Cellular Physiology and Biochemistry

Cell Physiol Biochem 2018;48:2549-2562
DOI: 10.1159/000492699
Published online: 16 August, 2018
Accepted: 7 August, 2018

Original Paper

Identification and Validation of Two Novel Prognostic lncRNAs in Kidney Renal Clear Cell Carcinoma

Jukun Song[a] Juxiang Peng[b] Chen Zhu[b] Guohui Bai[c] Yongda Liu[d] Jianguo Zhu[e] Jianguo Liu[c]

[a]Department of Oral and Maxillofacial Surgery, Guizhou Provincial People's Hospital, Guiyang, [b]Guiyang Hospital of Stomatology, Guiyang, [c]Special Key Laboratory of Oral Diseases Research, Stomatological Hospital Affiliated to Zunyi Medical College, Guiyang, [d]Department of Urology, Minimally Invasive Surgery Center, the First Affiliated Hospital of Guangzhou Medical Universtiy, Guangzhou, [e]Department of Urology, Guizhou Provincial People's Hospital, Zunyi, China

Key Words
Long non-coding RNAs • Kidney Renal Clear • Cell Carcinoma • Competitive endogenous RNA

Abstract
Background/Aims: Kidney renal clear cell carcinoma (KIRC) is one of the most fatal malignancies due to late diagnosis and poor treatment. To improve its prognosis, a screening for molecular biomarkers of KIRC is urgently needed. Long non-coding RNAs (lncRNAs) play important roles in tumorigenesis and prognosis of cancers. However, it is not clear whether lncRNAs can be used as molecular biomarkers in predicting the survival of KIRC patients. ***Methods:*** In this study, our aim was to identify lncRNAs/mRNAs signatures and their prognostic values in KIRC. The aberrant expression profile of mRNAs and lncRNAs in 529 KIRC tissues and 72 adjacent non-tumor pancreatic tissues were obtained from the Cancer Genome Atlas (TCGA). A weighted gene co-expression network analysis (WGCNA) of two key lncRNAs was constructed. We constructed an aberrant lncRNA-mRNA-miRNA ceRNA network in CESC. In addition, Gene Ontology (GO) and KEGG pathway analysis were performed. ***Results:*** Using lncRNA/mRNA expression profiling data, the overall analysis revealed that two novel lncRNA signatures (DNM1P35 and MIR155HG) and several mRNAs were found to be significantly correlated with KIRC patient's overall analysis. Based on the target gene of the two lncRNA in co-expression network, the GO and KEGG analysis were also performed. A dysregulated lncRNA-related ceRNA network was also observed. ***Conclusion:*** These results suggested that the two novel lncRNAs signatures may act as independent prognostic biomarkers for predicting the survival of KIRC patient.

J. Song, J. Peng and C. Zhu contributed equally to this work.

Jianguo Zhu
and Jianguo Liu

Department of Urology, Guizhou Provincial People's HospitalGuiyang; Special Key Laboratory of Oral Diseases Research, Stomatological Hospital Affliated to Zunyi Medical College, Zunyi (China)
E-Mail doctorzhujianguo@163.com; liujg_001@163.com

Introduction

Kidney cancer (KC), a genitourinary type of cancer, accounts for 2%-3% of all adult malignancies worldwide [1, 2]. It has one of the highest incidence rates among the urinary system tumors. The morbidity and mortality of KC are increasingly and the most common subtype of KC is kidney renal clear cell carcinoma (KIRC)[3]. In 2017, approximately 63, 990 new cases and 14, 440 deaths associated with kidney cancer were reported in the United States [4]. The low survival rates are due to late diagnosis of KIRC and ineffective therapeutic methods currently available. To improve prognosis and decrease mortality and morbidity due to KIRC, diagnostic biomarkers are required to facilitate early detection and risk stratification of KIRC and to guide decisions related to the choice of proper treatments. However, numerous studies indicated that KIRC is heterogeneous in various aspects including clinicopathological, molecular, and cellular heterogeneity. Therefore, identification of molecular biomarkers or therapeutic targets to conquer KIRC is imperative.

Due to rapid technological advancements in genome-wide sequencing, research of clinical disease and pathological mechanisms in various cancers has made tremendous achievement [5]. The sequencing of the human genome produced remarkable findings that protein-coding genes comprise <3% of human DNA. Yet over 80% of our genome is actively transcribed to a set of RNA transcripts without protein-coding functions [6, 7]. Long non-coding RNAs (lncRNAs) are broadly defined as RNAs over 200 nucleotides (nt) in length, which have many structural features of the mRNAs [8-11]. Accumulating evidence has demonstrated that lncRNAs play important molecular roles in regulating gene expression at the level of transcription and post-transcriptional, and chromatin modification [12-14]. Although small ncRNAs, in particular miRNAs, have been extensively studied for over 20 years and many aspects of their biology have been unraveled, still very little is known about the functions of lncRNAs. In this study, aberrant expression profiles of lncRNAs and mRNAs in 529 KIRC patients and 72 non-tumor samples were acquired from the TCGA database and a panel of lncRNAs and mRNAs were detected. Additionally, aberrant lncRNA related ceRNA network was constructed via WGCNA in KIRC. This study can help clinicians to understand the function of lncRNAs through lncRNA-associated ceRNA network in KIRC and provide new lncRNAs as novel diagnostic biomarkers.

Materials and Methods

Raw data

The mRNA expression profiles and corresponding clinical information of KIRC patients were obtained from TCGA data portal (https://tcga-data.nci.nih.gov/tcga/), which was imputed on IlluminaHiSeq RNA-Seq platform, containing 529 KIRC tissues and 72 adjacent non-tumor pancreatic tissues. The lncRNA expression profiles were identified based on the annotation from the GENCODE project (http://www.gencodegenes.org) [15]. Both mRNA profiles data and clinical characteristics of KIRC are publicly available and in open-access platforms. Therefore, approval by local ethics committee was not needed.

Screening of differentially expressed lncRNAs and mRNAs

The differential expression of lncRNAs (DElncRNAs) and mRNAs (DEmRNAs) between CESC and adjacent tissue were calculated using R/Bioconductor package of edgeR[16]. The differentially expressed genes (DEGs) of the data set with |log2 fold change| ≥ 1.0 and P-value less than 0.05 were selected for subsequent analysis.

Functional enrichment analysis of DEmRNAs

To better understand the biological effects and pathways of the aberrantly expression DEmRNA, Gene Ontology (GO) Biological Process, Kyoto Encyclopedia of Genes, and Genomes (KEGG) pathway analyses were conducted using the R/Bioconductor package of Clusteprofiler [17]. Functional enrichment analysis was based on the threshold of P-value <0.05.

Prognostic analysis

A univariate Cox model was employed to impute the relationship between the expression level of each lncRNAs/mRNAs and patient's overall survival (OS). The lncRNAs with P-values less than 0.05 were considered to be statistically significant in univariate Cox analysis. Thereafter, multivariate Cox analysis was employed to evaluate the contribution of lncRNAs/mRNAs as independent prognostic factors of patient survival. These analyses were conducted using the R package of survival and KMsurv.

Weighted co-expression network construction with WGCNA and target prediction

We analyzed the incorporated network using weighted gene co-expression network analysis (WGCNA), which enables the description of the correlation patterns and gene expression profiles [18, 19]. The WGCNA R package was used to evaluate the significance of the two lncRNAs and their module membership. We assessed the weighted co-expression relationship among all dataset subjects in an adjacency matrix using the pairwise Pearson correlation. The appropriate soft threshold power was automatically calculated and generated as described for the standard scale-free network. In the study, the soft threshold was set at $\beta=11$ (scale-free R2= 0.85). Following the identification of weighted correlation, characteristics of the network were presented by Cytoscape 3.4.0[20]. We also predicted the target genes of the two lncRNAs via co-expression network.

Functional enrichment analysis of lncRNAs target genes

The target genes of the candidate lncRNAs were predicted via co-expression network using WGCNA. The enrichment analysis of these target genes was determined using the R/Bioconductor package of Clusteprofiler. Functional enrichment analysis was conducted at the threshold of P-value <0.05.

Validation of the differentially expressed lncRNAs with GEO data

To verify the DElncRNAs/DEmRNAs obtained from TCGA database, we screened the mRNA datasets of KIRC on the GEO database. To identify eligible studies, we employed the following search strategies: "Kidney renal clear cell carcinoma" or "KIRC". The lncRNAs expression level was also extracted for further analysis. The differentially expressed genes were also imputed using the R package of Limma.

ONCOMINE and CCLE analysis

The two DElncRNAs in different types of cancers and cell lines were determined through analysis on the ONCOMINE database (www.oncomine.org) and CCLE database (https://portals.broadinstitute.org/ccle/home), which is a publicly accessible online microarray database to facilitate discovery and identification of genome-wide expression analyses.

Prediction of lncRNA-related ceRNA

The target miRNA of lncRNAs were predicted and minimum free energy (MFE) of miRNA-lncRNA duplexes was imputed using the RNAhybrid program. The data of lncRNA-miRNA interactions were downloaded from highly reliable online miRNA reference database of miRcode (http://www.mircode.org/) [21]. The integrated lncRNA-miRNA pairs were predicted using miRcode, in combination with the selected miRNA. The filter threshold of selected miRNA was max energy <= - 20 and score> 160. Then, the prediction of targeted mRNAs of miRNAs was retrieved from databases of starBase [22]. The lncRNA-associated ceRNA was constructed and visualized using Cytoscape 3.3.2[20]

Results

Pre-treated data

The mRNA expression and corresponding clinical information of KIRC patients were obtained from TCGA data portal (https://tcga-data.nci.nih.gov/tcga/), which was based on the IlluminaHiSeq RNA-Seq platform; containing 529 KIRC tissues and 72 adjacent non-tumor pancreatic tissues. The raw counts of KIRC mRNA expression profiles (level 3 data) were downloaded from the TCGA databases. After initial screening, 16, 315 genes were obtained.

The mRNA expression profiles were transformed lncRNA expression profiles according to the annotation of GENCODE project. Both mRNA profiles data and clinical characteristics of PDAC are publicly available and in open-access platforms. Therefore, approval by a local ethics committee was not needed.

Screening of differentially expressed lncRNAs/mRNAs in KIRC

Analysis of mRNAs/lncRNAs expression profiles in KIRC patients tissues (n=529) and in normal tissues (n=72) identified 4, 544 differentially expressed genes, including differentially expressed mRNAs (n=4497) and lncRNAs (n=47) (Fig. 1). Of these, 3032 genes were over-expressed and 1512 genes were down-expressed. Among the differentially expressed lncRNA, 43 were over-expressed while 4 were down-expressed.

Functional enrichment analysis of DEmRNAs

To explore the functional implication of DEmRNAs, we performed the functional enrichment analysis of GO and KEGG category. The results of GO analysis revealed that GO-enriched categories in the biological process (n=1031), cellular component (n=80), and molecular function (n=117) were among the up-regulated genes. The top five GO-enriched categories are shown in Fig. 2A. A total of 52 KEGG pathways related to biological pathways were enriched among the upregulated genes, including hsa04940; type I diabetes mellitus, hsa05330; allograft rejection, hsa05332; graft-versus-host disease, hsa05150; Staphylococcus aureus infection, and hsa04060; cytokine-cytokine receptor

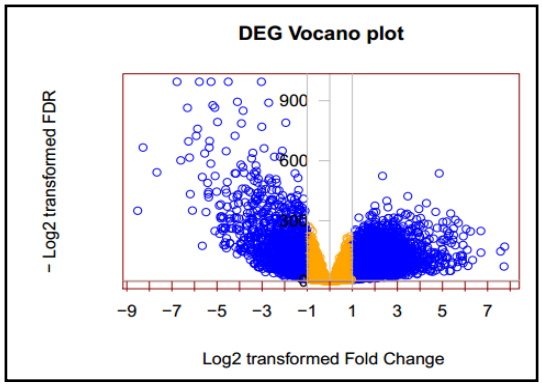


Fig. 1. Volcano plot of the differentially expressed genes. Blue color indicates differentially expressed genes with |log2 fold change| ≥ 1.0 and P-value < 0.05.

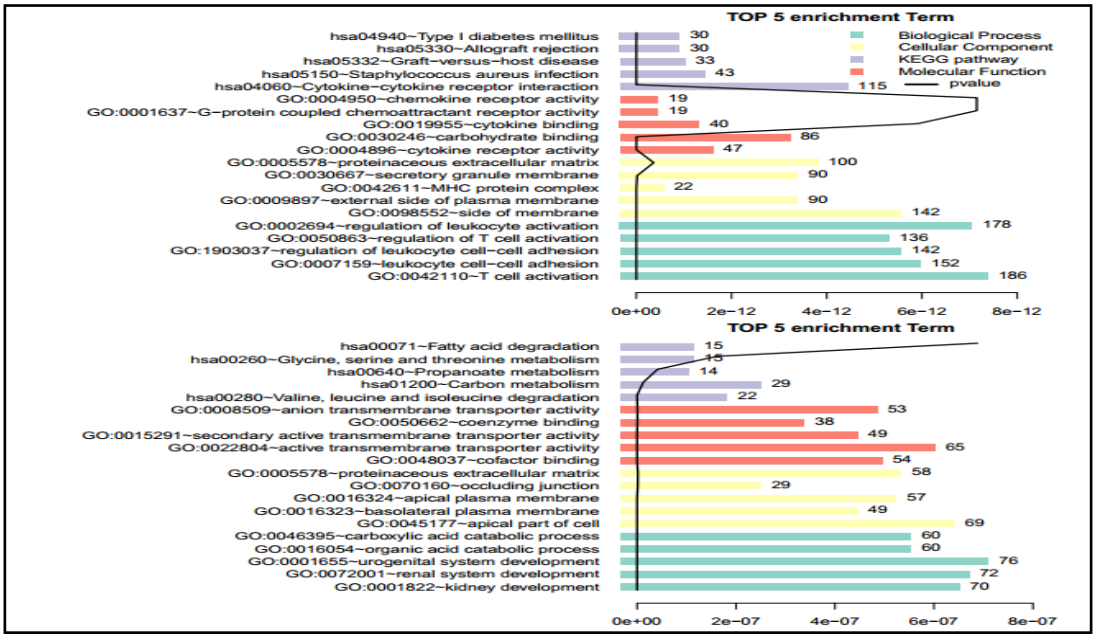


Fig. 2. The top five of GO terms and pathways. Above and below showing upregulated and downregulated DEmRNAs, respectively.

interaction. The results of GO analysis revealed GO-enriched categories in the biological process (n=347), cellular component (n=48), and molecular function (n=76) among the downregulated genes. The top five GO categories are shonw in Fig. 2B. A total of 36 KEGG pathways related to the biological pathways were enriched among the down-regulated genes, including hsa00280; valine, leucine, and isoleucine degradation, hsa01200; carbon metabolism, hsa00640; propanoate metabolism, hsa00260; glycine, serine, and threonine metabolism, and hsa00071; fatty acid degradation.

Prognostic assessment of DElncRNAs/DEmRNAs profiles and clinical characteristics

The univariate Cox regression was performed between DEmRNAs/DElncRNAs profiles and KIRC patients, and the results showed that 932 DEmRNAs and DElncRNAs were significantly associated with overall survival (OS) (P<0.05). In addition, multivariate Cox proportional regression was performed, and we found that 2 DElncRNAs and 78 DEmRNAs are independent prognostic indicators of KIRC (Table 1). The high expression of two DElncRNAs (DNM1P35 and MIR155HG) was significantly associated with overall survival (P<0.05) and exhibited positive effects (See Fig. 3, Table 2). To further examine the association between clinical features and prognosis of KIRC patients, we conducted univariate and multivariate Cox proportional regression. The results are shown in Table 3. In univariate Cox analysis, STAGE, Primary Tumor-Lymph Nodes-Metastasis (TNM), Metastasis (M), and TNM, Tumor (T) were significantly associated with OS of KIRC patients. However, only TNM and M were significantly associated with OS of KIRC patients in the multivariate Cox analysis.

Table 1. multivariate Cox regression analysis for DEmRNAs/DElncRNAs

GENE	z	Pr(>|z|)	HR	lower .95 CI	upper .95 CI
NTN5.126147	12.19952683	0	3.08197E+12	3.24E-13	30365823159
DAND5.199699	2.305107202	0.02116057	14290711362	7.00E-11	33.15233195
CCDC154.645811	2.5422388	0.011014489	3967357315	2.52E-10	157.921916
SPC24.147841	2.005602176	0.04489871	217454121.8	4.60E-09	1.547817557
GSG2.83903	4.619587573	3.85E-06	29199840.07	3.42E-08	19862.08382
KIAA1751.85452	2.702769962	0.006876432	3538398.747	2.83E-07	63.06923325
KIR3DL2.3812	2.654365794	0.007945764	122971.9579	8.13E-06	21.45512661
KIF14.9928	3.31932974	0.000902338	14416.54726	6.94E-05	50.48696484
C19orf35.374872	7.194473204	6.27E-13	11387.39047	8.78E-05	894.0203852
ACRC.93953	2.018081802	0.043582743	393.4687612	0.002541498	1.187762575
RFX8.731220	4.462430509	8.10E-06	211.6676026	0.004724389	20.14621038
DNM1P35.100128285	6.88901596	5.62E-12	184.7264857	0.005413409	41.84935869
SLC46A2.57864	2.416639361	0.015664529	134.4850386	0.007435771	2.524981468
KIF19.124602	2.360379384	0.018256254	123.3742125	0.008105422	2.263371662
NEK2.4751	2.916382649	0.003541159	89.79681353	0.011136253	4.370849735
BIRC5.332	2.450085553	0.014282228	76.26165933	0.013112749	2.379805703
C11orf35.256329	2.287471633	0.022168309	75.6248348	0.013223169	1.857705537
ATAD5.79915	7.665621903	1.78E-14	71.17200222	0.014050469	23.91698747
CENPI.2491	2.112468878	0.034646253	59.08051924	0.016926053	1.34241068
AGER.177	9.743421932	0	58.85309657	0.016991459	25.92785913
BRIP1.83990	8.413027803	0	45.60954225	0.021925236	18.7304957
LILRA2.11027	6.661521776	2.71E-11	43.3130434	0.023087733	14.29190209
PRSS36.146547	5.997315922	2.01E-09	39.91612632	0.025052531	11.96415722
CTLA4.1493	2.164790164	0.030403758	37.12412616	0.026936661	1.407720159
C17orf56.146705	2.053378029	0.040035928	34.9844756	0.028584107	1.175534312
PBK.55872	2.552003603	0.010710541	32.34339146	0.030918217	2.240037714
C7orf53.286006	5.204250003	1.95E-07	28.81653836	0.034702294	8.126967483
YPEL4.219539	5.240353704	1.60E-07	25.40633521	0.039360262	7.576602174
DNHD1.144132	4.950855672	7.39E-07	24.67361368	0.040529126	6.935379726
CDC6.990	5.844255669	5.09E-09	16.04076471	0.062341168	6.324609189
TRIM46.80128	2.877116629	0.004013272	12.59502614	0.079396421	2.242414431
KIFC1.3833	5.422624466	5.87E-08	12.39903469	0.08065144	4.991046599
CHEK2.11200	2.420179979	0.015512827	11.75166498	0.085094325	1.597672283
KLRA1.10748	3.170680295	0.001520824	11.21686761	0.089151449	2.517044144
FAM72B.653820	2.275818345	0.022856884	9.933711631	0.100667307	1.375264644
E2F1.1869	3.818919341	0.000134038	8.855395077	0.112925509	2.89119489
MIR155HG.114614	1.966408188	0.049251478	5.158421696	0.193857745	1.005391063
PRC1.9055	5.157125624	2.51E-07	4.754986139	0.210305555	2.629050281
TBC1D3B.414059	6.07734797	1.22E-09	4.271648978	0.234101633	2.674400786
CSAD.51380	7.014674294	2.30E-12	3.755325539	0.266288499	2.594691996
LOC100129637.100129637	3.379853768	0.000725244	3.018593475	0.331280117	1.590624988
CACNB1.782	3.3862544	0.000708537	3.000687683	0.333256942	1.588566529
TAZ.6901	7.407599	1.29E-13	2.978921007	0.335692017	2.231662564
NBPF9.400818	9.273790424	0	2.713152825	0.368574889	2.197161021
SPHK1.8877	2.136610756	0.032629664	2.603337852	0.384122253	1.082316972
SLC17A9.63910	6.859553794	6.91E-12	2.303251316	0.434168861	1.814723246
LOC150776.150776	3.144974807	0.001661011	2.245102622	0.445413938	1.356263417
CNTNAP1.8506	4.037768108	5.40E-05	2.218508274	0.450753334	1.506885769
C19orf48.84798	2.588200482	0.009647881	1.900039563	0.52630483	1.16859584
AP1G2.8906	5.675000698	1.39E-08	1.777835859	0.562481623	1.457427846
ENGASE.64772	3.142676429	0.001674108	1.721206472	0.580987822	1.226741902
ZNF692.55657	2.167070295	0.030229495	1.704913133	0.586540147	1.052310056
SRM.6723	9.005689118	0	1.69965789	0.588353695	1.514350863
QTRT1.81890	2.062210245	0.039187724	1.640477695	0.609578541	1.024845562
PPFIA4.8497	4.701273023	2.59E-06	1.618873361	0.617713543	1.324316784
UBXN11.91544	3.213728919	0.001310233	1.618168053	0.617982785	1.206551662
PLAUR.5329	9.283725821	0	1.614429847	0.619413722	1.459158712
ASAP1.50807	3.759319685	0.000170376	1.602477364	0.624033776	1.253202696

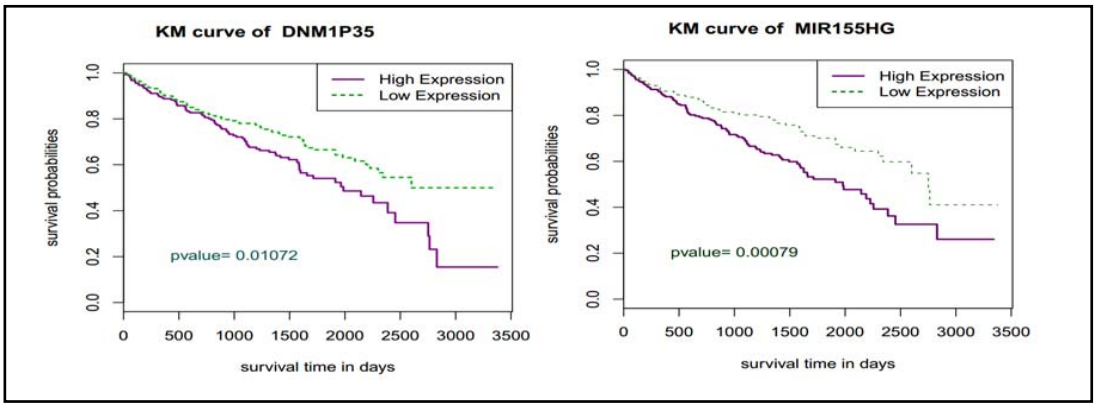


Fig. 3. Kaplan-Meier survival curves for overall survival outcomes according to the risk cutoff point for two lncRNAs. The p value of the log-rank test was less than 0.01.

Table 2. Two lncRNAs were significantly correlated with overall survival in multivariate Cox regression analysis

GENE	z	Pr(>|z|)	HR	lower .95 CI	upper .95 CI
DNM1P35	6.88901596	5.62E-12	184.7264857	41.84935869	815.3977887
MIR155HG	1.966408188	0.049251478	5.158421696	1.005391063	26.46663112

Table 3. Univariate and multivariate Cox regression analysis for clinical features in KIRC cohort

Clinical features	Univariate analysis		Multivariate analysis	
	P-value	HR(95%CI)	P-value	HR(95%CI)
GENDER	1.0578(0.8827-1.2676)	0.541811087		
AGE	0.9999(0.9927-1.0071)	0.976078039		
STAGE	1.1269(1.0457-1.2144)	0.015231324	0.172389844	0.8770(0.7264-1.0589)
TNM_M	1.6316(1.3782-1.9315)	1.04E-07	2.56E-07	1.6871(1.3828-2.0584)
TNM_N	1.0160(0.9300-1.1100)	0.118369952		
TNM_T	1.0461(1.0143-1.0788)	0.007735265	0.084015258	1.0674(0.9913-1.1493)
GRADE	1.1142(0.9890-1.2553)	0.078317061		

Target prediction and functional enrichment of the two lncRNAs

The biological functions of lncRNAs were still unknown. Therefore, to explore the target genes of the two key lncRNAs, we employed the Weighted Co-expression network construction to examine. A total of 11 modules were detected in the co-expression network. The lncRNA DNM1P35 and MIR155HG were included in the turquoise module and yellow module, respectively. The number of genes in the turquoise module was 466, of which 113 genes were identified to be potentially regulated by the lncRNA DNM1P35 in the co-expression network (see Fig. 4). Among these target genes in co-expression network, ACRC, AGER, AP1G2, C17orf56, CCDC14, CSAD, DNHD1, ENGASE, GOLGA8B, LOC100129637, LOC150776, LOC339047, LPIN3, UBXN11, ZNF692, and DNM1P35 were up-regulated in KIRC and hence may act as independent prognostic factors for the OS of KIRC patients. The number of genes in the yellow module was 221, of which 107 genes were identified to be potentially regulated by the lncRNA MIR155HG in the co-expression network (see Fig. 4). Among these target genes in the co-expression network, CTLA4, MIR155HG, DNM1P35, ACRC, AGER,

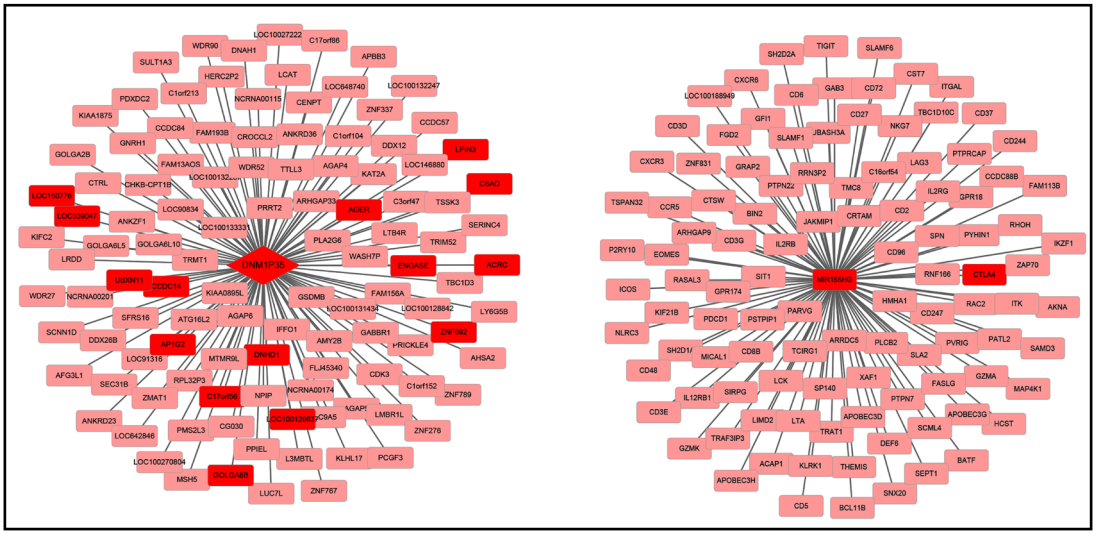


Fig. 4. The target genes of two key lncRNAs in the coexpression network (the crimson in the Fig. is the prognostic risk gene, the diamond represents lncRNAs, and the red represents the up-regulation).

AP1G2, C17orf56, CCDC14, CSAD, DNHD1, ENGASE, GOLGA8B, LOC100129637, LOC150776, LOC339047, LPIN3, UBXN11, and ZNF692 were up-regulated in KIRC and hence may act as independent prognostic factors for the OS of KIRC patients.

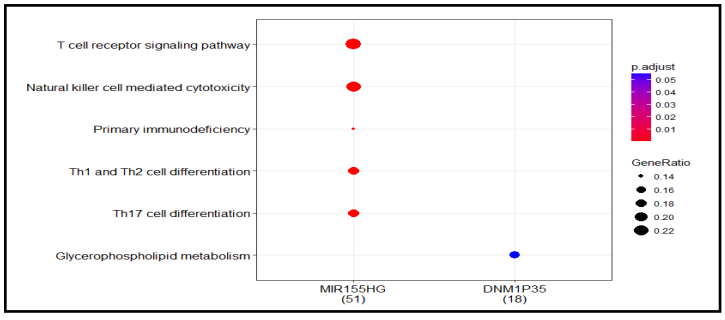


Fig. 5. Comparison of GO and KEGG enrichment results for target genes of two lncRNAs.

The enrichment analysis was conducted to describe the biological function of the target genes of two key lncRNAs. We found enrichment of 334 gene ontology categories. There are 238 GO terms with the biological process (BP), 26 GO terms with cellular component, and 70 GO categories with molecular function. A total of 14 KEGG pathways were enriched by target genes of two key lncRNAs. Comparison of the enrichment analysis of the two lncRNAs demonstrated that their target gene MIR155HG was enriched in T cell receptor signaling pathway (See Fig. 5).

Validation of DEmRNAs/DElncRNAs with GEO data

One study (GSE53757) was considered eligible for GEO. Analysis of mRNA expression profiles identified a total of 3135 differentially expressed mRNA. Among these genes, 1628 DEmRNAs were over-expressed and 1507 DEmRNAs were down-expressed. In both TCGA KIRC and GSE53757 studies, 1127 genes were over-expressed while 736 genes were down-expressed. Among significant prognostic risk of 78 mRNAs and 2 lncRNAs, 51 genes which FDR was less than 0.01 were up-regulated. As for expression profiles of the two lncRNAs, only lncRNA DNM1P35 was detected in the GSE53757. However, the differential expression profiles of lncRNA DNM1P35 were not statistically significant in tumor and normal group. lncRNA MIR155HG was not detected in GSE53757.

Expression level of two lncRNAs in Oncomine and CCLE database

Two lncRNAs (DNM1P35 and MIR155HG) were found in human solid tumors (Fig. 6). Only two studies reported the expression of lncRNA DNM1P35 in breast cancer [23] and kidney cancer of TCGA cohort. The study on TCGA kidney cancer suggested that the expression of lncRNA DNM1P35 was significantly higher in clear cell renal cell carcinoma than other types of renal cell carcinoma as shown in Fig. 7. In Oncomine database, several studies reported the expression level of lncRNAs MIR155HG in clear cell renal cell carcinoma samples [24, 25]. The pooled results revealed that lncRNAs MIR155HG was significantly higher in KIRC compared to other cancer samples (cancer vs. cancer, P<0.05) (Fig. 8). In addition, the CCLE database analysis demonstrated that the lncRNA expression level of DNM1P35 and MIR155HG was high in many cancer cell-lines (Fig. 9).

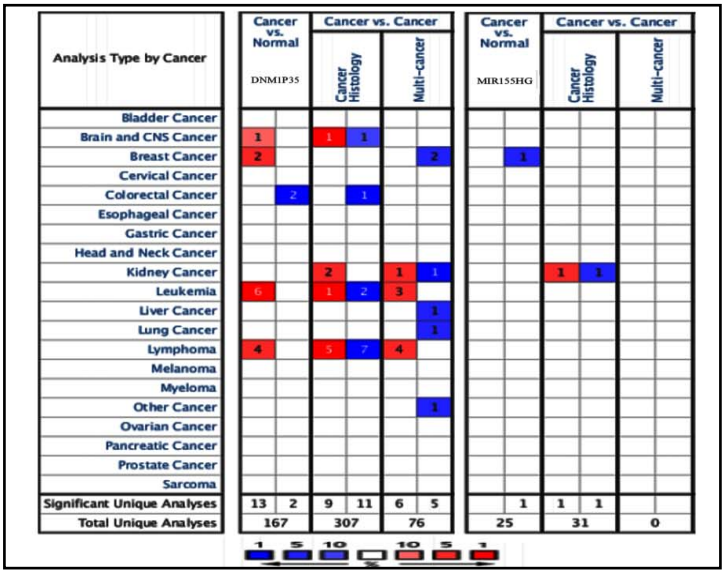


Fig. 6. The mRNA expression levels of two lncRNAs (DNM1P35 and MIR155HG) in different tumor types in Oncomine database.

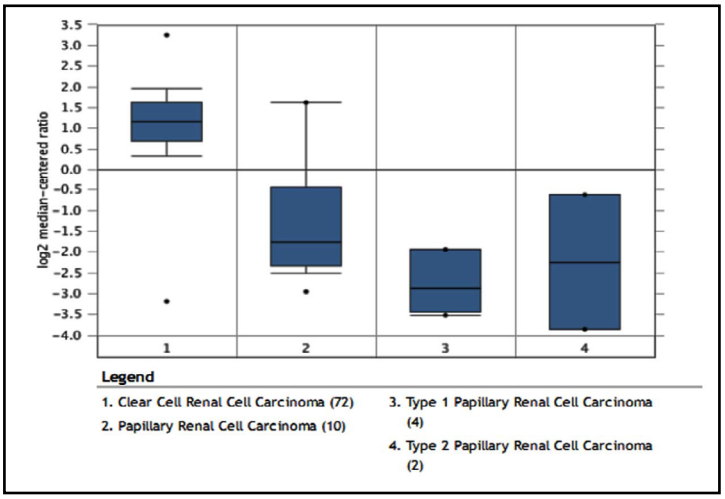
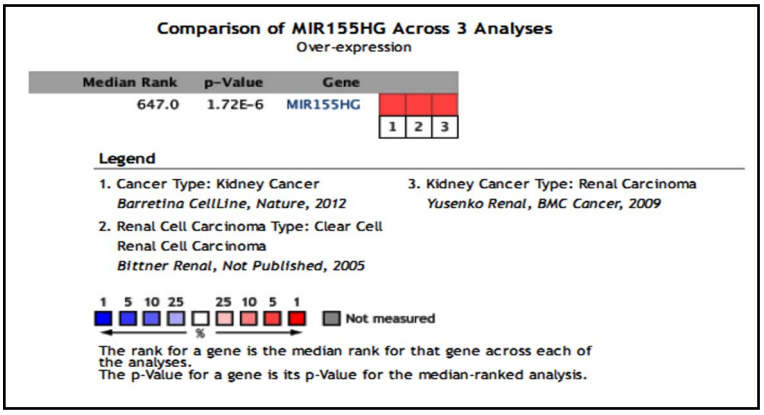


Fig. 7. The expression level of lncRNA DNM1P35 was significantly higher in Clear Cell Renal Cell Carcinoma compared to other types of Renal Cell Carcinoma.

Fig. 8. The combined result of lncRNA MIR155HG expression across three studies in Oncomine database.



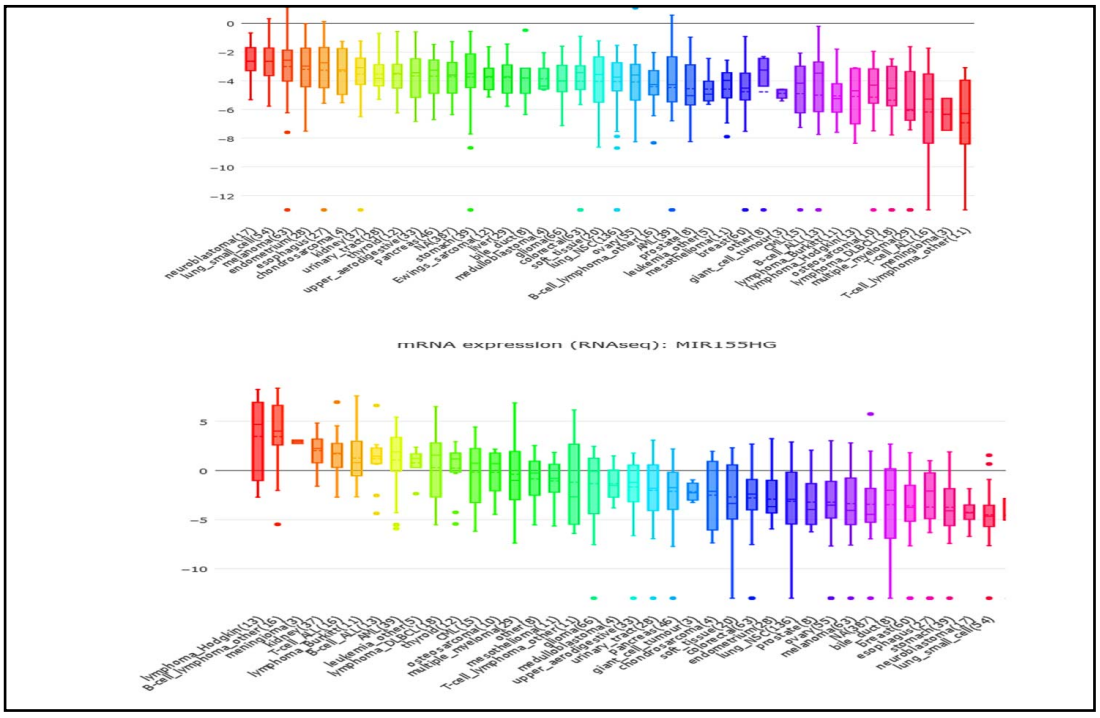


Fig. 9. The two lncRNAs (DNM1P35 and MIR155HG) were highly expressed in kidney cancer cell lines as revealed from CCLE analysis.

Construction and analysis of lncRNA-related ceRNA network

A total of 295 microRNAs that bind to DNM1P35 were identified at the set threshold, and 115 microRNAs which are associated with MIR155HG were screened. The miRNA-lncRNA network is illustrated in Fig. 10. The starBase database was used to explore the downstream target genes of microRNAs that bind to lncRNA, and those which are prognostic genes for KIRC. A ceRNA network was constructed based on a combination of microRNAs, lncRNAs, and target genes. The network contains 62 nodes and 135 edges, in which 32 miRNAs, 2 lncRNAs, and 28 target genes were detected (Fig. 11). The combination of MIR135HG with hsa-miR-155-5p and hsa-miR-223-3p is also mentioned in the starBase database.

Discussion

Kidney Renal Clear Cell Carcinoma (KIRC) is an aggressive and malignant type of kidney cancer which has poor prognosis and is characterized by a complex molecular and cellular heterogeneity [26]. Although KIRC is curable at early-stage by surgery resection, the prognosis of metastatic KIRC is poor. Researchers have demonstrated that the 5-year survival rate of KIRC patients is less than 10%[27]. In the past, great efforts have been made to provide insights into the molecular mechanisms underlying KIRC, but the focus has been on protein-coding genes or miRNA[28, 29]. Therefore, understanding the molecular mechanism of KIRC may provide clinicians with new treatment strategies for this disease. To date, there are no conventional molecular biomarkers for KIRC prognosis. In biological processes, lncRNAs acts through miscellaneous mechanisms [30, 31]. It has been documented that lncRNAs are a novel class of non-protein-coding transcripts found in cancer biogenesis and prognosis. Aberrant lncRNAs expression has been reported in many cancers, suggesting that they may function as oncogenes or tumor suppressors [32-34]. Integration genomic studies have demonstrated that lncRNAs play important roles which have attracted growing research attention. Up-to-date, some lncRNAs are thought to improve the clinical outcome of

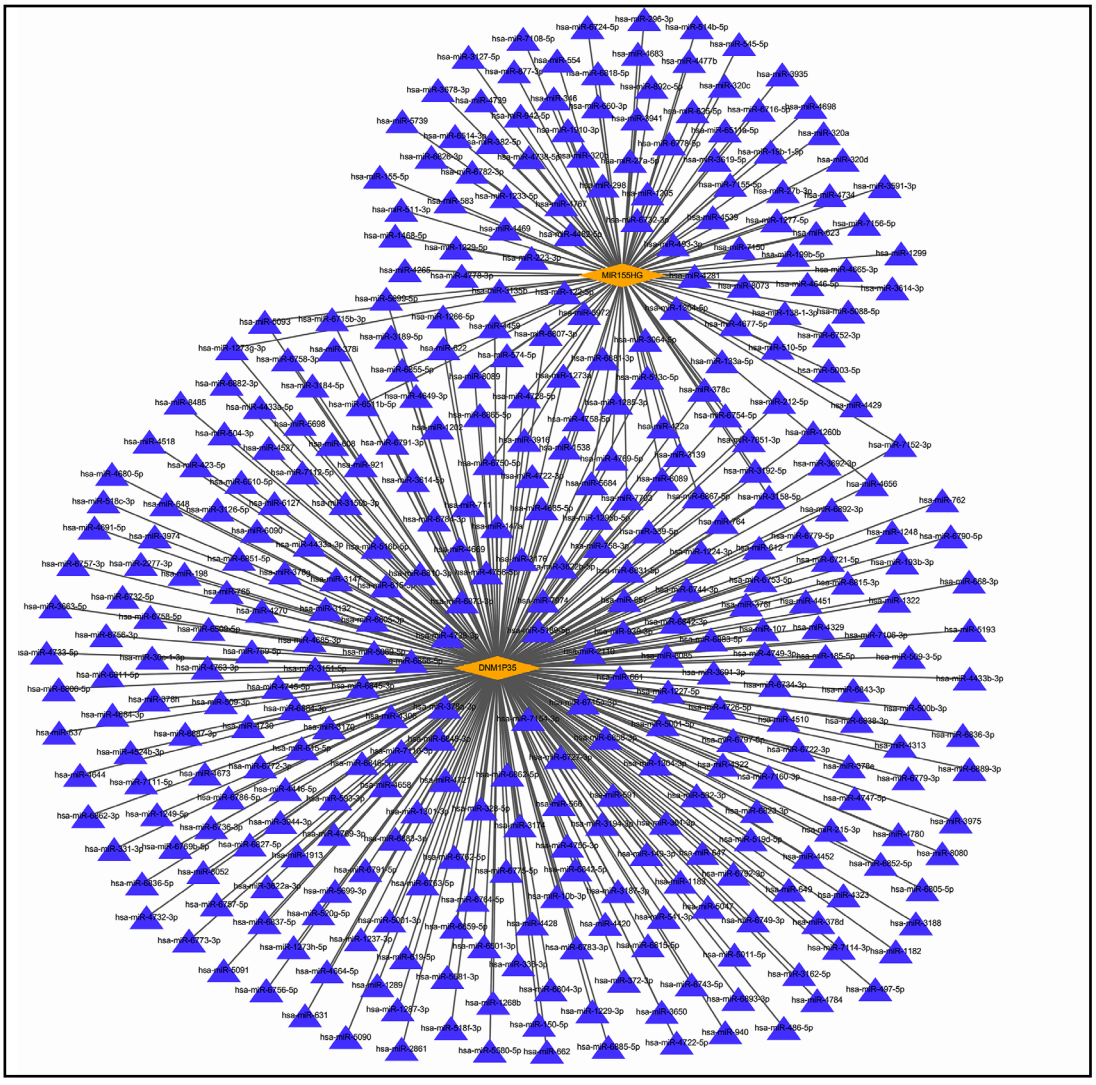


Fig. 10. The lncRNA-miRNA based ceRNA network (triangle represents miRNAs, and diamond represents lncRNAs).

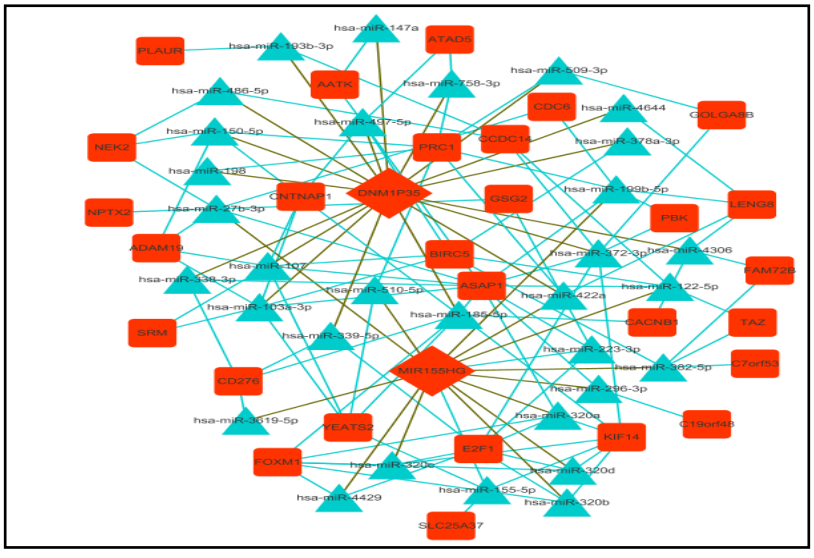


Fig. 11. The integrated ceRNA network (squares are target genes, triangles are miRNAs, diamonds are lncRNAs, blue lines indicates that microRNAs regulate target genes, dark yellow lines represent the reaction between miRNAs and lncRNAs).

KIRC patients [35, 36]. However, no specific biomarker has been found to display therapeutic efficiency for KIRC, and therefore there is need for specific prognostic factors to improve the treatment of KIRC patients.

In the present work, we compared gene expression profiles between KIRC samples and adjacent non-tumor tissues to identify potential lncRNA molecular biomarkers using a database of TCGA. The differentially expressed mRNAs/lncRNAs were identified after which the univariate and multivariate Cox analysis was performed to determine whether these mRNAs/lncRNAs can be used for prediction of KIRC prognosis. We found two key lncRNAs and several mRNAs as key predictors of KIRC prognosis. Among the lncRNAs, overexpression of lncRNAs (DNM1P35 and MIR155HG) was associated with poor prognosis of patients with KIRC. We further used the GEO data to confirm these findings. Indeed, the results revealed that DEmRNAs/lncRNAs can be independent prognostic predictors for patients with KIRC. Through a Weighted Co-expression network using WGCNA, we found that two lncRNAs were correlated with 230 protein-coding genes. In addition, a lncRNA-miRNA-mRNA ceRNA network was constructed based on the starBase and miRanda database. Using the enrichment and functional analysis using the R/Bioconductor package of Clusteprofiler, we found that MIR155HG and it's target gene Cytotoxic T lymphocyte antigen 4 (CTLA-4) were up-regulated in KIRC and both were significantly associated with OS in patients with KIRC. CTLA-4, a transmembrane protein encoded by the CTLA-4 gene, is expressed in activated CD4 + and CD8 + T cells. CTLA-4 binds to its ligand B7 molecule to produce inhibitory signals that prevent activation of T cell and protect tumor cells from T lymphocytes. Thus, blocking the immune effect of CTLA-4 can stimulate cell proliferation of many immunity cells, thereby inducing or enhancing the anti-tumor immune response [37-39]. The target gene of DNM1P35 did not show enrichment function. However, the expression of DNM1P35 and its target genes: ACRC, AGER, AP1G2, C17orf56, CCDC14, CSAD, DNHD1, ENGASE, GOLGA8B, LOC100129637, LOC150776, LOC339047, LPIN3, UBXN11, and ZNF692 were significantly increased in KIRC and are prognostic risk genes for this disease.

Several studies have reported that lncRNA MIR155HG plays key roles in the development of myeloproliferative disorders, leukemia, lymphoma, and glioma [40-42]. lncRNA MIR155HG can suppress lymphoma by regulating the activity of NF-κB and mesenchymal genes [42]. However, the contribution of lncRNA DNM1P35 to the development of KIRC is still not known, and further studies are required to address these aspects.

The expression of two incRNAs (MIR155HG and DNM1P35) and target genes of CTLA-4, ACRC, AGER, AP1G2, C17orf56, CCDC14, CSAD, DNHD1, ENGASE, GOLGA8B, LOC100129637, LOC150776, LOC339047, LPIN3, UBXN11, and ZNF692 was significantly increased in KIRC and thus they are prognostic risk genes for this disease. Among these key mRNAs, diseases associated with AGER include Diabetic Angiopathy and Thymic Hyperplasia. Among its related pathways are activated TLR4 signaling and a-beta signaling pathways. GO annotations related to this gene include identical protein binding and transmembrane signaling receptor activity [43, 44].

AP1G2 has been reported to be involved in cardiac arrest and long Qt syndrome 1. Among its related pathways are CTLA-4 signaling and clathrin derived vesicle budding. GO annotations related to this gene include binding and protein transporter activity [45, 46].

CSAD was associated with Disuse Amblyopia. Among its related pathways are β-alanine metabolism (TR) and sulfur amino acid metabolism. GO annotations related to this gene include pyridoxal phosphate binding and carboxy-lyase activity [47].

Diseases associated with ENGASE include congenital isorder of deglycosylation. Among its related pathways is transport to the golgi and subsequent modification and metabolism of proteins. GO annotations related to this gene include mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase activity [48].

LPIN3 (Lipin 3) is a protein-coding gene. Among its related pathways are mitotic prophase and synthesis of PC. GO annotations related to this gene include transcription coactivator activity and phosphatidate phosphatase activity [49].

Several limitations of the study should be considered. Firstly, the two lncRNAs signature were analyzed and validated only in the TCGA and GEO data set, and no other KIRC-related lncRNAs expression profile was used for further validation. Secondly, the TCGA data was obtained from a single central source, and the ethnic composition of the population in TCGA database was mainly white and black, hence the findings in the work cannot be extrapolated to other ethnicities. Thirdly, no experimental data on the mechanisms of lncRNA have been reported, and further experimental studies are needed to improve our understanding of the functional role of the lncRNA in KIRC.

Findings in this study demonstrated that two key lncRNAs and several mRNAs can be used as molecular biomarkers and prognostic factors to predict the survival rate in patients with KIRC. However, the biological function of the two lncRNAs needs to be further validated through experiments.

Acknowledgements

This study was supported by the NSFC (Natural Science Foundation of China) (81560186), NSFC (Natural Science Foundation of China) (81660426), and Science and Technology Project of Guizhou Province in 2017([2017]5803). Thank Dr. Chris Lou (Blogger, WeChat Official Accounts) for the bioinformatics technology support.

Disclosure Statement

The authors declare to have no competing financial interests.

References

1 Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM: Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010;127:2893-2917.
2 Calvo E, Schmidinger M, Heng DY, Grunwald V, Escudier B: Improvement in survival end points of patients with metastatic renal cell carcinoma through sequential targeted therapy. Cancer Treat Rev 2016;50:109-117.
3 Chaffer CL, Weinberg RA: A perspective on cancer cell metastasis. Science 2011;331:1559-1564.
4 Siegel RL, Miller KD, Jemal A: Cancer Statistics, 2017. CA Cancer J Clin 2017;67:7-30.
5 Stratton MR, Campbell PJ, Futreal PA: The cancer genome. Nature 2009;458:719-724.
6 Consortium EP: An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57-74.
7 Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Roder M, Kokocinski F, Abdelhamid RF, Alioto T, et al.: Landscape of transcription in human cells. Nature 2012;489:101-108.
8 Gloss BS, Dinger ME: The specificity of long noncoding RNA expression. Biochim Biophys Acta 2016;1859:16-22.
9 Heery R, Finn SP, Cuffe S, Gray SG: Long Non-Coding RNAs: Key Regulators of Epithelial-Mesenchymal Transition, Tumour Drug Resistance and Cancer Stem Cells. Cancers (Basel) 2017;9:pii:E38.
10 Li Y, Wang X: Role of long noncoding RNAs in malignant disease (Review). Mol Med Rep 2016;13:1463-1469.
11 Zhang F, Zhang L, Zhang C: Long noncoding RNAs and tumorigenesis: genetic associations, molecular mechanisms, and therapeutic strategies. Tumour Biol 2016;37:163-175.
12 Li CH, Chen Y: Targeting long non-coding RNAs in cancers: progress and prospects. Int J Biochem Cell Biol 2013;45:1895-1910.

13 Fatica A, Bozzoni I: Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet 2014;15:7-21.
14 Rinn JL, Chang HY: Genome regulation by long noncoding RNAs. Annu Rev Biochem 2012;81:145-166.
15 Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, et al.: The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012;22:1775-1789.
16 Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139-140.
17 Yu G, Wang LG, Han Y, He QY: clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284-287.
18 Zhang B, Horvath S: A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 2005;4:Article17.
19 Langfelder P, Horvath S: WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559.
20 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-2504.
21 Jeggari A, Marks DS, Larsson E: miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. Bioinformatics 2012;28:2062-2063.
22 Li JH, Liu S, Zhou H, Qu LH, Yang JH: starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 2014;42:D92-97.
23 Gluck S, Ross JS, Royce M, McKenna EF Jr, Perou CM, Avisar E, Wu L: TP53 genomics predict higher clinical and pathologic tumor response in operable early-stage breast cancer treated with docetaxel-capecitabine +/- trastuzumab. Breast Cancer Res Treat 2012;132:781-791.
24 Yusenko MV, Kuiper RP, Boethe T, Ljungberg B, van Kessel AG, Kovacs G: High-resolution DNA copy number and gene expression analyses distinguish chromophobe renal cell carcinomas and renal oncocytomas. BMC Cancer 2009;9:152.
25 Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, Berger MF, Monahan JE, Morais P, Meltzer J, Korejwa A, Jane-Valbuena J, Mapa FA, et al.: The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603-607.
26 Patard JJ, Leray E, Rioux-Leclercq N, Cindolo L, Ficarra V, Zisman A, De La Taille A, Tostain J, Artibani W, Abbou CC, Lobel B, Guille F, Chopin DK, Mulders PF, Wood CG, Swanson DA, Figlin RA, Belldegrun AS, Pantuck AJ: Prognostic value of histologic subtypes in renal cell carcinoma: a multicenter experience. J Clin Oncol 2005;23:2763-2771.
27 Krambeck AE, Dong H, Thompson RH, Kuntz SM, Lohse CM, Leibovich BC, Blute ML, Sebo TJ, Cheville JC, Parker AS, Kwon ED: Survivin and b7-h1 are collaborative predictors of survival and represent potential therapeutic targets for patients with renal cell carcinoma. Clin Cancer Res 2007;13:1749-1756.
28 Wang C, Cai L, Liu J, Wang G, Li H, Wang X, Xu W, Ren M, Feng L, Liu P, Zhang C: MicroRNA-30a-5p Inhibits the Growth of Renal Cell Carcinoma by Modulating GRP78 Expression. Cell Physiol Biochem 2017;43:2405-2419.
29 Xiao W, Lou N, Ruan H, Bao L, Xiong Z, Yuan C, Tong J, Xu G, Zhou Y, Qu Y, Hu W, Gao Y, Ru Z, Liu L, Xiao H, Chen K, Yang H, Zhang X: Mir-144-3p Promotes Cell Proliferation, Metastasis, Sunitinib Resistance in Clear Cell Renal Cell Carcinoma by Downregulating ARID1A. Cell Physiol Biochem 2017;43:2420-2433.
30 Guttman M, Rinn JL: Modular regulatory principles of large non-coding RNAs. Nature 2012;482:339-346.
31 Nagano T, Fraser P: No-nonsense functions for long noncoding RNAs. Cell 2011;145:178-181.
32 Ponting CP, Oliver PL, Reik W: Evolution and functions of long noncoding RNAs. Cell 2009;136:629-641.
33 Yin D, He X, Zhang E, Kong R, De W, Zhang Z: Long noncoding RNA GAS5 affects cell proliferation and predicts a poor prognosis in patients with colorectal cancer. Med Oncol 2014;31:253.
34 Enfield KS, Pikor LA, Martinez VD, Lam WL: Mechanistic Roles of Noncoding RNAs in Lung Cancer Biology and Their Clinical Implications. Genet Res Int 2012;2012:737416.
35 Qiao HP, Gao WS, Huo JX, Yang ZS: Long non-coding RNA GAS5 functions as a tumor suppressor in renal cell carcinoma. Asian Pac J Cancer Prev 2013;14:1077-1082.

36 Yu G, Yao W, Wang J, Ma X, Xiao W, Li H, Xia D, Yang Y, Deng K, Xiao H, Wang B, Guo X, Guan W, Hu Z, Bai Y, Xu H, Liu J, Zhang X, Ye Z: LncRNAs expression signatures of renal clear cell carcinoma revealed by microarray. PLoS One 2012;7:e42377.
37 Saitou M, Sugimoto J, Hatakeyama T, Russo G, Isobe M: Identification of the TCL6 genes within the breakpoint cluster region on chromosome 14q32 in T-cell leukemia. Oncogene 2000;19:2796-2802.
38 Buchbinder E, Hodi FS: Cytotoxic T lymphocyte antigen-4 and immune checkpoint blockade. J Clin Invest 2015;125:3377-3383.
39 Buchbinder EI, McDermott DF: Cytotoxic T-lymphocyte antigen-4 blockade in melanoma. Clin Ther 2015;37:755-763.
40 Wu X, Wang Y, Yu T, Nie E, Hu Q, Wu W, Zhi T, Jiang K, Wang X, Lu X, Li H, Liu N, Zhang J, You Y: Blocking MIR155HG/miR-155 axis inhibits mesenchymal transition in glioma. Neuro Oncol 2017;19:1195-1205.
41 Cui B, Chen L, Zhang S, Mraz M, Fecteau JF, Yu J, Ghia EM, Zhang L, Bao L, Rassenti LZ, Messer K, Calin GA, Croce CM, Kipps TJ: MicroRNA-155 influences B-cell receptor signaling and associates with aggressive disease in chronic lymphocytic leukemia. Blood 2014;124:546-554.
42 Baytak E, Gong Q, Akman B, Yuan H, Chan WC, Kucuk C: Whole transcriptome analysis reveals dysregulated oncogenic lncRNAs in natural killer/T-cell lymphoma and establishes MIR155HG as a target of PRDM1. Tumour Biol 2017;39:1010428317701648.
43 Hudson BI, Stickland MH, Futers TS, Grant PJ: Effects of novel polymorphisms in the RAGE gene on transcriptional regulation and their association with diabetic retinopathy. Diabetes 2001;50:1505-1511.
44 Goulart AC, Germer S, Rexrode KM, Martin M, Zee RY: Polymorphisms in advanced glycosylation end product-specific receptor (AGER) gene, insulin resistance, and type 2 diabetes mellitus. Clin Chim Acta 2008;398:95-98.
45 Aouizerat BE, Vittinghoff E, Musone SL, Pawlikowska L, Kwok PY, Olgin JE, Tseng ZH: GWAS for discovery and replication of genetic loci associated with sudden cardiac arrest in patients with coronary artery disease. BMC Cardiovasc Disord 2011;11:29.
46 Rost M, Doring T, Prange R: gamma2-Adaptin, a ubiquitin-interacting adaptor, is a substrate to coupled ubiquitination by the ubiquitin ligase Nedd4 and functions in the endosomal pathway. J Biol Chem 2008;283:32119-32130.
47 Winge I, Teigen K, Fossbakk A, Mahootchi E, Kleppe R, Skoldberg F, Kampe O, Haavik J: Mammalian CSAD and GADL1 have distinct biochemical properties and patterns of brain expression. Neurochem Int 2015;90:173-184.
48 Suzuki T, Yano K, Sugimoto S, Kitajima K, Lennarz WJ, Inoue S, Inoue Y, Emori Y: Endo-beta-N-acetylglucosaminidase, an enzyme involved in processing of free oligosaccharides in the cytosol. Proc Natl Acad Sci U S A 2002;99:9691-9696.
49 Reue K, Zhang P: The lipin protein family: dual roles in lipid biosynthesis and gene expression. FEBS Lett 2008;582:90-96.