
CLDN10-AS1	miR-221; miR-222
MIR181A2HG	miR-205
LINC00365	miR-519d
LINC00457	miR-144
SFTA1P	miR-221; miR-222
LINC00475	miR-205
LINC00423	miR-31
HOTAIR	miR-519d; miR-221; miR-222; miR-375
HCG22	miR-31
MIR4500HG	miR-144; miR-31
MIR205HG	miR-205; miR-221; miR-222; miR-31
CYP1B1-AS1	miR-205
LINC00460	miR-221; miR-222; mir-506
LINC00284	miR-519d; miR-205
AL158206.1	miR-372; miR-373; miR-221; miR-222
AC068594.1	miR-221; miR-222
AC011383.1	miR-31
SYNPR-AS1	miR-375
GDNF-AS1	miR-187
OPCML-IT1	miR-372; miR-373; miR-519d; miR-184; miR-205; miR-375
AP001029.2	miR-31

Table 7. Differentially expressed microRNAs (miRNAs) targeted mRNAs in the competing endogenous RNA (ceRNA) network.

miRNA	mRNA
miR-221	PCDHA1; PCDHAC2; CYP1B1
miR-205	SHISA6; LRRK2;
miR-144	GRIK3
miR-519d	SALL3; FOXQ1; E2F1
miR-373	TMEM100; TBC1D2
miR-372	TMEM100
miR-31	HOXC13

Table 8. Differentially expressed RNAs involved in the competing endogenous RNA (ceRNA) network.

RNAs	Regulation	Fold change	P-value
AL158206.1(lncRNA)	Up-regulation	2.042202918	1.82E-64
MIR181A2HG(lncRNA)	Up-regulation	2.158854422	1.70E-42
AC006305.1(lncRNA)	Down-regulation	−2.584610457	2.31E-41
LINC00475(lncRNA)	Up-regulation	3.437126593	1.92E-38
AC068594.1(lncRNA)	Up-regulation	2.707393493	3.53E-38
AP001029.2(lncRNA)	Up-regulation	2.729408357	3.02E-37
CYP1B1-AS1(lncRNA)	Up-regulation	2.316826339	6.46E-36
LINC00284(lncRNA)	Up-regulation	3.959715814	2.63E-30
AC010336.2(lncRNA)	Up-regulation	2.383032924	5.07E-29
AC022148.1(lncRNA)	Up-regulation	3.182439942	9.27E-29
AC002511.1(lncRNA)	Down-regulation	−2.243661781	1.21E-28
AC011383.1(lncRNA)	Up-regulation	2.414400943	2.39E-27
LINC00423(lncRNA)	Up-regulation	3.213132636	3.71E-21
LINC00460(lncRNA)	Up-regulation	3.65846406	3.54E-20
MIR4500HG(lncRNA)	Down-regulation	−2.427833048	5.33E-20
AC004832.1(lncRNA)	Down-regulation	−2.149556707	1.09E-19
HCG22(lncRNA)	Up-regulation	2.810057237	1.14E-19
SFTA1P(lncRNA)	Up-regulation	2.23122951	1.75E-19
AP000525.1(lncRNA)	Up-regulation	2.569810266	2.64E-17
LINC00365(lncRNA)	Up-regulation	2.433475211	2.44E-14
OPCML-IT1(lncRNA)	Up-regulation	5.036087373	3.60E-14
CLDN10-AS1(lncRNA)	Up-regulation	4.240702414	1.28E-13
GDNF-AS1(lncRNA)	Up-regulation	3.624884288	4.24E-13
UCA1(lncRNA)	Up-regulation	3.546251043	3.43E-12
LINC00457(lncRNA)	Up-regulation	2.891942882	5.67E-12
LINC00313(lncRNA)	Up-regulation	2.211716173	1.36E-11

Table 8 continued. Differentially expressed RNAs involved in the competing endogenous RNA (ceRNA) network.

RNAs	Regulation	Fold change	P-value
MIR205HG(lncRNA)	Up-regulation	2.703446156	3.27E-10
IGF2-AS(lncRNA)	Up-regulation	3.984846876	4.91E-08
LINC00302(lncRNA)	Up-regulation	3.377197372	6.49E-08
SYNPR-AS1(lncRNA)	Up-regulation	2.698830196	3.62E-07
HOTAIR(lncRNA)	Up-regulation	2.456805552	2.12E-06
miR-221(miRNA)	Down-regulation	3.275967528	1.04E-58
miR-222(miRNA)	Down-regulation	2.934265554	2.01E-54
miR-144(miRNA)	Down-regulation	−2.214636301	5.73E-53
miR-187(miRNA)	Down-regulation	3.08531177	1.96E-20
miR-31(miRNA)	Down-regulation	2.382671166	6.19E-19
miR-184(miRNA)	Down-regulation	3.332322739	2.54E-11
miR-372(miRNA)	Down-regulation	3.340999366	1.05E-08
miR-205(miRNA)	Down-regulation	2.014735591	2.11E-07
miR-519d(miRNA)	Down-regulation	4.562776545	5.31E-06
miR-373(miRNA)	Down-regulation	2.460474501	7.99E-05
LRRK2(mRNA)	Up-regulation	4.030832163	7.66E-51
TBC1D2(mRNA)	Up-regulation	2.011561395	9.64E-42
E2F1(mRNA)	Up-regulation	2.000905201	2.33E-39
SHISA6(mRNA)	Up-regulation	3.922870545	5.16E-28
CYP1B1(mRNA)	Up-regulation	3.089583272	1.54E-24
TMEM100(mRNA)	Up-regulation	2.831889424	1.89E-20
PCDHAC2(mRNA)	Up-regulation	2.163599995	2.46E-20
FOXQ1(mRNA)	Up-regulation	2.112598705	1.62E-19
GRIK3(mRNA)	Up-regulation	2.834216694	1.76E-19
PCDHA1(mRNA)	Up-regulation	2.035813016	8.10E-14
HOXC13(mRNA)	Up-regulation	2.952974119	1.55E-07
SALL3(mRNA)	Up-regulation	2.533075495	5.71E-06

combined gene expressions was good with receiver operating characteristic (ROC) area under the curve (AUC) discriminative value of 0.88. However, none of these seven genes (ADRA1B, FNDC4, GGCT, NLGN1, OR4D10, PCOLCE2, TEKT1) have previously been reported to be associated with thyroid cancer. The prognostic value of this set of seven genes remains to be confirmed and validated in future studies.

The KEGG pathway analysis in this study showed that the mRNAs of 39 genes were enriched in the neuroactive ligand-receptor interaction, including GABRB2 and MLNR. Previous studies have reported that these two genes have a strong correlation with lymph node metastasis in papillary thyroid cancer [21,22]. In addition to studies on mRNAs, an increasing number of studies have shown that lncRNAs play a critical role in thyroid cancer. However, the relationship between differentially expressed lncRNAs, miRNAs, and mRNAs remains unclear. However, the findings of the present study showed that all the differentially expressed lncRNAs in the ceRNA network were upregulated, except for AC006305.1,
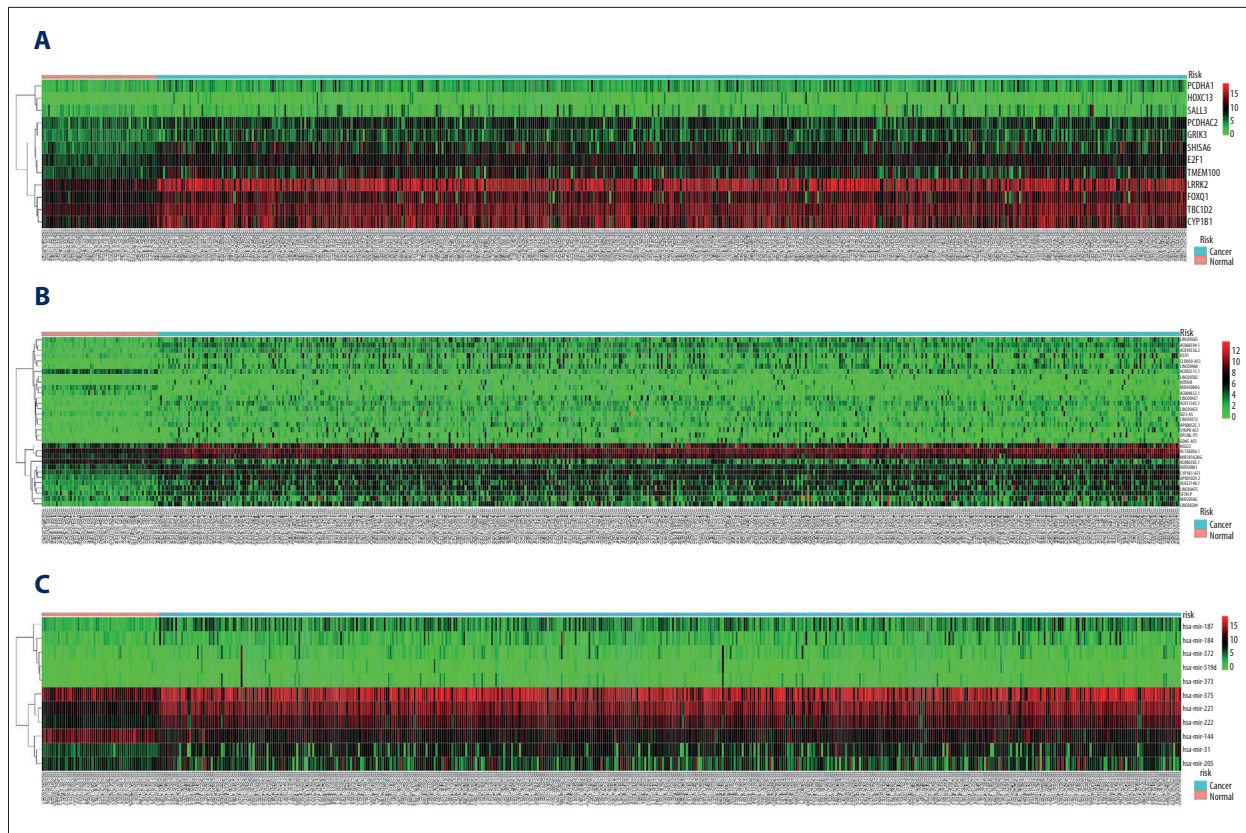
Figure 4. Differentially expressed genes involved in the competing endogenous RNA (ceRNA) network. (**A**) A heat map of differentially expressed mRNAs involved in the competing endogenous RNA (ceRNA) network. (**B**) A heat map of differentially expressed long noncoding RNAs (lncRNAs) involved in the ceRNA network. (**C**) A heat map of differentially expressed microRNAs (miRNAs) involved in the ceRNA network.

AC002511.1, MIR4500HG, and AC004832.1. None of these four lncRNAs have been previously reported in the literature in association with thyroid cancer.

However, in the present study, all differentially expressed miRNAs were down-regulated, and of these differentially expressed miRNAs, miR-144 and miR-31 have previously been reported to be associated with thyroid cancer. Guan et al. reported that down-regulation of miR-144 could promote thyroid cancer cell invasion by targeting the ZEB1 and ZEB2 genes, suggesting that miR-144 is acting as a tumor suppressor in thyroid cancer [23]. Wu et al. showed that miR-31 inhibited proliferation of thyroid cancer cells by targeting the RNA-binding protein, HuR, which modulates mRNAs [24]. Therefore, the results of the present study indicate that all the differentially expressed miRNAs involved in the constructed ceRNA network may function as tumor suppressor genes, although the specific functions need to be validated by further mechanistic studies.

In the present study, the ceRNA network that was developed for thyroid cancer included only two lncRNAs (MIR181A2HG and OPCML-IT1) and one miRNA (Mir-184) that could predict patient

survival. The MIR181A2HG gene, as a MIR181A2 host gene, has been previously reported to be overexpressed in the thyroid, but the function of MIR181A2HG remains unknown [25]. Also, the function of OPCML-IT1 also remains unclear. Mir-184, which was found to be down-regulated in the ceRNA network in this study and was associated with increased OS, has previously been reported to have a role in increasing cell proliferation oral squamous cell carcinoma and in accelerating the development of resistance to chemotherapy [26]. However, the function of miR-184 in thyroid cancer requires further study.

Conclusions

This study used data from patients with thyroid cancer from the Cancer Genome Atlas (TCGA), the Kyoto Encyclopedia of Genes and Genomes (KEGG), and the Gene Ontology (GO) databases to construct a competing endogenous RNA (ceRNA) network of mRNAs, long noncoding RNAs (lncRNAs), and microRNAs (miRNAs). The study identified differentially expressed mRNAs, lncRNAs and miRNAs related to thyroid cancer progression and prognosis. Some of these RNAs may represent
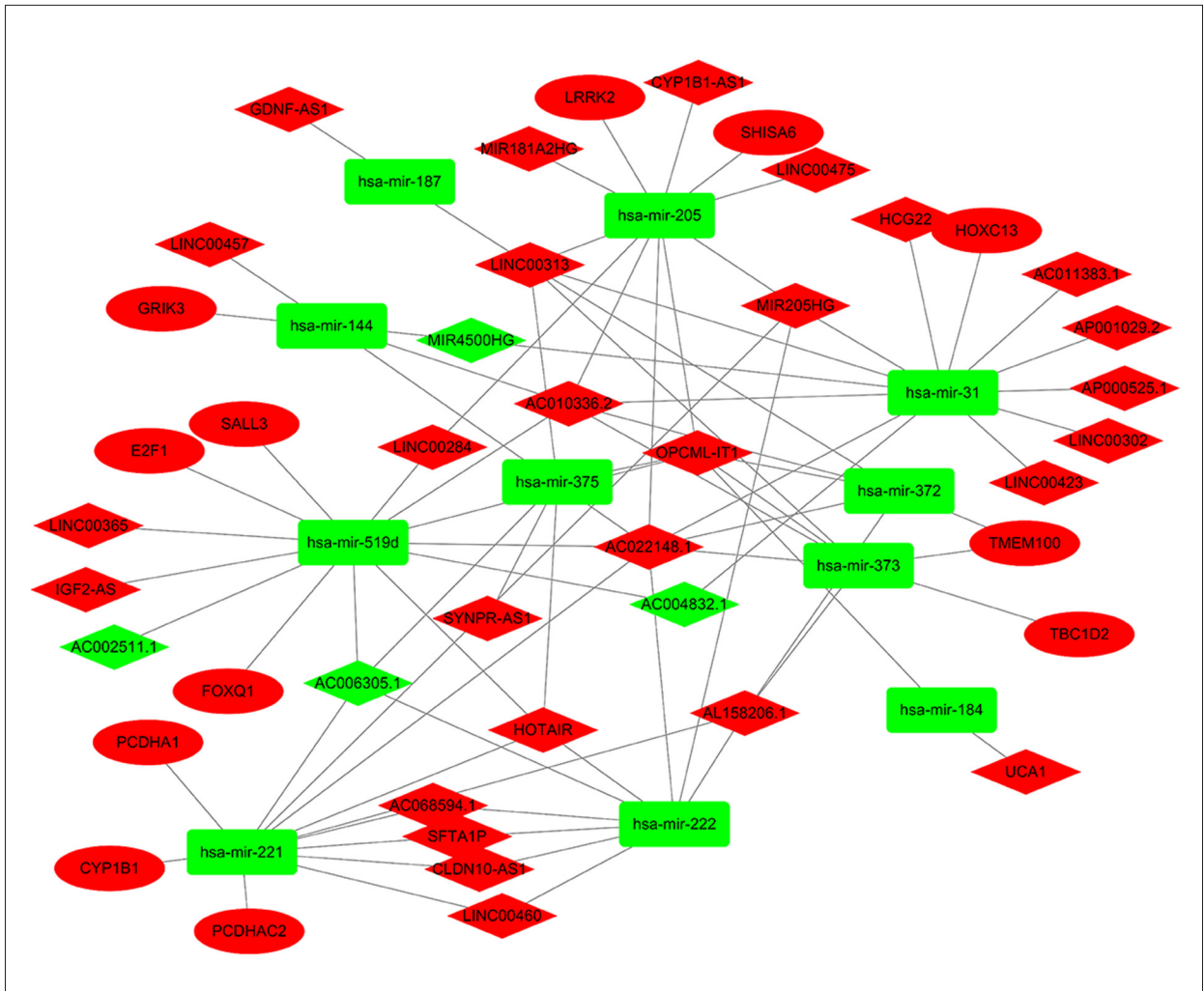
Figure 5. The competing endogenous RNA (ceRNA) network of thyroid cancer. Downregulated genes are shown as a green color, and upregulated genes are shown as a red color. The mRNAs are represented as an ellipse, the long noncoding RNAs (lncRNAs) are represented as a rhombus, and the microRNAs (miRNAs) are represented as a rectangle.
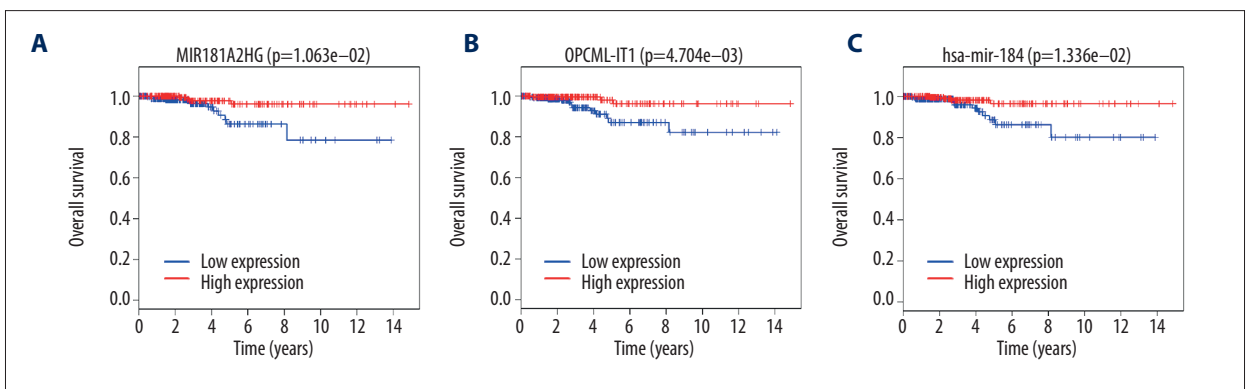

Figure 6. The overall survival (OS) associated with the two long noncoding RNAs (lncRNAs), MIR181A2HG and OPCML-IT1, and the microRNA, miR-184 in thyroid cancer.

diagnostic or prognostic biomarkers in thyroid cancer but require further validation and study. Also, future studies with functional experiments are needed to determine the specific functional roles of the components of the ceRNA network in thyroid cancer.

Supplementary Figure
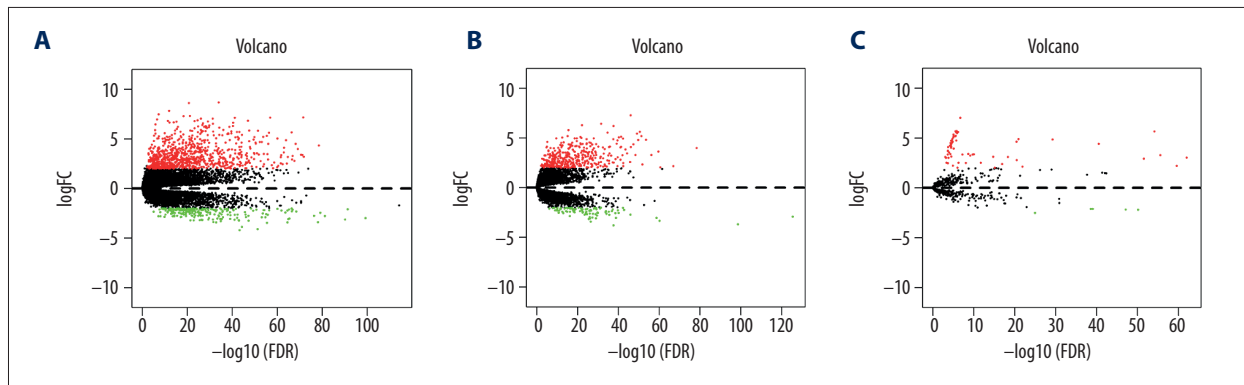
Supplementary Figure 1. Differentially expressed RNAs in thyroid cancer. Volcano plot to compare the RNAs between thyroid cancer tissues and normal thyroid gland tissues. The red dots represent upregulated genes and the green dots represent down-regulated genes. (**A**) Differentially expressed mRNAs in thyroid cancer. (**B**) Differentially expressed long noncoding RNAs (lncRNAs) in thyroid cancer. (**C**) Differentially expressed microRNAs (miRNAs) in thyroid cancer.

Conflict of interest

None.

References:

1. Yapa S, Mulla O, Green V et al: The role of chemokines in thyroid carcinoma. Thyroid, 2017; 27: 1347–59

2. Davies L, Welch HG: Current thyroid cancer trends in the United States. JAMA Otolaryngol Head Neck Surg, 2014; 140: 317–22

3. Siegel RL, Miller KD, Jemal A: Cancer Statistics, 2017. Cancer J Clin, 2017; 67: 7–30

4. Lloyd RV, Osamura RY, Klöppel G, Rosai J: WHO classification of tumours of endocrine organs, Fourth Edition, Volume 10. IARC Press, Lyon, 2017

5. Fahiminiya S, de Kock L, Foulkes WD: Biologic and clinical perspectives on thyroid cancer. N Engl J Med, 2016; 375: 2306–7

6. Fugazzola L, Puxeddu E, Avenia N et al: Correlation between B-RAFV600E mutation and clinicopathologic parameters in papillary thyroid carcinoma: data from a multicentric Italian study and review of the literature. Endocr Relat Cancer, 2006; 13: 455–64

7. Alberti L, Carniti C, Miranda C et al: RET and NTRK1 proto-oncogenes in human diseases. J Cell Physiol, 2003; 195: 168–86

8. Fagin JA, Mitsiades N: Molecular pathology of thyroid cancer: Diagnostic and clinical implications. Best Pract Res Clin Endocrinol Metab, 2008; 22: 955–69

9. Chan WL, Yuo CY, Yang WK et al: Transcribed pseudogene psiPPM1K generates endogenous siRNA to suppress oncogenic cell growth in hepatocellular carcinoma. Nucleic Acids Res, 2013; 41: 3734–47

10. Gou Q, Gao L, Nie X et al: Long noncoding RNA AB074169 inhibits cell proliferation via modulation of KHSRP-mediated p21 expression in papillary thyroid carcinoma. Cancer Res, 2018; 78(15): 4163–74

11. Zhang XF, Ye Y, Zhao SJ: LncRNA Gas5 acts as a ceRNA to regulate PTEN expression by sponging miR-222-3p in papillary thyroid carcinoma. Oncotarget, 2018; 9: 3519–30

12. McCarter JP, Mitreva MD, Martin J et al: Analysis and functional classification of transcripts from the nematode *Meloidogyne incognita*. Genome Biol, 2003; 4: R26

13. 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

14. Hsu SD, Tseng YT, Shrestha S et al: miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. Nucleic Acids Res, 2014; 42: D78–85

15. Morani F, Phadngam S, Follo C et al: PTEN regulates plasma membrane expression of glucose transporter 1 and glucose uptake in thyroid cancer cells. J Mol Endocrinol, 2014; 53: 247–58

16. Badziong J, Ting S, Synoracki S et al: Differential regulation of monocarboxylate transporter 8 expression in thyroid cancer and hyperthyroidism. Eur J Endocrinol, 2017; 177: 243–50

17. Luo D, Chen H, Lu P et al: CHI3L1 overexpression is associated with metastasis and is an indicator of poor prognosis in papillary thyroid carcinoma. Cancer Biomark, 2017; 18: 273–84

18. Zhang Z, Li SY, Zhang LB: LncRNA RGMB-AS1 is activated by E2F1 and promotes cell proliferation and invasion in papillary thyroid carcinoma. Eur Rev Med Pharmacol Sci, 2018; 22: 1979–86

19. Barros-Filho MC, Marchi FA, Pinto CA et al: High diagnostic accuracy based on CLDN10, HMGA2, and LAMB3 transcripts in papillary thyroid carcinoma. J Clin Endocrinol Metab, 2015; 100: E890–99

20. de Figueiredo-Feitosa NL, Martelli PG, Ciliao AD et al: HLA-G 3' untranslated region polymorphic sites associated with increased HLA-G production are more frequent in patients exhibiting differentiated thyroid tumours. Clin Endocrinol (Oxf), 2017; 86: 597–605

21. Jin Y, Jin W, Zheng Z et al: GABRB2 plays an important role in the lymph node metastasis of papillary thyroid cancer. Biochem Biophys Res Commun, 2017; 492: 323–30

22. Yip J, Orlov S, Orlov D et al: Predictive value of metastatic cervical lymph node ratio in papillary thyroid carcinoma recurrence. Head Neck, 2013; 35: 592–98

23. Guan H, Liang W, Xie Z et al: Down-regulation of miR-144 promotes thyroid cancer cell invasion by targeting ZEB1 and ZEB2. Endocrine, 2015; 48: 566–74

24. Wu D, Wang B, Shang J et al: miR-31 reduces cell growth of papillary thyroid carcinoma by RNA-binding protein HuR. Clin Lab, 2015; 61: 1625–34

25. Fagerberg L, Hallstrom BM, Oksvold P et al: Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics, 2014; 13: 397–406

26. Fang Z, Zhao J, Xie W et al: LncRNA UCA1 promotes proliferation and cisplatin resistance of oral squamous cell carcinoma by suppressing miR-184 expression. Cancer Med, 2017; 6: 2897–908