

Comprehensive analysis of a long noncoding RNA-associated competing endogenous RNA network in colorectal cancer

Qiaowei Fan^{1,2}
Bingrong Liu¹

¹Department of Gastroenterology and Hepatology, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, People's Republic of China; ²Translational Medicine Research and Cooperation Center of Northern China, Heilongjiang Academy of Medical Sciences, Harbin, Heilongjiang, People's Republic of China

Purpose: This study was aimed to develop a lncRNA-associated competing endogenous RNA (ceRNA) network to provide further understanding of the ceRNA regulatory mechanism and pathogenesis in colorectal cancer (CRC).

Patients and methods: Expression profiles of mRNAs, lncRNAs, and miRNAs, and clinical information for CRC patients were obtained from The Cancer Genome Atlas. The differentially expressed mRNAs, lncRNAs, and miRNAs (referred to as “DEmRNAs”, “DElncRNAs”, and “DEmiRNAs”, respectively) were screened out between 539 CRC samples and 11 normal samples. The interactions between DElncRNAs and DEmiRNAs were predicted by miRcode. The DEmRNAs targeted by the DEmiRNAs were retrieved according to TargetScan, miRTarBase, and miRDB. The lncRNA–miRNA–mRNA ceRNA network was constructed based on the DEmiRNA–DElncRNA and DEmiRNA–DEmRNA interactions. Functional enrichment analysis revealed the biological processes and pathways of DEmRNAs involved in the development of CRC. Key lncRNAs were further analyzed for their associations with overall survival and clinical features of CRC patients.

Results: A total of 1,767 DEmRNAs, 608 DElncRNAs, and 283 DEmiRNAs were identified as CRC-specific RNAs. Three hundred eighty-two DEmiRNA–DElncRNA interactions and 68 DEmiRNA–DEmRNA interactions were recognized according to the relevant databases. The lncRNA–miRNA–mRNA ceRNA network was constructed using 25 DEmiRNAs, 52 DEmRNAs, and 64 DElncRNAs. Two DElncRNAs, five DEmiRNAs, and six DEmRNAs were demonstrated to be related to the prognosis of CRC patients. Four DElncRNAs were found to be associated with clinical features. Twenty-eight Gene Ontology terms and 10 Kyoto Encyclopedia of Genes and Genomes pathways were found to be significantly enriched by the DEmRNAs in the ceRNA network.

Conclusion: Our results showed cancer-specific mRNA, lncRNA, and miRNA expression patterns and enabled us to construct an lncRNA-associated ceRNA network that provided new insights into the molecular mechanisms of CRC. Key RNA transcripts related to the overall survival and clinical features were also found with promising potential as biomarkers for diagnosis, survival prediction, and classification of CRC.

Keywords: colorectal cancer, competing endogenous RNA network, long noncoding RNA, survival analysis

Correspondence: Bingrong Liu
Department of Gastroenterology and Hepatology, The Second Affiliated Hospital of Harbin Medical University, 246 Xue-Fu Road, Nan-Gang District, Harbin 150086, People's Republic of China
Tel/fax +86 451 8660 5980
Email LiuBingrongLBR@126.com

Introduction

Noncoding RNAs (ncRNAs) are RNAs that lack protein-coding functions.¹ On the basis of their characteristics, ncRNAs can be divided into several classes, including lncRNAs, miRNAs, small interfering RNAs, small nuclear RNAs, and ribosomal RNAs.¹ miRNAs are a class of small ncRNAs with important regulatory roles, ~22 nt

in length.² They can posttranscriptionally regulate gene expression through binding to miRNA response elements (MREs) on their target transcripts.^{3,4} A single miRNA can regulate up to hundreds of target genes, and each gene may also be modulated by multiple miRNAs.^{5,6} Aberrantly expressed miRNAs are associated with many diseases including cancer.^{7–10}

In 2011, Salmena et al presented a competing endogenous RNA (ceRNA) hypothesis, which proposed a new “language” among different types of RNA transcripts.¹¹ The hypothesis postulates that RNA molecules can communicate with each other through shared MREs. MRNAs, lncRNAs, and other RNA transcripts could act as endogenous miRNA sponges to inhibit miRNA function, thereby impact the multiple targets of multiple miRNAs.¹¹ RNA transcripts sharing one or more MREs can actively regulate their respective expression levels by competing for a limited pool of miRNAs. This ceRNA crosstalk forms a large-scale regulatory network across the transcriptome, including protein-coding and noncoding RNAs.¹¹ lncRNAs, a class of ncRNAs >200 nt in length, are suggested to participate in various biological processes (BPs) by regulating gene expression at transcriptional, post-transcriptional, and epigenetic levels.^{12–14} Recent studies have provided emerging support for the ceRNA activity of lncRNAs involved in the progression of various types of cancer.^{15–18} lncRNA-associated ceRNA networks have been constructed and analyzed in lung squamous cell carcinoma, papillary renal cell carcinoma, hepatocellular carcinoma, and gastric cancer.^{19–22} However, there have been few integrated analyses of lncRNA-associated ceRNA networks in colorectal cancer (CRC).

In the present study, we conducted a comprehensive analysis of the mRNA, lncRNA, and miRNA expression profiles in CRC and constructed a CRC-specific lncRNA-associated ceRNA network using a large cohort from The Cancer Genome Atlas (TCGA). To the best of our knowledge, our study is the first to establish an lncRNA-associated ceRNA network in CRC using a combined dataset of colon adenocarcinoma and rectum adenocarcinoma from TCGA. The results are expected to help to further elucidate the lncRNA–miRNA–mRNA crosstalk in CRC and generate insight into the molecular mechanisms involved in the tumorigenesis and progression of CRC.

Patients and methods

Selection of CRC patients

The mRNA, lncRNA, and miRNA expression quantification files and the corresponding clinical information for CRC

patients were obtained from TCGA (<https://cancergenome.nih.gov/>). The inclusion criteria were as follows: 1) patients with complete clinical information including age, gender, pathological stage, TNM stage, survival status, and survival time; 2) patients with complete RNA-Seq data of mRNAs and lncRNAs, and complete miRNA-Seq data; 3) patients with a follow-up survival time of $\leq 2,000$ days. After filtering with the inclusion criteria, a total of 539 CRC patients and 11 normal controls were finally enrolled in the study. The clinicopathological characteristics of the CRC patients are summarized in Table 1.

Expression profile analysis of mRNAs, lncRNAs, and miRNAs in CRC and normal tissues

The differentially expressed mRNAs, lncRNAs, and miRNAs (hereafter referred to as “DEmRNAs”, “DElncRNAs”, and “DEmiRNAs”, respectively) in CRC and normal tissues were screened out individually using the edgeR package in the R language with the threshold of

Table 1 Clinicopathological characteristics of 539 CRC patients

Clinicopathological characteristics	Patients (N=539)	
	n	%
Age category (years)		
<65	220	40.82
≥ 65	319	59.18
Gender		
Female	252	46.75
Male	287	53.25
Pathological stage		
Stage I	99	18.37
Stage II	192	35.62
Stage III	163	30.24
Stage IV	85	15.77
Pathologic T		
Tis	1	0.18
T1	18	3.34
T2	95	17.63
T3	360	66.79
T4	65	12.06
Pathologic N		
N0	301	55.85
N1	138	25.60
N2	100	18.55
Pathologic M		
M0	404	74.95
M1	83	15.40
Mx	52	9.65
Survival status		
Alive	440	81.63
Dead	99	18.37

Abbreviation: CRC, colorectal cancer.

$|\log_2 \text{fold change} (\log_2 \text{FC})| > 2$ and false discovery rate (FDR) < 0.01 .²³ Hierarchical clustering analysis was then performed using the heatmap package in R (Version 1.0.8) to evaluate the specificity of the differentially expressed RNAs.²⁴

Construction of the ceRNA network

The interactions between DElncRNAs and DEmiRNAs were predicted by miRcode.²⁵ The mRNAs targeted by the DEmiRNAs were retrieved according to TargetScan, miRTarBase, and miRDB.^{26–28} Only mRNAs recognized by all three databases were considered as candidate mRNAs and intersected with the DEmRNAs to screen out the DEmRNAs targeted by the DEmiRNAs. Then, the lncRNA–miRNA–mRNA ceRNA network was constructed based on the DEmiRNA–DElncRNA and DEmiRNA–DEmRNA interactions and visualized using the Cytoscape software.²⁹ A flowchart of the lncRNA–miRNA–mRNA ceRNA network analysis is displayed in Figure 1. Linear regression analysis was performed

to evaluate the correlation between ceRNA expression levels.

Survival analysis

For each of the DEmRNAs, DElncRNAs, and DEmiRNAs in the ceRNA network, the CRC patients were classified into either a high expression group or a low expression group using the median expression value of the specific RNA as the cutoff value. The differences in patients' survival time between the high expression group and the low expression group were then evaluated by Kaplan–Meier survival curve and log-rank test analysis to find the mRNAs, lncRNAs, and miRNAs that were significantly associated with the survival of CRC patients. $P < 0.05$ was set as the threshold for statistical significance.

The associations between key lncRNAs and clinical features

Based on the bioinformatics analysis and the ceRNA network, key lncRNAs were screened out. The lncRNAs were

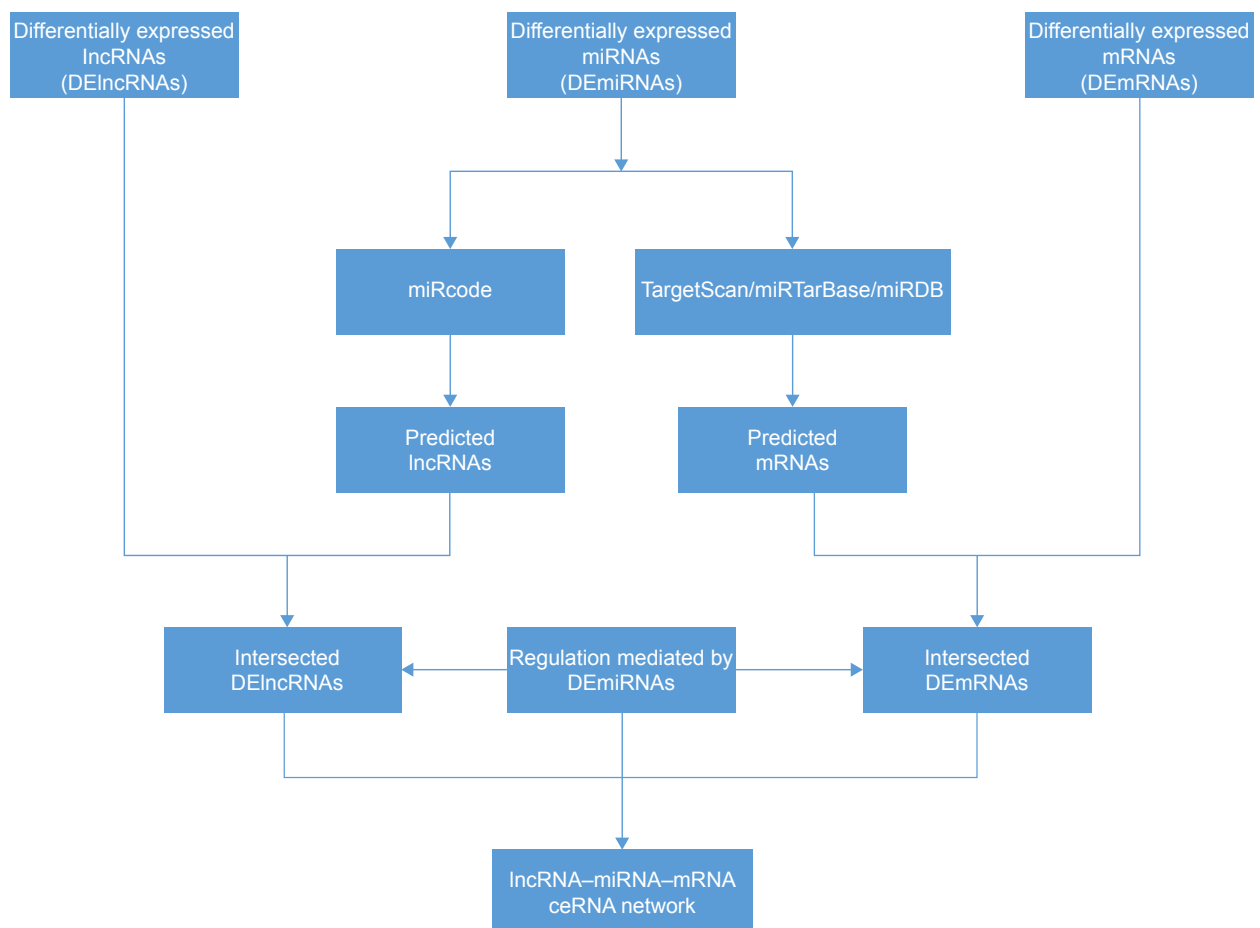


Figure 1 Flowchart of the lncRNA–miRNA–mRNA ceRNA network analysis.

then analyzed to determine whether they were differentially expressed between different clinical feature categories, including gender (male vs female), T stage (T3 + T4 vs T1 + T2), N stage (N1 + N2 vs N0), M stage (M1 vs M0), and pathological stage (stage III + IV vs stage I + II). Differentially expressed key lncRNAs with $|\log_2FC| > 1$ and $FDR < 0.05$ were considered significantly associated with the relevant clinical feature.

Functional enrichment analysis

Functional enrichment analysis was performed for the DEmRNAs in the ceRNA network to reveal the functional implications of these mRNAs in the tumorigenesis of CRC. Gene Ontology (GO) functional enrichment analysis was conducted using the Database for Annotation, Visualization and Integrated Discovery (DAVID, Version 6.8) online tool.³⁰ Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was carried out using the KEGG Orthology-Based Annotation System (KOBAS).³¹ $P < 0.05$ was set as the statistical significance threshold criterion for both GO and KEGG enrichment analysis.

Results

Cancer-specific mRNAs, lncRNAs, and miRNAs in CRC

According to the cutoff threshold, a total of 1,767 DEmRNAs (including 854 upregulated and 913 downregulated) were

screened out between CRC tissues and normal tissues. The expressions of 608 DElncRNAs (including 370 upregulated and 238 downregulated) and 283 DEmiRNAs (including 201 upregulated and 82 downregulated) were altered significantly in CRC tissues compared with normal tissues. Unsupervised hierarchical clustering analysis was performed for the top 100 upregulated and top 100 downregulated mRNAs and lncRNAs (owing to the large amount of DEmRNAs and DElncRNAs), respectively, and all the DEmiRNAs. The heatmaps showed that the CRC samples could be clearly distinguished from the normal samples with respect to the expression of these differentially expressed RNAs (Figure 2A–C).

DElncRNA–DEmiRNA interactions predicted by miRcode

The lncRNAs that are targeted by miRNAs were predicted by miRcode. Only interactions between DElncRNAs and DEmiRNAs were selected for ceRNA network construction. As a result, 479 interactions between 67 DElncRNAs (including KCNQ1OT1, ADAMTS9-AS2, and AGAP11) and 32 DEmiRNAs (including hsa-mir-106a, hsa-mir-141, and hsa-mir-143) were identified for further analysis.

Prediction of DEmRNAs targeted by DEmiRNAs

The 32 DEmiRNAs recognized in the above step were mapped into TargetScan, miRTarBase, and miRDB

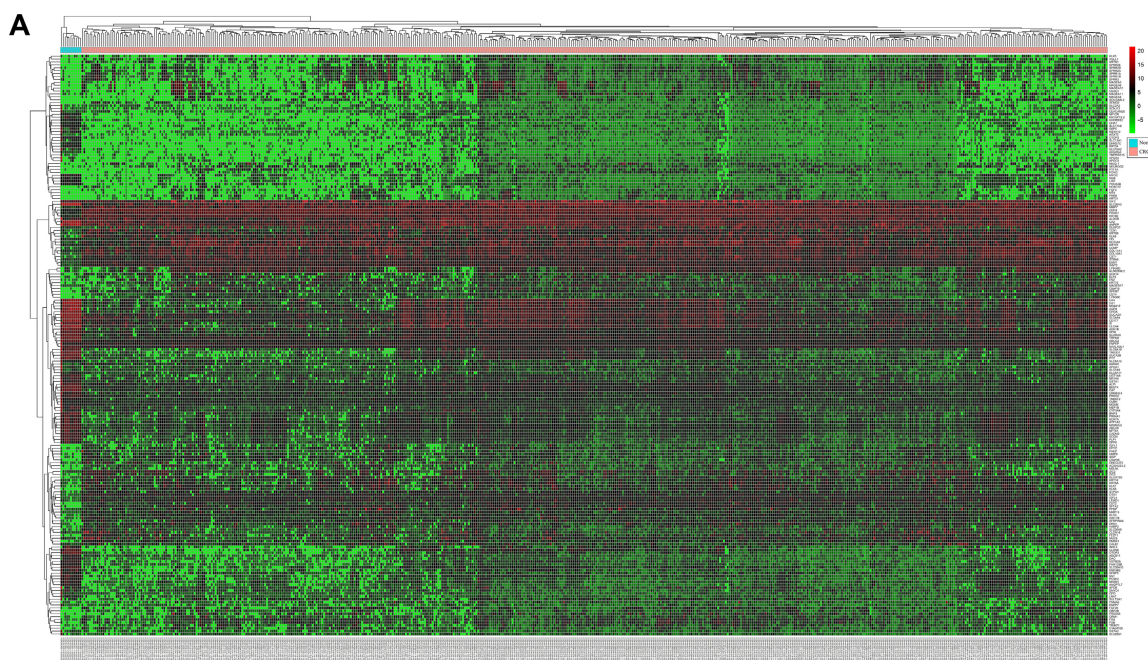


Figure 2 (Continued)

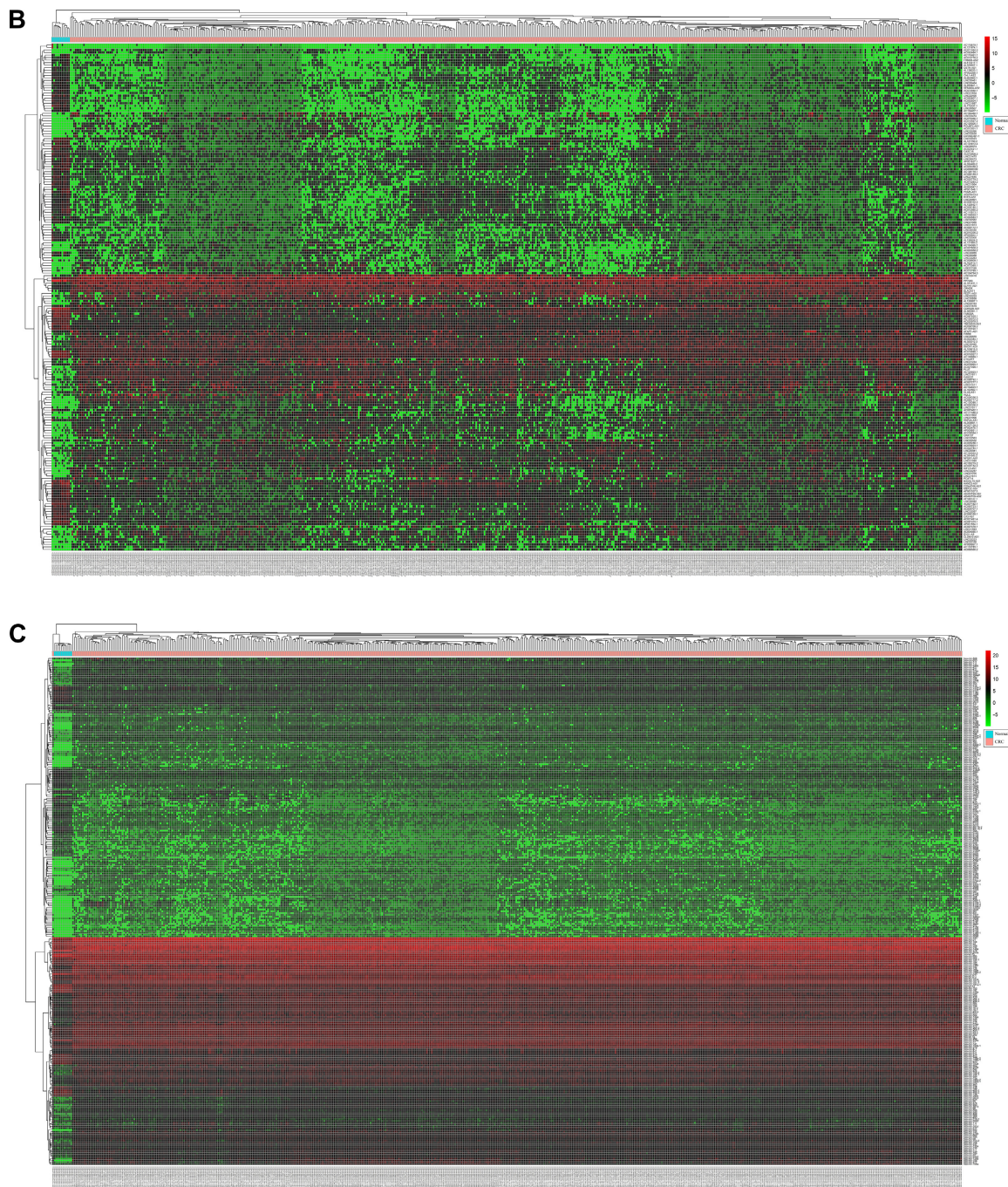


Figure 2 Heatmaps based on the differentially expressed RNAs in CRC and normal tissues.

Notes: (A) The heatmap with top 100 upregulated and top 100 downregulated mRNAs. (B) The heatmap with top 100 upregulated and top 100 downregulated lncRNAs. (C) The heatmap with all the DE miRNAs.

Abbreviation: CRC, colorectal cancer.

databases to search for their target mRNAs. A total of 1,416 mRNAs were found to be targeted by these DE miRNAs in all three databases. These candidate mRNAs were then intersected with the DE mRNAs, and 52 of the 1,416 mRNAs were identified as the DE mRNAs targeted by the DE miRNAs. Among the 32 candidate DE miRNAs, only 25 were found to interact with these 52 DE mRNAs, the

remaining seven DE miRNAs without specific DE mRNA targets were excluded from further analysis. To highlight the ceRNA features of the network, only DE miRNAs interacting with both DE mRNAs and DE lncRNAs were introduced into the ceRNA network. Finally, 25 DE miRNAs, 52 DE mRNAs, and 64 DE lncRNAs were selected for ceRNA network construction.

Table 2 The top 11 DElncRNAs putatively targeted by most DEmiRNAs in the ceRNA network

DElncRNAs	DEmiRNAs
KCNQ1OT1	hsa-mir-106a, hsa-mir-141, hsa-mir-143, hsa-mir-145, hsa-mir-150, hsa-mir-152, hsa-mir-17, hsa-mir-182, hsa-mir-192, hsa-mir-193b, hsa-mir-206, hsa-mir-217, hsa-mir-223, hsa-mir-301b, hsa-mir-32, hsa-mir-338, hsa-mir-372, hsa-mir-375, hsa-mir-424, hsa-mir-429, hsa-mir-454, hsa-mir-96, hsa-mir-98
ADAMTS9-AS2	hsa-mir-106a, hsa-mir-141, hsa-mir-143, hsa-mir-144, hsa-mir-145, hsa-mir-150, hsa-mir-152, hsa-mir-182, hsa-mir-193b, hsa-mir-223, hsa-mir-301b, hsa-mir-32, hsa-mir-338, hsa-mir-372, hsa-mir-375, hsa-mir-454, hsa-mir-96, hsa-mir-98
DLX6-AS1	hsa-mir-106a, hsa-mir-141, hsa-mir-144, hsa-mir-145, hsa-mir-150, hsa-mir-152, hsa-mir-17, hsa-mir-192, hsa-mir-193b, hsa-mir-206, hsa-mir-223, hsa-mir-338, hsa-mir-372, hsa-mir-424, hsa-mir-429
AGAP11	hsa-mir-106a, hsa-mir-141, hsa-mir-143, hsa-mir-145, hsa-mir-150, hsa-mir-152, hsa-mir-17, hsa-mir-182, hsa-mir-193b, hsa-mir-206, hsa-mir-21, hsa-mir-217, hsa-mir-372, hsa-mir-375, hsa-mir-424
AC010336.2	hsa-mir-106a, hsa-mir-141, hsa-mir-144, hsa-mir-145, hsa-mir-150, hsa-mir-17, hsa-mir-182, hsa-mir-192, hsa-mir-193b, hsa-mir-338, hsa-mir-372, hsa-mir-429
LINC00461	hsa-mir-106a, hsa-mir-141, hsa-mir-143, hsa-mir-144, hsa-mir-145, hsa-mir-150, hsa-mir-192, hsa-mir-32, hsa-mir-338, hsa-mir-372, hsa-mir-424, hsa-mir-96
LINC00484	hsa-mir-106a, hsa-mir-141, hsa-mir-143, hsa-mir-152, hsa-mir-206, hsa-mir-217, hsa-mir-223, hsa-mir-32, hsa-mir-338, hsa-mir-372, hsa-mir-424, hsa-mir-98
LIFR-AS1	hsa-mir-106a, hsa-mir-144, hsa-mir-150, hsa-mir-182, hsa-mir-192, hsa-mir-193b, hsa-mir-206, hsa-mir-32, hsa-mir-372, hsa-mir-375, hsa-mir-96
LINC00330	hsa-mir-106a, hsa-mir-145, hsa-mir-150, hsa-mir-17, hsa-mir-192, hsa-mir-206, hsa-mir-301b, hsa-mir-372, hsa-mir-424, hsa-mir-454, hsa-mir-98
FAM95B1	hsa-mir-141, hsa-mir-143, hsa-mir-150, hsa-mir-152, hsa-mir-182, hsa-mir-206, hsa-mir-338, hsa-mir-375, hsa-mir-96, hsa-mir-98
PVT1	hsa-mir-106a, hsa-mir-143, hsa-mir-145, hsa-mir-150, hsa-mir-152, hsa-mir-17, hsa-mir-21, hsa-mir-217, hsa-mir-372, hsa-mir-424

The ceRNA network in CRC

Based on the above data, an lncRNA–miRNA–mRNA ceRNA network consisting of 141 molecules and 450 interactions (382 DEmiRNA–DElncRNA interactions and 68 DEmiRNA–DEmRNA interactions) was constructed. The top 11 DElncRNAs targeted by most DEmiRNAs are shown in Table 2. All the 64 DElncRNAs in the ceRNA network with their name, Ensembl ID, \log_2FC , and FDR are listed in Table S1. All the 25 DEmiRNAs with their target DEmRNAs in the ceRNA network are shown in Table 3. The network was visualized with Cytoscape and is displayed in Figure 3. According to the ceRNA network, there is a possibility of indirect interactions between DElncRNAs and DEmRNAs. To confirm our finding, linear regression analysis was performed. The results suggested a good positive correlation between ceRNA expression levels (Figures 4A–I and 5A–H). For example, mediated by hsa-mir-106a, LINC00461 interacted with CADM2, CFL2, and FAM129A while mediated by hsa-mir-424, LINC00461 interacted with TMEM100 and TPM2 (Figure 4A, E, D, F, and H). Mediated by hsa-mir-152, LINC00484 interacted with BMP3 and KLF4 while mediated by hsa-mir-141, LINC00484 interacted with PHLPP2 (Figure 5 A, D, and C).

Table 3 The 25 DEmiRNAs with their target DEmRNAs in the ceRNA network

DEmiRNAs	DEmRNAs
hsa-mir-106a	CFL2, FOXQ1, FAM129A, CADM2, FJX1
hsa-mir-141	ELAVL4, PHLPP2, EPHA7, KIAA1549
hsa-mir-143	COL1A1, SERPINE1
hsa-mir-144	KCNQ5, GRIK3
hsa-mir-145	SERPINE1
hsa-mir-150	EREG, HILPDA, SLC7A11
hsa-mir-152	BMP3, NPTX1, KLF4
hsa-mir-17	FAM46C, FOXQ1, FJX1, CADM2, CYBRD1, FAM129A, CFL2
hsa-mir-182	CHL1, NPTX1, ULBP2
hsa-mir-192	GRHL1
hsa-mir-193b	PLAU
hsa-mir-206	STC2, SFRP1
hsa-mir-21	TGFBI
hsa-mir-217	DACH1
hsa-mir-223	EPB41L3
hsa-mir-301b	RBM20, CFL2
hsa-mir-32	PHLPP2, PBLD, PAX9, UGP2
hsa-mir-338	NOVA1
hsa-mir-372	CADM2, TMEM100, SLC7A11, MIXL1
hsa-mir-375	ELAVL4
hsa-mir-424	PSAT1, TMEM100, PHLPP2, CBX2, AXIN2, TPM2
hsa-mir-429	PMAIP1
hsa-mir-454	CFL2, RBM20
hsa-mir-96	ALK, TRIB3, SLC1A1
hsa-mir-98	HAND1, IGF2BP1, PRSS22, RGS16, TRIM71, IGF2BP3, CPA4

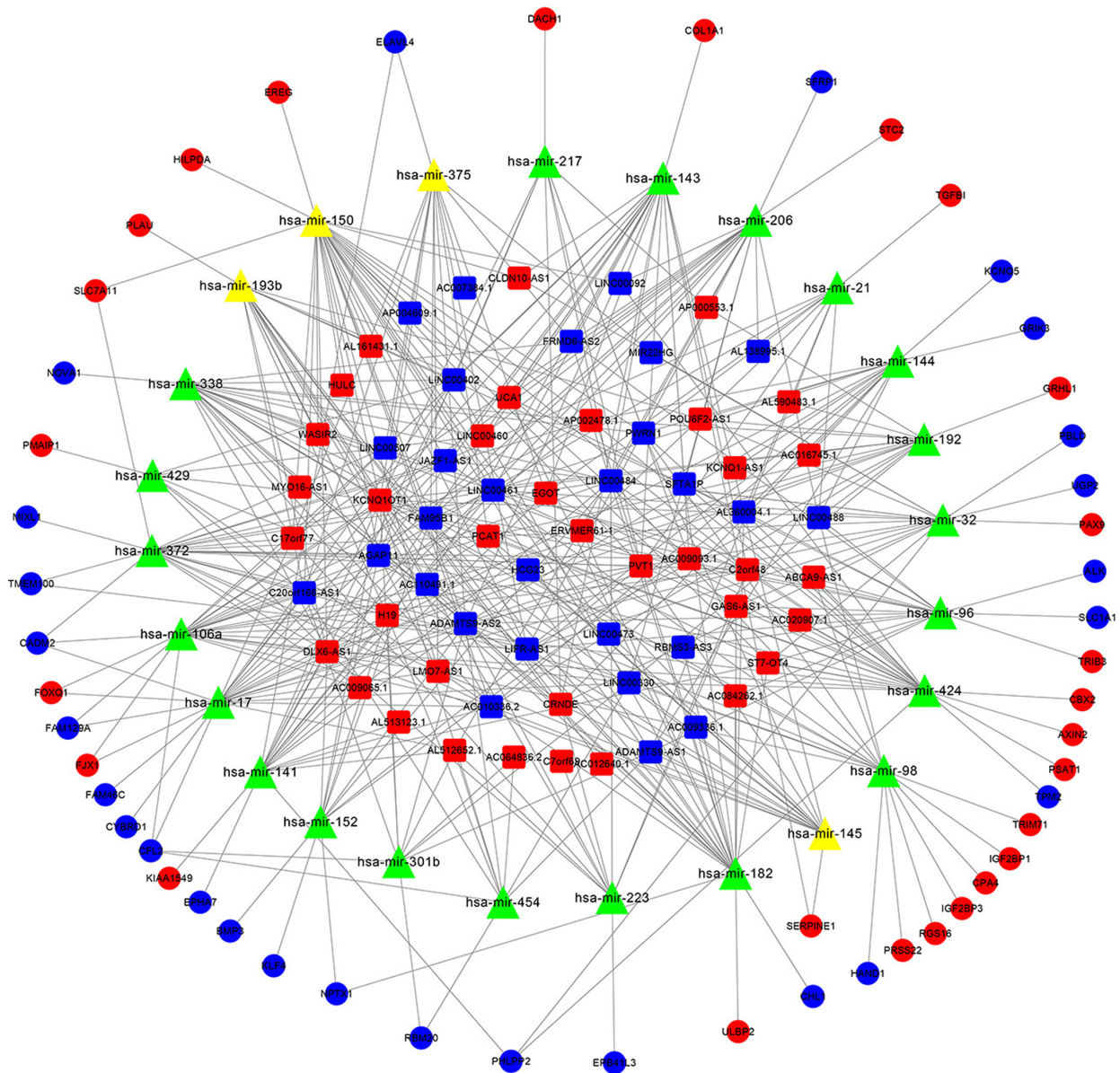


Figure 3 The lncRNA–miRNA–mRNA ceRNA network in CRC.

Notes: Red rounds represent upregulated DEmRNAs and blue rounds represent downregulated DEmRNAs. Red squares represent upregulated DElncRNAs and blue squares represent downregulated DElncRNAs. Green triangles represent upregulated DEmiRNAs and yellow triangles represent downregulated DEmiRNAs.

Abbreviation: CRC, colorectal cancer.

Important RNAs associated with the prognosis of CRC patients

Kaplan–Meier survival curve and log-rank test analysis were performed for each of the DElncRNAs, DEmiRNAs, and DEmRNAs in the ceRNA network to find the RNAs significantly associated with the survival of CRC patients. According to the results, two DElncRNAs (AP004609.1 and C2orf48), five DEmiRNAs (hsa-mir-96, hsa-mir-144, hsa-mir-145, hsa-mir-193b, and hsa-mir-375), and six DEmRNAs (FAM46C, FJX1, HAND1, SERPINE1, TPM2,

and ULBP2) were identified as prognosis-related RNAs ($p < 0.05$, Figures 6A and B, 7A–E, and 8A–F). Among these, hsa-mir-144, and hsa-mir-375 (DEmiRNA) and FAM46C (DEmRNA) appeared to be protective, as patients with higher expression levels of these RNAs tended to have longer overall survival, compared with those with lower expression levels of these RNAs (Figures 7B and E, and 8A). The remaining two DElncRNAs, three DEmiRNAs, and five DEmRNAs were considered risky, as their high expression was associated with shorter overall survival of CRC patients.

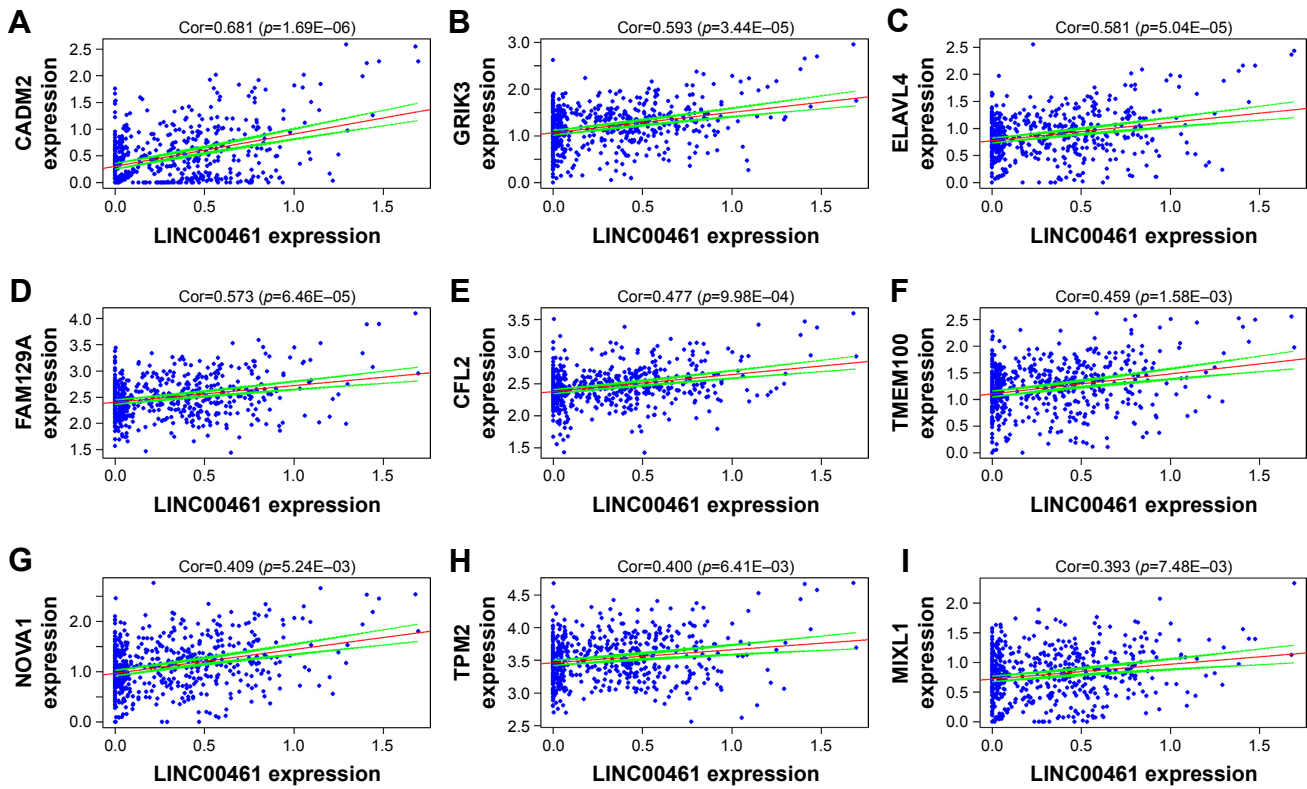


Figure 4 Linear regression of ceRNA expression levels for LINC00461.

Notes: LINC00461 vs (A) CADM2, (B) GRIK3, (C) ELAVL4, (D) FAM129A, (E) CFL2, (F) TMEM100, (G) NOVA1, (H) TPM2, and (I) MIXL1. The red line represents the linear model fitted by the dots in each figure, while green lines represent the 95% CI. “Cor” represents the correlation between the expressions of LINC00461 and the specific mRNA.

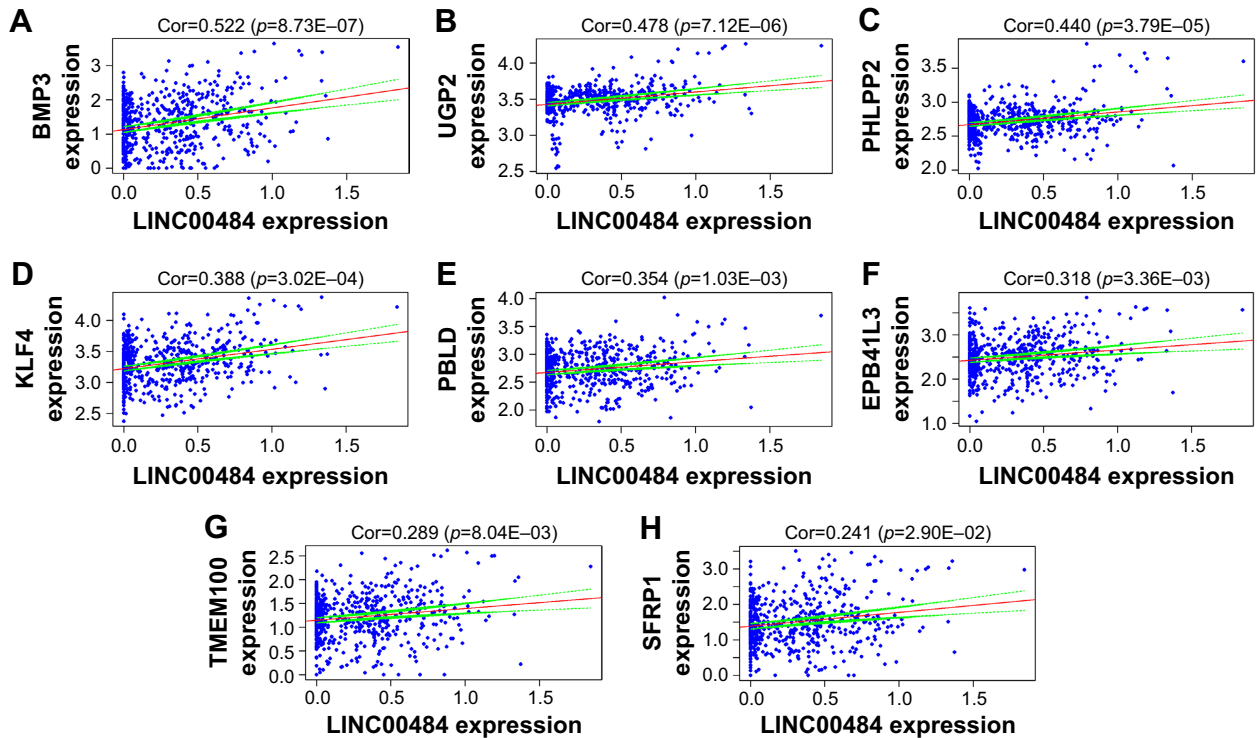


Figure 5 Linear regression of ceRNA expression levels for LINC00484.

Notes: LINC00484 vs (A) BMP3, (B) UGP2, (C) PHLPP2, (D) KLF4, (E) PBLD, (F) EPB41L3, (G) TMEM100, and (H) SFRP1. The red line represents the linear model fitted by the dots in each figure while green lines represent the 95% CI. “Cor” represents the correlation between the expressions of LINC00484 and the specific mRNA.

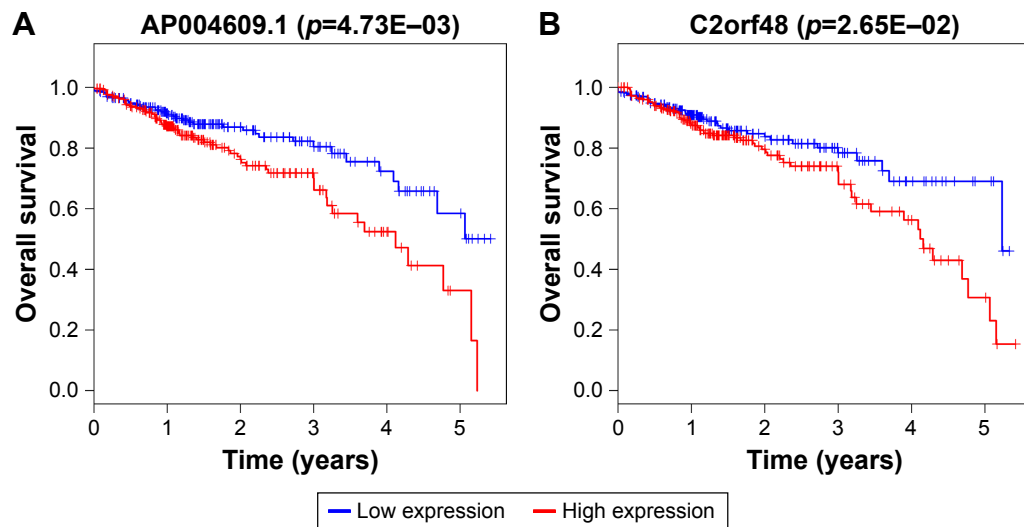


Figure 6 Kaplan–Meier survival curves for the two DElncRNAs significantly associated with overall survival of CRC patients.

Note: Kaplan–Meier survival curves for (A) AP004609.1 and (B) C2orf48.

Abbreviation: CRC, colorectal cancer.

Key lncRNAs relevant to clinical features

The 64 DElncRNAs from the ceRNA network were further analyzed to find their associations with clinical features. The lncRNA expression profiles of CRC patients were compared

between different genders, T, N, and M (tumor, node, and metastasis) stages, and pathological stages. Four key DElncRNAs were found to be significantly differentially expressed in clinical feature comparisons. AC110491.1 was significantly

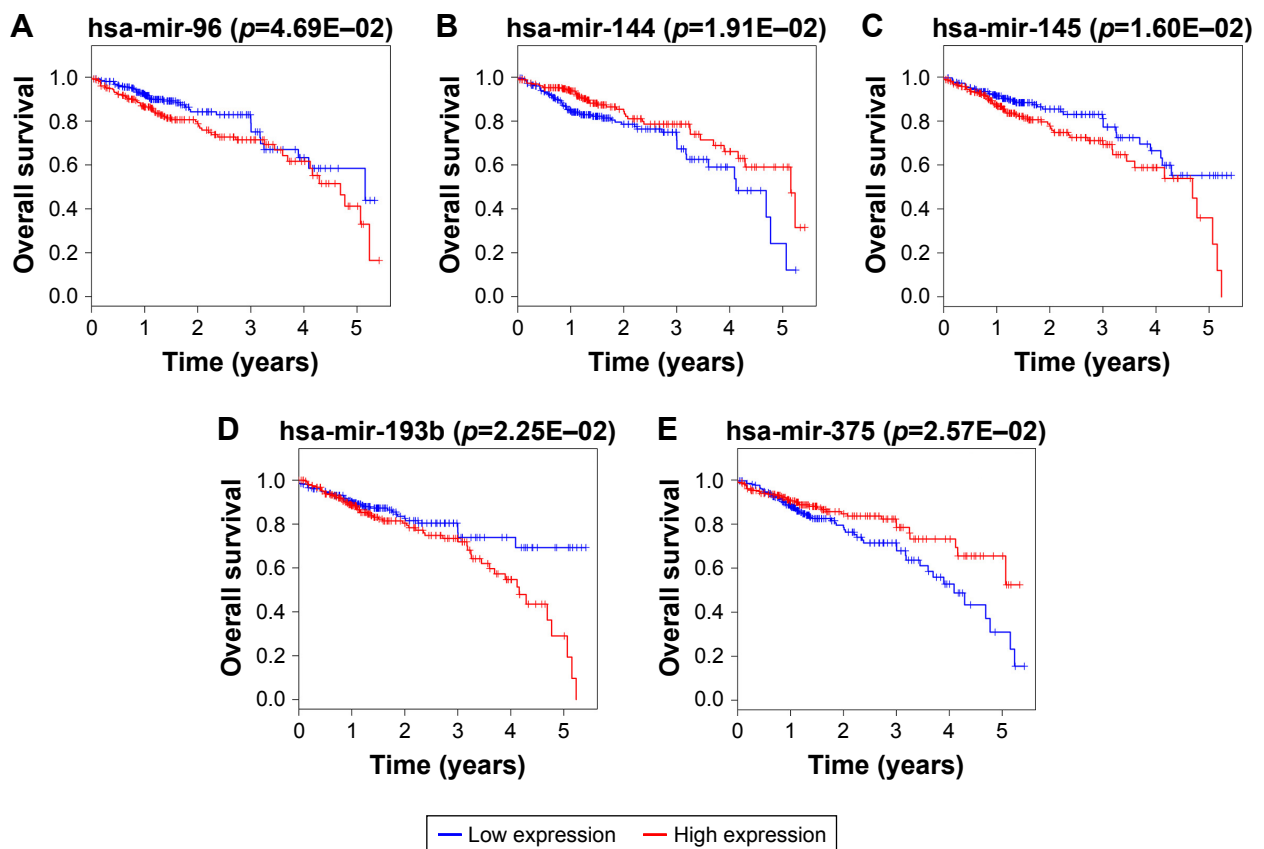


Figure 7 Kaplan–Meier survival curves for the five DEmiRNAs significantly associated with overall survival of CRC patients.

Note: Kaplan–Meier survival curves for (A) hsa-mir-96, (B) hsa-mir-144, (C) hsa-mir-145, (D) hsa-mir-193b, and (E) hsa-mir-375.

Abbreviation: CRC, colorectal cancer.

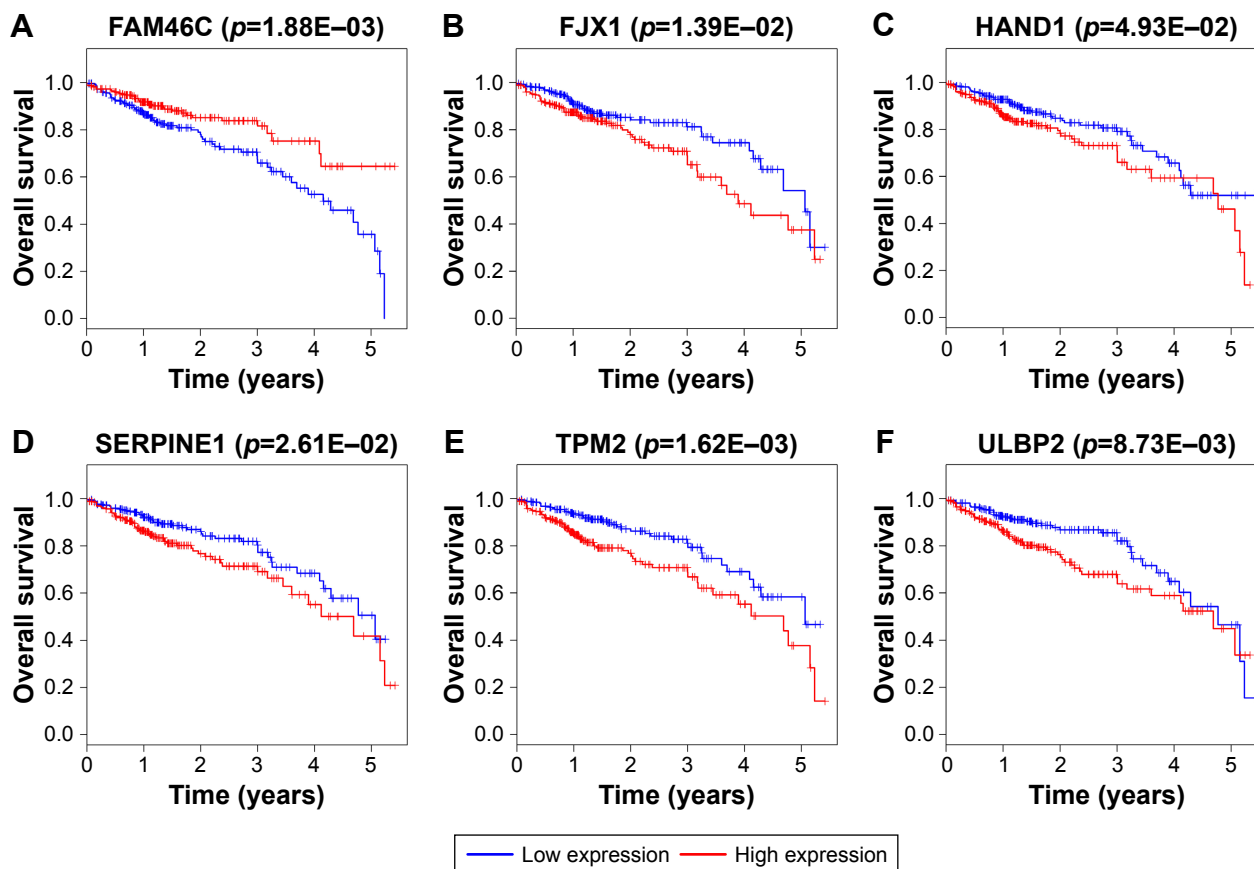


Figure 8 Kaplan–Meier survival curves for the six DE mRNAs significantly associated with overall survival of CRC patients. **Note:** Kaplan–Meier survival curves for (A) FAM46C, (B) FJX1, (C) HAND1, (D) SERPINE1, (E) TPM2, and (F) ULBP2. **Abbreviation:** CRC, colorectal cancer.

altered between genders (male vs female). H19 and HULC were not only differentially expressed in advanced-stage patients (stage III + IV) compared with earlier-stage patients (stage I + II), but also appeared to be associated with N stage (N1 + N2 vs N0) and M stage (M1 vs M0). As well as H19 and HULC, CLDN10-AS1 was also indicated to be an M stage-associated DE lncRNA. The key DE lncRNAs relevant to clinical features are summarized in Table 4.

Functional enrichment analysis

Functional enrichment analysis revealed that a total of 28 GO terms, including 20 BP terms, 4 cellular component

terms, and 4 molecular function terms were enriched by the DE mRNAs in the ceRNA network. Some of the GO terms were concerned with regulation of transcription, such as GO:0045892–negative regulation of transcription, DNA-templated and GO:0001190–transcriptional activator activity, RNA polymerase II transcription factor binding (Table 5).

The KEGG pathway enrichment analysis was also performed for the DE mRNAs. Ten KEGG pathways were found to be significantly enriched by the DE mRNAs in the ceRNA network, among which hsa04115: p53 signaling pathway, hsa05206: MicroRNAs in cancer, hsa04310: Wnt signaling pathway, and hsa04390: Hippo signaling pathway were closely associated with the pathogenesis of cancer (Table 6).

Table 4 The correlations between DE lncRNAs from the ceRNA network and clinical features of CRC patients

Comparisons	Upregulated	Downregulated
Gender (male vs female)		AC110491.1
N stage (N1 + N2 vs N0)	H19	HULC
M stage (M1 vs M0)	CLDN10-AS1, H19	HULC
Pathological stage (III + IV vs I + II)	H19	HULC

Abbreviation: CRC, colorectal cancer.

Discussion

CRC is one of the most common malignancies worldwide, ranking the third in men and the second in women.³² Despite numerous studies endeavoring to reveal the mechanisms of CRC, the exact pathogenesis has remained to be elucidated. In recent years, a novel hypothesis has been proposed that different RNA transcripts could interact with each

Table 5 The top 10 predominant BP terms and all the four CC terms and MF terms in GO functional enrichment analysis

Categories	Terms	Count	P-value
GO BP	GO:0045892~negative regulation of transcription, DNA-templated	7	3.16 E-03
	GO:0048147~negative regulation of fibroblast proliferation	3	3.68 E-03
	GO:0001525~angiogenesis	5	4.02 E-03
	GO:0043434~response to peptide hormone	3	7.30 E-03
	GO:0071363~cellular response to growth factor stimulus	3	7.62 E-03
	GO:0050680~negative regulation of epithelial cell proliferation	3	0.012
	GO:0017148~negative regulation of translation	3	0.012
	GO:0001957~intramembranous ossification	2	0.017
	GO:0042035~regulation of cytokine biosynthetic process	2	0.020
	GO:0060346~bone trabecula formation	2	0.023
GO CC	GO:0005615~extracellular space	13	1.79 E-04
	GO:0070062~extracellular exosome	14	0.031
	GO:0005886~plasma membrane	18	0.036
	GO:0031012~extracellular matrix	4	0.045
GO MF	GO:0003730~mRNA 3'-UTR binding	3	8.90 E-03
	GO:0045182~translation regulator activity	2	0.017
	GO:0048027~mRNA 5'-UTR binding	2	0.026
	GO:0001190~transcriptional activator activity, RNA polymerase II transcription factor binding	2	0.045

Abbreviations: BP, biological process; CC, cellular component; GO, Gene Ontology; MF, molecular function; UTR, untranslated region.

other through shared MREs.¹¹ This is called “the ceRNA hypothesis” and presents a new pattern of gene expression regulation that could be used to further understand the mechanisms of various diseases including cancer.

In the present study, we identified DElncRNAs, DEmiRNAs, and DEmRNAs between CRC tissues and normal tissues, using a large cohort from the TCGA database. Then, an lncRNA-miRNA-mRNA ceRNA network was constructed, which provided an integrated view of the ceRNA regulatory crosstalk among these CRC-specific RNA transcripts. The DElncRNAs in the ceRNA network were analyzed for their associations with the survival and clinical features of CRC patients. Functional enrichment analysis

Table 6 KEGG pathways enriched by the DEmRNAs in the ceRNA network

ID	Terms	P-value
hsa04550	Signaling pathways regulating pluripotency of stem cells	3.51 E-03
hsa04115	p53 signaling pathway	9.51 E-03
hsa04610	Complement and coagulation cascades	0.012
hsa00750	Vitamin B6 metabolism	0.014
hsa04974	Protein digestion and absorption	0.016
hsa05206	MicroRNAs in cancer	0.019
hsa04933	AGE-RAGE signaling pathway in diabetic complications	0.020
hsa04724	Glutamatergic synapse	0.024
hsa04310	Wnt signaling pathway	0.036
hsa04390	Hippo signaling pathway	0.041

Abbreviation: KEGG, Kyoto Encyclopedia of Genes and Genomes.

further revealed the BPs and pathways associated with the DEmRNAs involved in the development of CRC.

Dysregulated lncRNAs have been observed in various cancers, and some of them have been demonstrated to be associated with patient survival.³³⁻³⁶ For instance, decreased expression of the lncRNA MLLT4 antisense RNA 1 indicates poor prognosis in gastric cancer patients, whereas overexpression of the lncRNA linc-UBC1 is associated with poor prognosis in CRC.^{34,35} In the present study, two DElncRNAs from the ceRNA network, namely AP004609.1 and C2orf48, were found to be significantly associated with the survival of patients with CRC. Both of them could be considered life threatening, as patients with high expression levels of these two lncRNAs tended to have shorter survival times than patients with low expression. Although little is known about the functions of these two DElncRNAs, our results indicate the prognostic potential of the lncRNAs in survival prediction and provide new directions for further investigation of the two lncRNAs involved in CRC.

With respect to the associations between DElncRNAs and clinical features, we found that four DElncRNAs were related to clinical features, including gender, N stage, M stage, and tumor pathological stage. Among the four DElncRNAs, H19, a widely studied lncRNA, has been suggested as an indicator of tumor progression in multiple types of cancers, including CRC.³⁷⁻⁴⁰ H19 has been reported to indicate poor prognosis and promote tumor growth in CRC.³⁹ Overexpression of H19 is associated with decreased overall survival rates in CRC

patients, as well as increased viability and migration in colon cancer cells.⁴⁰ In our study, the expression of H19 was significantly upregulated in advanced tumor stages (stages N1/N2, M1, and stage III/IV) compared with earlier tumor stages (stages N0, M0, and stage I/II). The findings of our study were consistent with previous reports and further confirmed the important function of H19 in the progression of CRC.

MiRNAs, the hub elements of the ceRNA network, exhibited key roles in the crosstalk among different RNA transcripts. Mediated by hsa-mir-424, LINC00461 interacted with TMEM100 and TPM2 while LINC00484 interacted with TMEM100 and PHLPP2 in the present CRC-specific ceRNA network. Mediated by hsa-mir-141, LINC00461 interacted with ELAVL4 while LINC00484 interacted with PHLPP2. Both hsa-mir-424 and hsa-mir-141 were demonstrated to be involved in the pathogenesis of CRC.^{41,42} We also found five miRNAs to be significantly associated with the prognosis in CRC. High expression of hsa-mir-96, hsa-mir-145, and hsa-mir-193b indicated an unfavorable clinical outcome, while upregulation of hsa-mir-144 and hsa-mir-375 suggested longer survival times of patients with CRC. All the key DE miRNAs show promising potential as biomarkers for survival prediction in CRC.

mRNAs comprise another important part of the ceRNA network, through being directly targeted by miRNAs or having indirect interactions with lncRNAs mediated by miRNAs. Similar to lncRNAs and miRNAs, some mRNAs were also found to be associated with the survival of CRC patients, such as the protective mRNA FAM46C, and the risky mRNAs including FJX1, HAND1, SERPINE1, TPM2, and ULBP2. Furthermore, functional enrichment analysis revealed that DE mRNAs were enriched in some BPs and pathways related to transcription regulation and cancer, such as negative regulation of transcription, the p53 signaling pathway, and the Wnt signaling pathway.^{43,44} The close correlation between the enriched pathways and the ceRNA network demonstrates the credibility of our results.

Conclusion

The present study identified CRC-specific lncRNAs, miRNAs, and mRNAs using a large cohort from the TCGA database. We successfully constructed an lncRNA-associated ceRNA network, which provided insight into the newly proposed crosstalk among different types of RNA transcripts. Key lncRNAs were analyzed to determine significant associations with overall survival and clinical features of CRC patients. Hub miRNAs also showed promising potential as prognostic biomarkers in CRC. Our study provides better understanding of the ceRNA regulatory mechanism and helps

to elucidate the pathogenesis involved in the tumorigenesis and progression of CRC.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Song E, editor. *Advances in experimental medicine and biology. The Long and Short Non-Coding RNAs in Cancer Biology*. Singapore: Springer; 2016.
2. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004;116(2):281–297.
3. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009;136(2):215–233.
4. Cai Y, Yu X, Hu S, Yu J. A brief review on the mechanisms of miRNA regulation. *Genomics Proteomics Bioinformatics*. 2009;7(4):147–154.
5. Thomas M, Lieberman J, Lal A. Desperately seeking microRNA targets. *Nat Struct Mol Biol*. 2010;17(10):1169–1174.
6. Wu S, Huang S, Ding J, et al. Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3' untranslated region. *Oncogene*. 2010;29(15):2302–2308.
7. Singh A, Sen D. MicroRNAs in Parkinson's disease. *Exp Brain Res*. 2017;235(8):2359–2374.
8. Banerjee J, Nema V, Dhas Y, Mishra N. Role of MicroRNAs in type 2 diabetes and associated vascular complications. *Biochimie*. 2017;139:9–19.
9. Barwari T, Joshi A, Mayr M. MicroRNAs in cardiovascular disease. *J Am Coll Cardiol*. 2016;68(23):2577–2584.
10. Abba ML, Patil N, Leupold JH, et al. MicroRNAs as novel targets and tools in cancer therapy. *Cancer Lett*. 2017;387:84–94.
11. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 2011;146(3):353–358.
12. Kornienko AE, Guenzi PM, Barlow DP, Pauler FM. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol*. 2013;11:59.
13. Shi X, Sun M, Liu H, Yao Y, Song Y. Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett*. 2013;339(2):159–166.
14. Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol*. 2013;20(3):300–307.
15. Chen D, Sun Q, Zhang L, et al. The lncRNA HOXA11-AS functions as a competing endogenous RNA to regulate PADI2 expression by sponging miR-125a-5p in liver metastasis of colorectal cancer. *Oncotarget*. 2017;8(41):70642–70652.
16. Lv M, Zhong Z, Huang M, Tian Q, Jiang R, Chen J. lncRNA H19 regulates epithelial-mesenchymal transition and metastasis of bladder cancer by miR-29b-3p as competing endogenous RNA. *Biochim Biophys Acta*. 2017;1864(10):1887–1899.
17. Chen Z, Liu H, Yang H, Gao Y, Zhang G, Hu J. The long noncoding RNA, TINCR, functions as a competing endogenous RNA to regulate PDK1 expression by sponging miR-375 in gastric cancer. *Onco Targets Ther*. 2017;10:3353–3362.
18. Cui M, Wang J, Li Q, Zhang J, Jia J, Zhan X. Long non-coding RNA HOXA11-AS functions as a competing endogenous RNA to regulate ROCK1 expression by sponging miR-124-3p in osteosarcoma. *Biomed Pharmacother*. 2017;92:437–444.
19. Sui J, Xu SY, Han J, et al. Integrated analysis of competing endogenous RNA network revealing lncRNAs as potential prognostic biomarkers in human lung squamous cell carcinoma. *Oncotarget*. 2017;8(39):65997–66018.
20. Huang C, Yuan N, Wu L, et al. An integrated analysis for long noncoding RNAs and microRNAs with the mediated competing endogenous RNA network in papillary renal cell carcinoma. *Onco Targets Ther*. 2017;10:4037–4050.

21. Zhang J, Fan D, Jian Z, Chen GG, Lai PB. Cancer specific long non-coding RNAs show differential expression patterns and competing endogenous RNA potential in hepatocellular carcinoma. *PLoS One*. 2015;10(10):e0141042.
22. Xia T, Liao Q, Jiang X, et al. Long noncoding RNA associated-competing endogenous RNAs in gastric cancer. *Sci Rep*. 2014;4:6088.
23. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–140.
24. Kolde R [webpage on the Internet]. Pheatmap: Pretty heatmaps. R package version 1.0.8: 2015. Available from: <https://cran.r-project.org/web/packages/pheatmap/index.html>. Accessed May 14, 2016.
25. Jeggari A, Marks DS, Larsson E. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics*. 2012; 28(15):2062–2063.
26. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *eLife*. 2015;4:e05005.
27. Chou CH, Chang NW, Shrestha S, et al. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res*. 2016;44(D1):D239–D247.
28. Wong N, Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res*. 2015; 43(Database issue):D146–D152.
29. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504.
30. Dennis G Jr, Sherman BT, Hosack DA, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol*. 2003;4(5):3.
31. Wu J, Mao X, Cai T, Luo J, Wei L. KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Res*. 2006;34(Web Server issue):W720–W724.
32. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359–E386.
33. Sun Y, Zhou Y, Bai Y, et al. A long non-coding RNA HOTTIP expression is associated with disease progression and predicts outcome in small cell lung cancer patients. *Mol Cancer*. 2017;16(1):162.
34. Lai Y, Xu P, Liu J, et al. Decreased expression of the long non-coding RNA MLLT4 antisense RNA 1 is a potential biomarker and an indicator of a poor prognosis for gastric cancer. *Oncol Lett*. 2017;14(3):2629–2634.
35. Gao X, Wen J, Gao P, Zhang G, Zhang G. Overexpression of the long non-coding RNA, linc-UBC1, is associated with poor prognosis and facilitates cell proliferation, migration, and invasion in colorectal cancer. *Oncotargets Ther*. 2017;10:1017–1026.
36. Zhang JH, Li AY, Wei N. Downregulation of long non-coding RNA LINC01133 is predictive of poor prognosis in colorectal cancer patients. *Eur Rev Med Pharmacol Sci*. 2017;21(9):2103–2107.
37. Liu N, Zhou Q, Qi YH, Wang H, Yang L, Fan QY. Effects of long non-coding RNA H19 and microRNA let7a expression on thyroid cancer prognosis. *Exp Mol Pathol*. 2017;103(1):71–77.
38. Chen JS, Wang YF, Zhang XQ, et al. H19 serves as a diagnostic biomarker and up-regulation of H19 expression contributes to poor prognosis in patients with gastric cancer. *Neoplasma*. 2016;63(2): 223–230.
39. Han D, Gao X, Wang M, et al. Long noncoding RNA H19 indicates a poor prognosis of colorectal cancer and promotes tumor growth by recruiting and binding to eIF4A3. *Oncotarget*. 2016;7(16): 22159–22173.
40. Chen SW, Zhu J, Ma J, et al. Overexpression of long non-coding RNA H19 is associated with unfavorable prognosis in patients with colorectal cancer and increased proliferation and migration in colon cancer cells. *Oncol Lett*. 2017;14(2):2446–2452.
41. Torres S, Garcia-Palmero I, Bartolomé RA, et al. Combined miRNA profiling and proteomics demonstrates that different miRNAs target a common set of proteins to promote colorectal cancer metastasis. *J Pathol*. 2017;242(1):39–51.
42. Wu PP, Zhu HY, Sun XF, Chen LX, Zhou Q, Chen J. MicroRNA-141 regulates the tumour suppressor DLC1 in colorectal cancer. *Neoplasma*. 2015;62(5):705–712.
43. Stegh AH. Targeting the p53 signaling pathway in cancer therapy – the promises, challenges and perils. *Expert Opin Ther Targets*. 2012; 16(1):67–83.
44. Duchartre Y, Kim YM, Kahn M. The Wnt signaling pathway in cancer. *Crit Rev Oncol Hematol*. 2016;99:141–149.

Supplementary material

Table S1 All the 64 DEIncRNAs in the ceRNA network

DEIncRNAs	Ensembl ID	log ₂ FC	FDR	DEIncRNAs	Ensembl ID	log ₂ FC	FDR
ABCA9-AS1	ENSG00000231749	3.9864	1.50E-03	ERVMER61-I	ENSG00000230426	6.2730	4.00E-03
AC007384.1	ENSG00000237513	-2.4205	6.92E-14	FAM95B1	ENSG00000223839	-2.0743	4.64E-06
AC009065.1	ENSG00000259780	2.4120	7.23E-03	FRMD6-AS2	ENSG00000258537	-3.1025	7.09E-09
AC009093.1	ENSG00000259807	3.0580	8.28E-05	GAS6-AS1	ENSG00000233695	3.5499	1.09E-07
AC009336.1	ENSG00000272729	-2.0597	6.03E-06	HI9	ENSG00000130600	4.1017	5.86E-04
AC010336.2	ENSG00000260500	-2.7897	2.57E-14	HCG23	ENSG00000228962	-2.5508	3.36E-13
AC012640.1	ENSG00000248968	2.4459	2.90E-04	HULC	ENSG00000251164	6.8799	6.67E-03
AC016745.1	ENSG00000234997	3.6141	4.61E-03	JAZF1-AS1	ENSG00000234336	-2.6263	2.05E-13
AC020907.1	ENSG00000179066	3.8949	3.35E-03	KCNQ1-AS1	ENSG00000229414	2.53016	6.52E-03
AC064836.2	ENSG00000272966	2.5934	4.09E-03	KCNQ1OT1	ENSG00000269821	2.0209	1.35E-05
AC084262.1	ENSG00000260253	3.2217	4.57E-03	LIFR-AS1	ENSG00000244968	-2.3659	2.13E-13
AC110491.1	ENSG00000261292	-4.8260	3.22E-22	LINC00092	ENSG00000225194	-2.8438	8.99E-30
ADAMTS9-AS1	ENSG00000241158	-3.9253	2.66E-28	LINC00330	ENSG00000235097	-2.9422	1.32E-07
ADAMTS9-AS2	ENSG00000241684	-2.8713	5.15E-21	LINC00402	ENSG00000235532	-2.7266	1.25E-09
AGAP11	ENSG00000271880	-2.1666	1.83E-07	LINC00460	ENSG00000233532	6.2479	1.30E-08
AL138995.1	ENSG00000277763	-2.4069	1.42E-17	LINC00461	ENSG00000245526	-2.9817	1.73E-13
AL161431.1	ENSG00000275216	5.7502	7.72E-07	LINC00473	ENSG00000223414	-3.3201	6.64E-17
AL360004.1	ENSG00000196979	-2.3594	1.39E-06	LINC00484	ENSG00000229694	-3.1173	2.08E-16
AL512652.1	ENSG00000275485	2.02862	2.74E-04	LINC00488	ENSG00000214381	-3.1742	1.39E-10
AL513123.1	ENSG00000236347	3.0582	5.80E-05	LINC00507	ENSG00000256193	-4.1785	5.34E-24
AL590483.1	ENSG00000229960	2.8803	7.85E-06	LMO7-AS1	ENSG00000261105	2.4135	6.21E-05
AP000553.1	ENSG00000272954	3.3636	2.14E-05	MIR22HG	ENSG00000186594	-2.0251	9.64E-22
AP002478.1	ENSG00000266401	4.3909	1.13E-04	MYO16-AS1	ENSG00000236242	3.8592	6.26E-03
AP004609.1	ENSG00000245869	-2.0890	7.59E-05	PCAT1	ENSG00000253438	2.4776	4.83E-07
C17orf77	ENSG00000182352	4.1557	1.43E-03	POU6F2-AS1	ENSG00000224122	5.8205	4.29E-03
C20orf166-AS1	ENSG00000174403	-3.4682	8.37E-20	PVT1	ENSG00000249859	2.6328	2.82E-18
C2orf48	ENSG00000163009	2.2264	3.27E-05	PWRNI	ENSG00000259905	-2.2617	1.55E-04
C7orf69	ENSG00000136275	3.2131	9.71E-03	RBMS3-AS3	ENSG00000235904	-2.7161	1.60E-13
CLDN10-AS1	ENSG00000223392	4.5630	6.72E-04	SFTA1P	ENSG00000225383	-2.7508	1.59E-11
CRNDE	ENSG00000245694	4.4129	1.47E-12	ST7-OT4	ENSG00000214188	2.4463	6.43E-04
DLX6-AS1	ENSG00000231764	4.3820	1.38E-03	UCA1	ENSG00000214049	3.9110	1.06E-04
EGOT	ENSG00000235947	2.3163	2.49E-03	WASIR2	ENSG00000231439	2.1913	8.49E-03

Abbreviations: FDR, false discovery rate; log₂FC, log₂fold change.

OncoTargets and Therapy

Dovepress

Publish your work in this journal

OncoTargets and Therapy is an international, peer-reviewed, open access journal focusing on the pathological basis of all cancers, potential targets for therapy and treatment protocols employed to improve the management of cancer patients. The journal also focuses on the impact of management programs and new therapeutic agents and protocols on

Submit your manuscript here: <http://www.dovepress.com/oncotargets-and-therapy-journal>

patient perspectives such as quality of life, adherence and satisfaction. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.