Integrated analysis of differentially expressed profiles and construction of a competing endogenous long non-coding RNA network in renal cell carcinoma

Qianwei Xing[*], Yeqing Huang[*], You Wu, Limin Ma and Bo Cai

Department of Urology, Affiliated Hospital of Nantong University, Nantong, Jiangsu, China
[*] These authors contributed equally to this work.

ABSTRACT

Background. Long non-coding RNAs (lncRNAs) play crucial roles in the initiation and progression of renal cell carcinoma (RCC) by competing in binding to miRNAs, and related competitive endogenous RNA (ceRNA) networks have been constructed in several cancers. However, the coexpression network has been poorly explored in RCC.

Methods. We collected RCC RNA expression profile data and relevant clinical features from The Cancer Genome Atlas (TCGA). A cluster analysis was explored to show different lncRNA expression patterns. Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses and gene set enrichment analysis (GSEA) were performed to analyze the functions of the intersecting mRNAs. Targetscan and miRanda bioinformatics algorithms were used to predict potential relationships among RNAs. Univariate Cox proportional hazards regression was conducted to determine the RNA expression levels and survival times.

Results. Bioinformatics analysis revealed that the expression profiles of hundreds of aberrantly expressed lncRNAs, miRNAs, and mRNAs were significantly changed between different stages of tumors and non-tumor groups. By combining the data predicted by databases with intersection RNAs, a ceRNA network consisting of 106 lncRNAs, 26 miRNAs and 69 mRNAs was established. Additionally, a protein interaction network revealed the main hub nodes (VEGFA, NTRK2, DLG2, E2F2, MYB and RUNX1). Furthermore, 63 lncRNAs, four miRNAs and 31 mRNAs were significantly associated with overall survival.

Conclusion. Our results identified cancer-specific lncRNAs and constructed a ceRNA network for RCC. A survival analysis related to the RNAs revealed candidate biomarkers for further study in RCC.

Subjects Bioinformatics, Oncology, Urology
Keywords lncRNA, Renal cell carcinoma, ceRNA, Expression profile

INTRODUCTION

Renal cell carcinoma (RCC) is among the ten leading cancer types for estimated new cancer cases in both males and females in the United States (*Siegel, Miller & Jemal, 2017*). Although the five-year relative survival rate for all cancers has increased due to advances in diagnosis and therapy over the past three decades, discriminative biological markers

have not been determined in RCC for early-stage diagnosis and therapy, unlike for other cancers. Increasing numbers of studies on RCC have been conducted to explore tumor initiation and progression, including protein-coding RNAs (mRNAs) and non-coding RNAs (ncRNAs) (*Frew & Moch, 2015*; *Lorenzen & Thum, 2016*; *Martens-Uzunova et al., 2014*; *Qin et al., 2014*; *Schmidt & Linehan, 2016*); however, the tumor-specific mechanisms in the regulation of tumor progression and biological behaviors are not fully understood. Recently, long non-coding RNAs (lncRNAs) have been reported to play important roles in the molecular mechanism of RCC (*Chen et al., 2016*; *Li, Shuch & Gerstein, 2017*; *Serghiou, Kyriakopoulou & Ioannidis, 2016*).

LncRNAs are a subtype of ncRNAs with transcript lengths over 200 nucleotides and were once viewed as transcriptional noise without the capacity to encode proteins (*Guttman et al., 2009*). However, growing evidence suggests that lncRNAs may play crucial biological roles in transcriptional regulation, cellular development, and RNA modification (*Schmitt & Chang, 2016*). Emerging reports have revealed that lncRNAs are associated with various human diseases, such as neurological diseases (*Bai et al., 2014*; *Lau, Frigerio & De Strooper, 2014*), cardiovascular diseases (*Vausort, Wagner & Devaux, 2014*; *Yang et al., 2014*), immune inflammatory diseases (*Kino et al., 2010*), and malignant tumors; moreover, lncRNAs can be drug targets (*Sana et al., 2012*). In urologic malignancies, lncRNAs are thought to be related to tumor formation, invasion, and metastasis (*Martens-Uzunova et al., 2014*). In RCC, several studies have reported that lncRNAs may function as oncogenes or tumor suppressors and that they can affect the long-term survival and mortality of patients. For example, TRIM52-AS1 overexpression can suppress cell migration and proliferation and can induce apoptosis of RCC (*Liu et al., 2016a*; *Liu et al., 2016b*). Additionally, down-regulated lncRNA H19 inhibits RCC carcinogenesis (*Wang et al., 2015*). Tumor formation and development are complex pathophysiological processes, and lncRNAs may interact with miRNAs or mRNAs or other molecules in regulatory capacities. There are multiple hypotheses regarding how cancer-specific molecules communicate with each other (*Gomes, Nolasco & Soares, 2013*; *Huang et al., 2017*).

In 2011, *Salmena et al. (2011)* proposed a competing endogenous RNA (ceRNA) hypothesis, which indicated that all types of RNA transcripts communicate with each other by competing for binding to miRNA response elements (MREs). This competition between mRNAs, lncRNAs, pseudogenes, and circular RNAs widely exists in tumor initiation and progression. *Xiao et al. (2015)* reported that the lncRNA MALAT1 participates in the ceRNA network by sponging miR-200s to regulate ZEB2 expression. Additionally, *Qu et al. (2016)* demonstrated that exosome-transmitted lncARSR acts as a ceRNA to promote sunitinib resistance in renal cancer. Nevertheless, studies incorporating large sample sizes and high-throughput detection methods to identify whether lncRNAs are correlated with overall survival, gender, or other clinical features or to determine the ceRNA mechanisms of lncRNAs with aberrant expression in RCC are currently lacking.

The Caner Genome Atlas (TCGA) (http://cancergenome.nih.gov) is a platform that can provide clinical information and RNA sequencing with miRNA, lncRNA, and mRNA data. We downloaded data sets from the TCGA that included RNA sequences of different stage (according to the American joint committee on cancer, AJCC) tumor tissues and

adjacent non-tumor renal tissues. To the best of our knowledge, this study is the first to use this large sequencing database to investigate the differential expression profiles of a cancer-specific lncRNA and ceRNA coexpression network in RCC. This new approach of predicting caner-specific lncRNAs and their ceRNA potential can help us elucidate the functions of lncRNAs in RCC.

MATERIALS AND METHODS

Patients and samples

A total of 537 patients with RCC were retrieved from the TCGA data portal. We selected data according to the following criteria: (1) the histopathological diagnosis was RCC; (2) complete data was required for all samples of all stages; and (3) the included patients had no other malignancies. Overall, a total of 457 RCC patients were enrolled in this study. Among these samples, we analyzed miRNA sequences from 64 adjacent non-tumor tissue samples and 213 tumor tissue samples (95 stage I samples, 27 stage II samples, 46 stage III samples, and 45 stage IV samples) and mRNA/lncRNA sequences from 65 adjacent non-tumor tissues and 454 tumor tissues (209 stage I samples, 51 stage II samples, 114 stage III samples, and 80 stage IV samples). This study was conducted in accordance with the publication guidelines provided by the TCGA (http://cancergenome.nih.gov/publications/publicationguidelines); therefore, further approval of the ethics committee was not required.

RNA sequencing data sets and computational analysis

The RCC RNA expression profile data (level 3) and relevant clinical features were acquired from the TCGA data portal (through Feb, 2017). The lncRNA and mRNA expression profiles were normalized using the RNASeqV2 system provided by the TCGA database. The miRNA expression profiles, downloaded from the TCGA database, were obtained using the Illumina HiSeq 2000 miRNA sequencing platform (Illumina Inc., San Diego, CA, USA). No further normalization was needed, as the TCGA database already had normalized the lncRNA, miRNA, and mRNA expression profile data. Then, we compared differentially expressed RNAs at four levels, including stage I RCC patient tumor tissues vs. non-tumor renal tissues, AJCC stage II RCC patient tumor tissues vs. non-tumor renal tissues, AJCC stage III RCC patient tumor tissues vs. non-tumor renal tissues, and AJCC stage IV RCC patient tumor tissues vs. non-tumor renal tissues. In the next step, we collected intersecting lncRNAs ($P \leq 0.01$, FDR ≤ 0.01, FC ≥ 2), miRNAs ($P < 0.05$, FDR < 0.05, FC ≥ 2) and mRNAs ($P \leq 0.01$, FDR ≤ 0.01, FC ≥ 3)for hierarchical clustering and further bioinformatics analysis. A flow chart was depicted to show the framework of this study (Fig. 1).

GO, KEGG pathway analysis and GSEA

Gene ontology (GO) analysis contains three domains: biological processes, cellular components, and molecular functions. We used a GO database to analyze differentially expressed intersecting mRNAs (http://www.geneontology.org). The molecular functions of up- and down-regulated genes were identified. The potential functions of the aberrantly
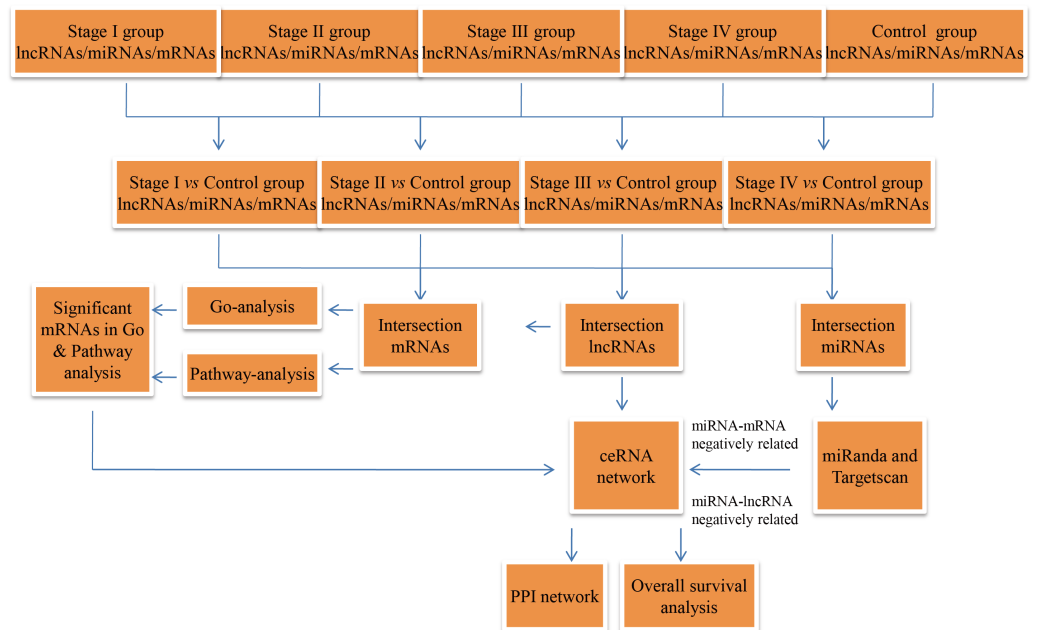
Figure 1 Flow chart of the ceRNA network construction.

Full-size ⬜ DOI: 10.7717/peerj.5124/fig-1

expressed genes involved in the pathways were analyzed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.kegg.jp). Besides, the above data were analyzed using gene set enrichment analysis (GSEA) (Broad Institute, Cambridge, MA, USA, accessed at http://software.broadinstitute.org/gsea/). Additionally, we identified intersecting mRNAs ($P < 0.05$, FDR < 0.05) by combing GO with KEGG analyses to build a ceRNA network.

Construction of a lncRNA-related ceRNA network

According to the ceRNA hypothesis, lncRNA can regulate mRNA expression by sequestering and binding miRNAs like miRNA sponges. We constructed the ceRNA network through three steps: (1) we identified intersecting lncRNAs, miRNAs, and mRNAs that were differentially expressed in four stages (including up- and down-regulation); (2) we predicted lncRNA-miRNA interactions by miRanda (http://www.microrna.org/microrna/home.do) and used the Targetscan (http://www.targetscan.org) and miRanda databases to find target genes; and (3) we integrated aberrantly expressed data from the TCGA and the predicted miRNA information. Through these steps, we constructed a ceRNA network with intersecting lncRNAs and mRNAs that were negatively regulated by miRNAs. Additionally, a figure for the ceRNA network was depicted using Cytoscape v3.0.

PPI network analysis

To clarify the potential relationships between the aberrantly expressed genes and RCC, we performed a protein-protein interaction (PPI) network analysis with the online software String (score >0.4).
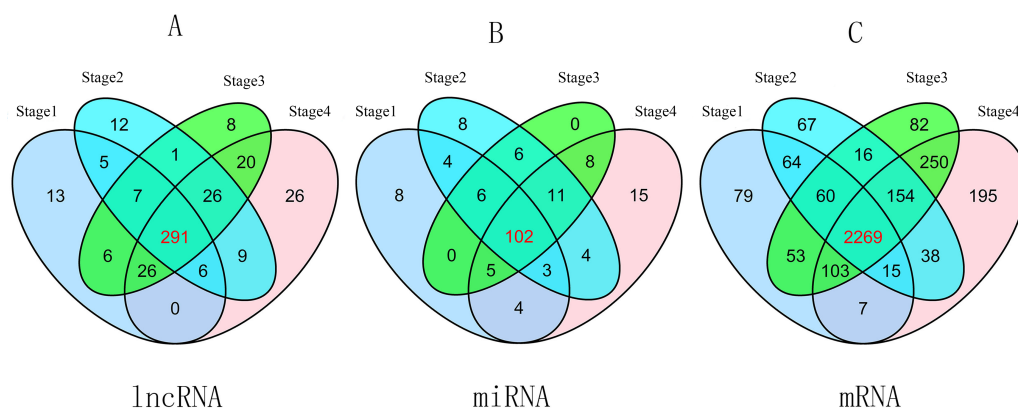
Figure 2 **Venn diagram of aberrant expression profiles of lncRNAs (A), miRNAs (B) and mRNAs (C) between stage I RCC group (tumor tissues according to AJCC)/control group (non-tumor tissues), stage II group/control group, stage III group/control group, and stage IV group/control group.** A total of 291 intersecting lncRNAs, 102 intersecting miRNAs and 2,269 mRNAs were identified.

Full-size ☒ DOI: 10.7717/peerj.5124/fig-2

Statistical analysis

Differentially expressed lncRNAs, miRNAs, and mRNAs were analyzed by BRB array tool software (https://brb.nci.nih.gov/BRB-ArrayTools/download.html). To identify the relationships between lncRNAs/miRNAs/mRNAs in the ceRNA network and to determine overall survival, we conducted univariate Cox proportional hazards regression and performed Kaplan–Meier survival analyses and log-rank tests to identify the relationships between RNA expression levels and survival time ($P < 0.05$ was considered statistically significant).

RESULTS

Cancer-specific lncRNAs in RCC

To compare stage I RCC tumor tissues and adjacent non-tumor RCC tissues from the TCGA database, we found 354 differentially expressed lncRNAs ($P \leq 0.01$, FDR ≤ 0.01, FC ≥ 2). Similarly, we identified 357 differentially expressed lncRNAs between stage II RCC tumor tissues and adjacent non-tumor RCC tissues. Three hundred eighty-five differentially expressed lncRNAs were identified between stage III RCC tumor tissues and non-tumor RCC tissues, and 404 differentially expressed lncRNAs were identified between stage IV RCC tumor tissues and non-tumor RCC tissues. Then, we found 291 intersecting lncRNAs from the above lncRNAs (92 up-regulated; 199 down-regulated) (Fig. 2A, Table S1). A heat map of these 291 aberrant lncRNAs was established using R software (Fig. S1), which revealed different expression patterns of lncRNAs in RCC.

GO, KEGG pathway analyses and GSEA

A total of 2,650 mRNAs ($P \leq 0.01$, FDR ≤ 0.01, FC ≥ 3) were differentially expressed between stage I RCC tumors and non-tumor RCC tissues from the TCGA database, 2,683 mRNAs were differentially expressed between stage II RCC tumors and non-tumor RCC tissues, 2,987 mRNAs were differentially expressed between stage III RCC tumors and

non-tumor RCC tissues, and 3,031 mRNAs were differentially expressed between stage IV RCC tumors and non-tumor RCC tissues. Next, we identified 2,269 intersecting mRNAs (Fig. 2C) from the above mRNAs (1,132 up-regulated; 1,147 down-regulated).

We further explored the biological functions of these up- and down-regulated protein-coding RNAs by GO and KEGG pathway enrichment analyses. Using a GO database for comprehensive analysis, we found that the top three processes of the up-regulated mRNAs involved immune responses, signal transduction, and inflammatory responses, while the top three processes of the down-regulated mRNAs were transmembrane transport, small molecule metabolic processes, and excretion (Fig. 3A). The significant pathways of the up-regulated mRNAs included cytokine-cytokine receptor interactions, Staphylococcus aureus infection, and cell adhesion molecules, which were determined by KEGG pathway enrichment analysis (Fig. 3B). The pathways correlated with the down-regulated mRNAs included metabolic pathways, neuroactive ligand–receptor interactions, protein digestion, and absorption (Fig. 3B). In addition, GSEA of up and down-regulated mRNAs was implemented using the gene sets of GO and KEGG pathway database (Tables S2–S5). Finally, we found 738 significant intersecting mRNAs by GO and KEGG analyses ($P < 0.05$, FDR < 0.05).

Construction of the ceRNA network

We found 132 RCC-related miRNAs that were differentially expressed between stage I tumors and non-tumor tissues ($P < 0.05$, FDR < 0.05, FC ≥ 2). Additionally, 144 miRNAs were differentially expressed between stage II tumors and non-tumor tissues, 138 miRNAs were differentially expressed between stage III tumors and non-tumor tissues, and 152 miRNAs were differentially expressed between stage IV tumor and non-tumor tissues. Subsequently, we identified 102 differentially expressed miRNAs from the above four groups (41 up-regulated; 61 down-regulated) (Fig. 2B).

Integrating 291 intersecting lncRNAs and 102 intersecting miRNAs with those that we identified from the prolife TCGA data, we finally chose 106 lncRNAs that were negatively related to 26 miRNAs that were predicted by the miRanda database. These 26 miRNAs were negatively related to 69 mRNAs following a comparison of data predicted by Targetscan and miRanda with the data from GO and KEGG analyses. Based on these data, we established a miRNA-lncRNA-mRNA ceRNA network (Fig. 4).

Construction of the PPI

To better understandthe critical genes predicted in RCC, we established a PPI network that included 44 genes with scores of >0.4 in String (Fig. 5). From the figure, we determined that the main hub nodes were VEGFA, NTRK2, DLG2, E2F2, MYB and RUNX1. Hence, the key genes associated with RCC can be predicted by our network.

RNAs in the ceRNA network are related to survival

To identify the miRNAs associated with prognosis, all miRNAs in the ceRNA network were analyzed via univariate Cox proportional hazards regression, and four miRNAs (miR-9-5p, miR-21-5p, miR-155-5p, and miR-244-5p) were observed to be significantly changed (log-rank $P < 0.05$). Additionally, we found 63 lncRNAs and 31 mRNAs that were
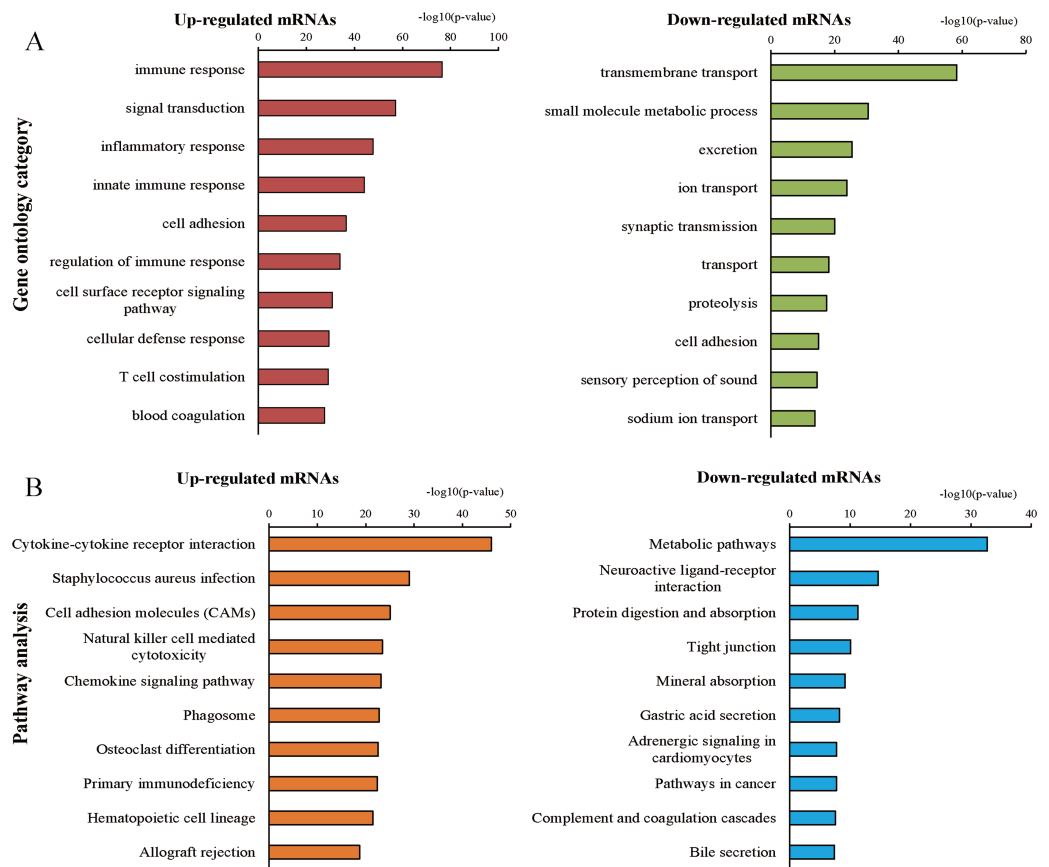
Figure 3 Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses. (A) Top 10 processes of up- and down-regulated genes in the GO analysis. (B) Top 10 pathways of up- and down-regulated genes in the pathway analysis.

Full-size ◳ DOI: 10.7717/peerj.5124/fig-3

significantly related to overall survival (Table S6). Here, we show figures for RNAs (Fig. 6), as the coexpression profiles of RNAs inthe lncRNA-miRNA-mRNA correlations were all significantly associated with overall survival. These relationships are shown in Table 1.

DISCUSSION

LncRNAs are emerging as having crucial roles in various cell biological processes and diverse malignant tumors (*Martens-Uzunova et al., 2014*; *Schmitt & Chang, 2016*). Although their exact mechanisms are unclear, a variety of theories and hypotheses have been described regarding how lncRNAs participate in regulating malignant biological behavior. Some studies have documented that lncRNAs can affect cancer-related mRNAs by interacting with miRNAs, yielding new perspectives regarding lncRNA functions (*Lv et al., 2016*; *Yue et al., 2016*). This hypothesis has been supported *in vivo* and *in vitro*; moreover, ceRNA coexpression networks have been established for several human tumors (*Li et al., 2016a*; *Li et al., 2016b*; *Liu et al., 2016a*; *Liu et al., 2016b*; *Wang et al., 2016a*; *Wang et al., 2016b*; *Yang et al., 2016*). A few studies have reported interactions between lncRNAs and miRNAs in
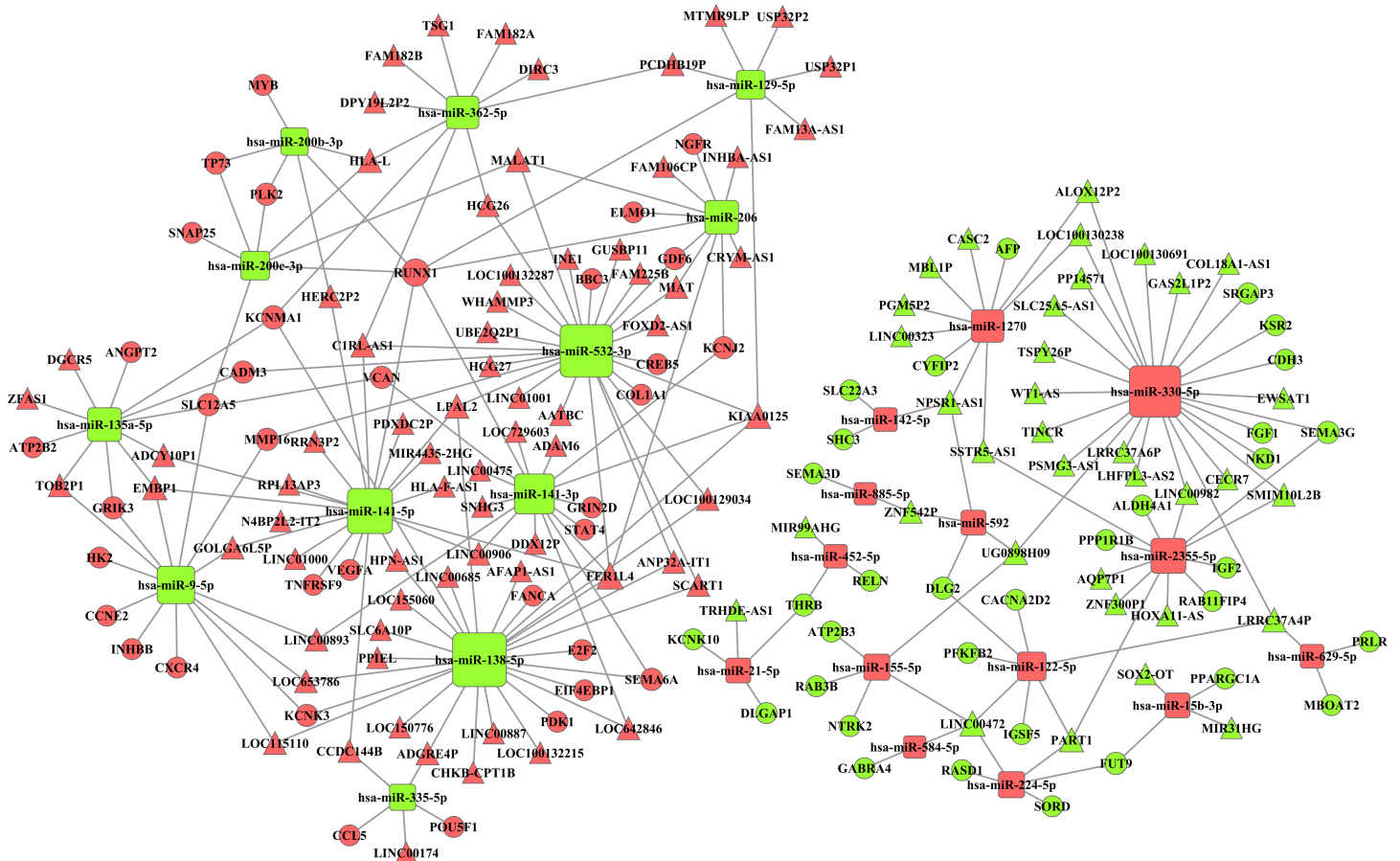
Figure 4 **The miRNA-lncRNA-mRNA ceRNA network is composed of 106 lncRNAs, 26 miRNAs and 69 mRNAs.** Red diamonds represent up-regulated miRNAs, while green diamonds represent down-regulated miRNAs. Triangles and circles represent lncRNAs and mRNAs, respectively, while red indicates up-regulation, and green indicates down-regulation.

RCC. Furthermore, one previous study suggested that some lncRNAs are associated with chromophobe RCC and that they may function as ceRNAs in complex ceRNA networks (*He et al., 2016*). Nevertheless, coexpression networks based on large-scale bioinformatics data and the potential roles of lncRNAs in ceRNA networks are still poorly understood in RCC. In the present study, we confirmed aberrantly expressed lncRNAs, miRNAs and mRNAs based on the RNA expression profiles of 457 RCC patient tissue samples from the TCGA. Additionally, we constructed a ceRNA network with cancer-specific lncRNAs, miRNAs and mRNAs. To understand the systematic biological roles of the abnormally expressed RNAs, we further investigated the association between these RNAs and overall survival, and we also constructed a PPI network to find key mRNAs.

After analyzing the RNA sequencing data from the TCGA, we finally identified 291 cancer-specific lncRNAs that were abnormally expressed in different stages of RCC. We then constructed a cluster analysis map to show the differential distributions of the 291 lncRNAs in RCC tumor tissues and adjacent non-tumor tissues. Among these lncRNAs, several known lncRNAs, such as GAS5, CASC2, TCL6, MALAT-1, UCA1 and HOTAIR,
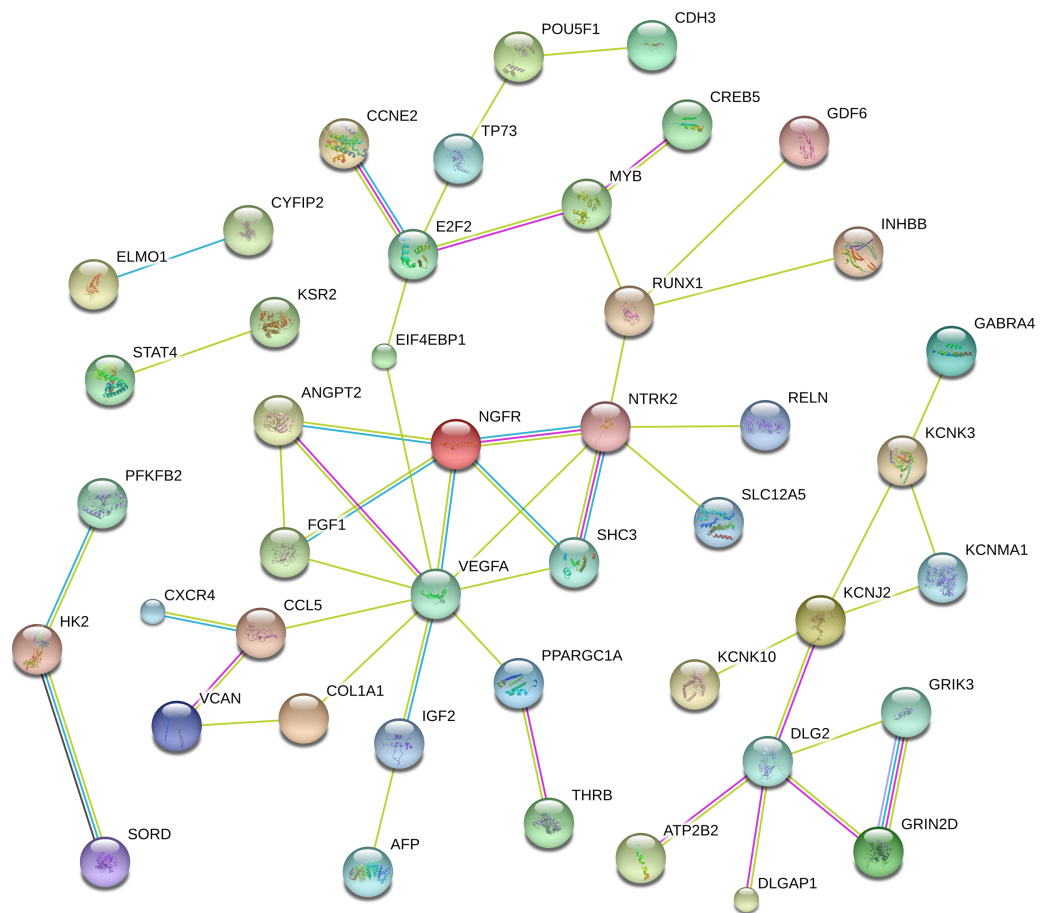
Figure 5 The protein-protein interaction network included 44 genes with scores of >0.4 in String.
Full-size ☒ DOI: 10.7717/peerj.5124/fig-5

have been found to be related to RCC cell proliferation, invasion and apoptosis, and they are important regulators in RCC occurrence and progression (*Cao et al., 2016*; *Chen et al., 2017*; *Hong et al., 2017*; *Li et al., 2016a*; *Li et al., 2016b*; *Qiao et al., 2013*; *Su et al., 2017*). Furthermore, the massive RNA sequencing data identified in the present study revealed hundreds of differentially expressed lncRNAs that have not yet been reported and may function as novel oncogenes and theranostic markers in RCC.

The GO and KEGG pathway analyses of the 2,269 the intersecting mRNAs revealed the main biological processes and pathways in RCC. Many of them were classical pathways comprising important areas of RCC research, such as immune responses, signal transduction, and metabolism (*Labrousse-Arias et al., 2017*; *Sakai, Miyake & Fujisawa, 2013*; *Wettersten et al., 2017*). However, some of biological processes and pathways, such as protein digestion and absorption, were reported in other malignancies (*Dong et al., 2017*; *Li et al., 2017*). From our enrichment analysis of large-scale samples focusing on RCC, we discovered multiple processes involved in neoplasm formation, which was supplementary to the results of previous studies. To enhance the validity and reliability of the ceRNA
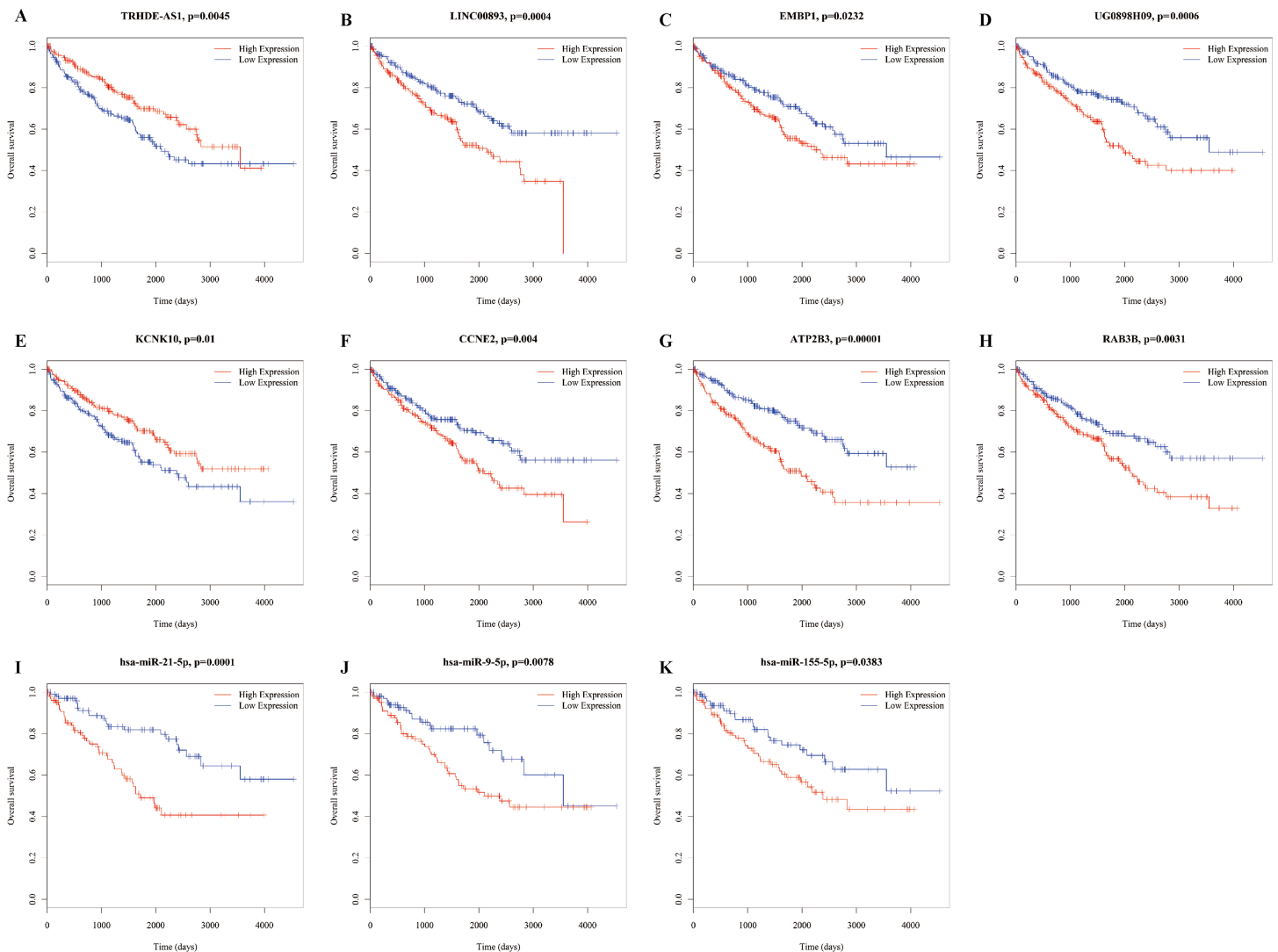
Figure 6 **Kaplan–Meier survival curves for 11 related RNAs (4 lncRNAs, 3 miRNAs and 4 mRNAs) in the ceRNA network (horizontal axis: overall survival times: days, vertical axis: survival function).** High expression of LINC00893 (B), EMBP1 (C), UG0898H09 (D), CCNE2 (F), ATP2B3 (G), RAB3B (H), has-miR-21-5p (I), has-miR-5p (J) and has-miR-155-5p (K) is associated with low proportion of overall survival ($P = 0.0004$, 0.0232, 0.0006, 0.004, 00001, 0.0031, 0.0001, 0.0078, 0.0383, respectively); high expression of TRHDE-AS1 (A) and KCNK10 (E) is associated with high proportion of overall survival ($P = 0.0045$ and 0.001, respectively).

Full-size 🖼 DOI: 10.7717/peerj.5124/fig-6

network, we selected the mRNAs that significantly participatedin both the GO and KEGG pathways for further analyses.

To provide a comprehensive view on how these lncRNAs participate in multiple biological processes, we constructed a ceRNA network combining the intersecting miRNAs and mRNAs, as many studies have elucidated the possibility that lncRNAs can be ceRNA regulators andcan communicate with other RNA transcripts (*Lv et al., 2016*; *Qu et al., 2016*; *Salmena et al., 2011*; *Xiao et al., 2015*; *Yue et al., 2016*). To increase the accuracy of the ceRNA network, we integrated the intersecting RNAs in different stages of RCC

Table 1 RNAs significantly associated with overall survival in lncRNA-miRNA-mRNA pathway.

lncRNA	miRNA	mRNA
TRHDE-AS	miR-21-5p	KCNK10
Lin00893	miR-9-5p	CCNE2
EMBP1	miR-9-5p	CCNE2
UG0898H09	miR-155-5p	ATP2B3
UG0898H09	miR-155-5p	PAB3B

and evaluated the relationships between lncRNAs and miRNAs and between miRNAs and mRNAsthat were predicted by the Targetscan and miRanda databases. Ultimately, 106 lncRNAs, 26 miRNAs and 69 mRNAs were involved in the ceRNA network. Recent studies have demonstrated that down-regulation of the lncRNA CASC2 by miRNA-21 increases RCC cell proliferation and migration (*Cao et al., 2016*) and that MALT1 promotes aggressive RCC through Ezh2 and interacts with miR-205 (*Hirata et al., 2015*). The key lncRNAs identified above were also found in our ceRNA network. Additionally, our study integratively predicted the potential relationships between novel lncRNAs, miRNAs and mRNAs in RCC. Construction of a PPI network increases the reliability of our ceRNA network in the following ways: First, some genes described as main hub nodes, such as VEGFA, NTRK2, DLG2, E2F2, MYB and RUNX1, are themselves vital genes in human tumors, and some of them are directly associated with RCC (*Chimge et al., 2016*; *Gao et al., 2016*; *Gonda & Ramsay, 2016*; *Jones et al., 2013*; *Ma et al., 2015*; *Zubakov, Stupar & Kovacs, 2006*). Second, some genes in the ceRNA network that have not been reported to be correlated with RCC may also be oncogenes, as they can interact with RCC-related genes. The reasons above validate the hypotheses that the genes in the ceRNA network caninfluence RCC and that lncRNAs can be regulators that modulate oncogene expression.

To further evaluate the roles of RNAs in RCC diagnosis and prognosis, we analyzed the associations between RNAs in the ceRNA network and patient survival. We found four miRNAs, 63 lncRNAs and 31 mRNAs that were significantly correlated with overall RCC survival. Up-regulation of the lncRNA MALAT1 has been reported to correlate with tumor progression and poor prognosis in RCC (*Zhang et al., 2015*), while some lncRNAs have been reported in other tumors, such as HOXA11-AS, which was regarded as a biomarker of poor progression in glioma (*Wang et al., 2016a*; *Wang et al., 2016b*). However, most of the 19 lncRNAs have not yet been reported. Subsequently, we found some RNAs in the lncRNA-miRNA-mRNA correlations that were uniformly related to overall survival, such as miR-21-5p-regulated mRNA KCNK10 and lncRNA TRHDE-AS1, which confirmed the internal relationships constructed in the ceRNA network in the present study.

Our work identified many valuable RNAs that are differentially expressed in RCC tissues via bioinformatics analysis. Some of these RNAs have been verified *in vivo* and *in vitro* experiments. However, most of the aberrant RNAs still need to be validated, and our ceRNA network, which was constructed *in silico*, needs to be validated with additional biological experiments.

CONCLUSION

The present study successfully identified hundreds of differentially expressed lncRNAs, miRNAs and mRNAs in RCC by bioinformatics analysis from candidate data from the TCGA. Moreover, we determined the biological processes and pathways via GO and KEGG pathway analyses with cancer-specific mRNAs in RCC. Importantly, we constructed a ceRNA network to explore the potential roles of lncRNAs in RCC, which can serve as a reference for further research. We also investigated the associations between RNAs and overall survival and found that some of the RNAs could be used as biomarkers for RCC diagnosis and prognosis.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Qianwei Xing conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables.
- Yeqing Huang, You Wu and Limin Ma performed the experiments, contributed reagents/materials/analysis tools, prepared figures and/or tables.
- Bo Cai conceived and designed the experiments, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

The following information was supplied regarding data availability:

BO CAI. (2018). TCGA-RCC [Data set]. Zenodo. DOI 10.5281/zenodo.1293051.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.5124#supplemental-information.

REFERENCES

Bai Y, Nie S, Jiang G, Zhou Y, Zhou M, Zhao Y, Li S, Wang F, Lv Q, Huang Y, Yang Q, Li Q, Li Y, Xia Y, Liu Y, Liu J, Qian J, Li B, Wu G, Wu Y, Wang B, Cheng X,

Yang Y, Ke T, Li H, Ren X, Ma X, Liao Y, Xu C, Tu X, Wang QK. 2014. Regulation of CARD8 expression by ANRIL and association of CARD8 single nucleotide polymorphism rs2043211 (p.C10X) with ischemic stroke. *Stroke* **45**:383–388 DOI 10.1161/STROKEAHA.113.003393.

Cao Y, Xu R, Xu X, Zhou Y, Cui L, He X. 2016. Downregulation of lncRNA CASC2 by microRNA-21 increases the proliferation and migration of renal cell carcinoma cells. *Molecular Medicine Reports* **14**:1019–1025 DOI 10.3892/mmr.2016.5337.

Chen J, Chen Y, Gu L, Li X, Gao Y, Lyu X, Chen L, Luo G, Wang L, Xie Y, Duan J, Peng C, Ma X. 2016. LncRNAs act as prognostic and diagnostic biomarkers in renal cell carcinoma: a systematic review and meta-analysis. *Oncotarget* **7**:74325–74336 DOI 10.18632/oncotarget.11101.

Chen S, Ma P, Zhao Y, Li B, Jiang S, Xiong H, Wang Z, Wang H, Jin X, Liu C. 2017. Biological function and mechanism of MALAT-1 in renal cell carcinoma proliferation and apoptosis: role of the MALAT-1-Livin protein interaction. *The Journal of Physiological Sciences* **67**:577–585 DOI 10.1007/s12576-016-0486-8.

Chimge NO, Little GH, Baniwal SK, Adisetiyo H, Xie Y, Zhang T, O'Laughlin A, Liu ZY, Ulrich P, Martin A, Mhawech-Fauceglia P, Ellis MJ, Tripathy D, Groshen S, Liang C, Li Z, Schones DE, Frenkel B. 2016. RUNX1 prevents oestrogen-mediated AXIN1 suppression and β-catenin activation in ER-positive breast cancer. *Nature Communications* **7**:10751 DOI 10.1038/ncomms10751.

Dong LF, Xu SY, Long JP, Wan F, Chen YD. 2017. RNA-sequence analysis reveals differentially expressed genes (degs) in patients exhibiting different risks of tumor metastasis. *Medical Science Monitor* **23**:2842–2849 DOI 10.12659/MSM.904789.

Frew IJ, Moch H. 2015. A clearer view of the molecular complexity of clear cell renal cell carcinoma. *Annual Review of Pathology* **10**:263–289 DOI 10.1146/annurev-pathol-012414-040306.

Gao Y, Ma X, Yao Y, Li H, Fan Y, Zhang Y, Zhao C, Wang L, Ma M, Lei Z, Zhang X. 2016. miR-155 regulates the proliferation and invasion of clear cell renal cell carcinoma cells by targeting E2F2. *Oncotarget* **7**:20324–20337 DOI 10.18632/oncotarget.7951.

Gomes AQ, Nolasco S, Soares H. 2013. Non-coding RNAs: multi-tasking molecules in the cell. *International Journal of Molecular Sciences* **14**:16010–16039 DOI 10.3390/ijms140816010.

Gonda TJ, Ramsay RG. 2016. Adenoid cystic carcinoma can be driven by MYB or MYBL1 rearrangements: new insights into MYB and tumor biology. *Cancer Discovery* **6**:125–127 DOI 10.1158/2159-8290.CD-15-1470.

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**:223–227 DOI 10.1038/nature07672.

He HT, Xu M, Kuang Y, Han XY, Wang MQ, Yang Q. 2016. Biomarker and competing endogenous RNA potential of tumor-specific long noncoding RNA in

chromophobe renal cell carcinoma. *OncoTargets and Therapy* **9**:6399–6406 DOI 10.2147/OTT.S116392.

Hirata H, Hinoda Y, Shahryari V, Deng G, Nakajima K, Tabatabai ZL, Ishii N, Dahiya R. 2015. Long noncoding RNA MALAT1 promotes aggressive renal cell carcinoma through Ezh2 and interacts with miR-205. *Cancer Research* **75**:1322–1331 DOI 10.1158/0008-5472.CAN-14-2931.

Hong Q, Li O, Zheng W, Xiao WZ, Zhang L, Wu D, Cai GY, He JC, Chen XM. 2017. LncRNA HOTAIR regulates HIF-1 α/AXL signaling through inhibition of miR-217 in renal cell carcinoma. *Cell Death & Disease* **8**:e2772 DOI 10.1038/cddis.2017.181.

Huang X, Xiao R, Pan S, Yang X, Yuan W, Tu Z, Xu M, Zhu Y, Yin Q, Wu Y, Hu W, Shao L, Xiong J, Zhang Q. 2017. Uncovering the roles of long non-coding RNAs in cancer stem cells. *Journal of Hematology & Oncology* **10**:62 DOI 10.1186/s13045-017-0428-9.

Jones DT, Hutter B, Jäger N, Korshunov A, Kool M, Warnatz HJ, Zichner T, Lambert SR, Ryzhova M, Quang DA, Fontebasso AM, Stütz AM, Hutter S, Zuckermann M, Sturm D, Gronych J, Lasitschka B, Schmidt S, Seker-Cin H, Witt H, Sultan M, Ralser M, Northcott PA, Hovestadt V, Bender S, Pfaff E, Stark S, Faury D, Schwartzentruber J, Majewski J, Weber UD, Zapatka M, Raeder B, Schlesner M, Worth CL, Bartholomae CC, Von KC, Imbusch CD, Radomski S, Lawerenz C, Van Sluis P, Koster J, Volckmann R, Versteeg R, Lehrach H, Monoranu C, Winkler B, Unterberg A, Herold-Mende C, Milde T, Kulozik AE, Ebinger M, Schuhmann MU, Cho YJ, Pomeroy SL, Von DA, Witt O, Taylor MD, Wolf S, Karajannis MA, Eberhart CG, Scheurlen W, Hasselblatt M, Ligon KL, Kieran MW, Korbel JO, Yaspo ML, Brors B, Felsberg J, Reifenberger G, Collins VP, Jabado N, Eils R, Lichter P, Pfister SM. 2013. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nature Genetics* **45**:927–932 DOI 10.1038/ng.2682.

Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP. 2010. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Science Signaling* **3**:ra8 DOI 10.1126/scisignal.2000568.

Labrousse-Arias D, Martínez-Alonso E, Corral-Escariz M, Bienes-Martínez R, Berridy J, Serrano-Oviedo L, Conde E, García-Bermejo ML, Giménez-Bachs JM, Salinas-Sánchez AS, Sánchez-Prieto R, Yao M, Lasa M, Calzada MJ. 2017. VHL promotes immune response against renal cell carcinoma via NF-κB-dependent regulation of VCAM-1. *The Journal of Cell Biology* **216**:835–847 DOI 10.1083/jcb.201608024.

Lau P, Frigerio CS, De Strooper B. 2014. Variance in the identification of microRNAs deregulated in Alzheimer's disease and possible role of lincRNAs in the pathology: the need of larger datasets. *Ageing Research Reviews* **17**:43–53 DOI 10.1016/j.arr.2014.02.006.

Li X, Dong W, Qu X, Zhao H, Wang S, Hao Y, Li Q, Zhu J, Ye M, Xiao W. 2017. Molecular dysexpression in gastric cancer revealed by integrated analysis of transcriptome data. *Oncology Letters* **13**:3177–3185 DOI 10.3892/ol.2017.5798.

Li CY, Liang GY, Yao WZ, Sui J, Shen X, Zhang YQ, Peng H, Hong WW, Ye YC, Zhang ZY, Zhang WH, Yin LH, Pu YP. 2016a. Integrated analysis of long non-coding RNA

competing interactions reveals the potential role in progression of human gastric cancer. *International Journal of Oncology* **48**:1965–1976 DOI 10.3892/ijo.2016.3407.

Li S, Shuch BM, Gerstein MB. 2017. Whole-genome analysis of papillary kidney cancer finds significant noncoding alterations. *PLOS Genetics* **13**:e1006685 DOI 10.1371/journal.pgen.1006685.

Li Y, Wang T, Li Y, Chen D, Yu Z, Jin L, Ni L, Yang S, Mao X, Gui Y, Lai Y. 2016b. Identification of long-non coding RNA UCA1 as an oncogene in renal cell carcinoma. *Molecular Medicine Reports* **13**:3326–3334 DOI 10.3892/mmr.2016.4894.

Liu D, Yu X, Wang S, Dai E, Jiang L, Wang J, Yang Q, Yang F, Zhou S, Jiang W. 2016a. The gain and loss of long noncoding RNA associated-competing endogenous RNAs in prostate cancer. *Oncotarget* **7**:57228–57238 DOI 10.18632/oncotarget.11128.

Liu Z, Yan HY, Xia SY, Zhang C, Xiu YC. 2016b. Downregulation of long non-coding RNA TRIM52-AS1 functions as a tumor suppressor in renal cell carcinoma. *Molecular Medicine Reports* **13**:3206–3212 DOI 10.3892/mmr.2016.4908.

Lorenzen JM, Thum T. 2016. Long noncoding RNAs in kidney and cardiovascular diseases. *Nature Reviews. Nephrology* **12**:360–373 DOI 10.1038/nrneph.2016.51.

Lv J, Fan HX, Zhao XP, Lv P, Fan JY, Zhang Y, Liu M, Tang H. 2016. Long non-coding RNA Unigene56159 promotes epithelial-mesenchymal transition by acting as a ceRNA of miR-140-5p in hepatocellular carcinoma cells. *Cancer Letters* **382**:166–175 DOI 10.1016/j.canlet.2016.08.029.

Ma X, Shen D, Li H, Zhang Y, Lv X, Huang Q, Gao Y, Li X, Gu L, Xiu S, Bao X, Duan J, Zhang X. 2015. MicroRNA-185 inhibits cell proliferation and induces cell apoptosis by targeting VEGFA directly in von Hippel-Lindau-inactivated clear cell renal cell carcinoma. *Urologic Oncology* **33**:169.e1–11 DOI 10.1016/j.urolonc.2015.01.003.

Martens-Uzunova ES, Böttcher R, Croce CM, Jenster G, Visakorpi T, Calin GA. 2014. Long noncoding RNA in prostate, bladder, and kidney cancer. *European Urology* **65**:1140–1151 DOI 10.1016/j.eururo.2013.12.003.

Qiao HP, Gao WS, Huo JX, Yang ZS. 2013. Long non-coding RNA GAS5 functions as a tumor suppressor in renal cell carcinoma. *Asian Pacific Journal of Cancer Prevention* **14**:1077–1082 DOI 10.7314/APJCP.2013.14.2.1077.

Qin C, Han Z, Qian J, Bao M, Li P, Ju X, Zhang S, Zhang L, Li S, Cao Q, Lu Q, Li J, Shao P, Meng X, Zhang W, Yin C. 2014. Expression pattern of long non-coding RNAs in renal cell carcinoma revealed by microarray. *PLOS ONE* **9**:e99372 DOI 10.1371/journal.pone.0099372.

Qu L, Ding J, Chen C, Wu ZJ, Liu B, Gao Y, Chen W, Liu F, Sun W, Li XF, Wang X, Wang Y, Xu ZY, Gao L, Yang Q, Xu B, Li YM, Fang ZY, Xu ZP, Bao Y, Wu DS, Miao X, Sun HY, Sun YH, Wang HY, Wang LH. 2016. Exosome-transmitted lncarsr promotes sunitinib resistance in renal cancer by acting as a competing endogenous RNA. *Cancer Cell* **29**:653–668 DOI 10.1016/j.ccell.2016.03.004.

Sakai I, Miyake H, Fujisawa M. 2013. Acquired resistance to sunitinib in human renal cell carcinoma cells is mediated by constitutive activation of signal transduction pathways associated with tumour cell proliferation. *BJU International* **112**:E211–E220 DOI 10.1111/j.1464-410X.2012.11655.x.

Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. 2011. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language. *Cell* **146**:353–358 DOI 10.1016/j.cell.2011.07.014.

Sana J, Faltejskova P, Svoboda M, Slaby O. 2012. Novel classes of non-coding RNAs and cancer. *Journal of Translational Medicine* **10**:103 DOI 10.1186/1479-5876-10-103.

Schmidt LS, Linehan WM. 2016. Genetic predisposition to kidney cancer. *Seminars in Oncology* **43**:566–574 DOI 10.1053/j.seminoncol.2016.09.001.

Schmitt AM, Chang HY. 2016. Long noncoding rnas in cancer pathways. *Cancer Cell* **29**:452–463 DOI 10.1016/j.ccell.2016.03.010.

Serghiou S, Kyriakopoulou A, Ioannidis JP. 2016. Long noncoding RNAs as novel predictors of survival in human cancer: a systematic review and meta-analysis. *Molecular Cancer* **15**:50 DOI 10.1186/s12943-016-0535-1.

Siegel RL, Miller KD, Jemal A. 2017. Cancer statistics, 2017. *CA: A Cancer Journal for Clinicians* **67**:7–30 DOI 10.3322/caac.21387.

Su H, Sun T, Wang H, Shi G, Zhang H, Sun F, Ye D. 2017. Decreased TCL6 expression is associated with poor prognosis in patients with clear cell renal cell carcinoma. *Oncotarget* **8**:5789–5799 DOI 10.18632/oncotarget.11011.

Vausort M, Wagner DR, Devaux Y. 2014. Long noncoding RNAs in patients with acute myocardial infarction. *Circulation Research* **115**:668–677 DOI 10.1161/CIRCRESAHA.115.303836.

Wang L, Cai Y, Zhao X, Jia X, Zhang J, Liu J, Zhen H, Wang T, Tang X, Liu Y, Wang J. 2015. Down-regulated long non-coding RNA H19 inhibits carcinogenesis of renal cell carcinoma. *Neoplasma* **62**:412–418 DOI 10.4149/neo_2015_049.

Wang H, Niu L, Jiang S, Zhai J, Wang P, Kong F, Jin X. 2016a. Comprehensive analysis of aberrantly expressed profiles of lncRNAs and miRNAs with associated ceRNA network in muscle-invasive bladder cancer. *Oncotarget* **7**:86174–86185 DOI 10.18632/oncotarget.13363.

Wang Q, Zhang J, Liu Y, Zhang W, Zhou J, Duan R, Pu P, Kang C, Han L. 2016b. A novel cell cycle-associated lncRNA, HOXA11-AS, is transcribed from the 5-prime end of the HOXA transcript and is a biomarker of progression in glioma. *Cancer Letters* **373**:251–259 DOI 10.1016/j.canlet.2016.01.039.

Wettersten HI, Aboud OA, Lara PN, Weiss RH. 2017. Metabolic reprogramming in clear cell renal cell carcinoma. *Nature Reviews. Nephrology* **13**:410–419 DOI 10.1038/nrneph.2017.59.

Xiao H, Tang K, Liu P, Chen K, Hu J, Zeng J, Xiao W, Yu G, Yao W, Zhou H, Li H, Pan Y, Li A, Ye Z, Wang J, Xu H, Huang Q. 2015. LncRNA MALAT1 functions as a competing endogenous RNA to regulate ZEB2 expression by sponging miR-200s in clear cell kidney carcinoma. *Oncotarget* **6**:38005–38015 DOI 10.18632/oncotarget.5357.

Yang S, Ning Q, Zhang G, Sun H, Wang Z, Li Y. 2016. Construction of differential mRNA-lncRNA crosstalk networks based on ceRNA hypothesis uncover key roles of lncRNAs implicated in esophageal squamous cell carcinoma. *Oncotarget* **7**:85728–85740 DOI 10.18632/oncotarget.13828.

Yang KC, Yamada KA, Patel AY, Topkara VK, George I, Cheema FH, Ewald GA, Mann DL, Nerbonne JM. 2014. Deep RNA sequencing reveals dynamic regulation of myocardial noncoding RNAs in failing human heart and remodeling with mechanical circulatory support. *Circulation* **129**:1009–1021 DOI 10.1161/CIRCULATIONAHA.113.003863.

Yue B, Cai D, Liu C, Fang C, Yan D. 2016. Linc00152 functions as a competing endogenous rna to confer oxaliplatin resistance and holds prognostic values in colon cancer. *Molecular Therapy* **24**:2064–2077 DOI 10.1038/mt.2016.180.

Zhang HM, Yang FQ, Chen SJ, Che J, Zheng JH. 2015. Upregulation of long non-coding RNA MALAT1 correlates with tumor progression and poor prognosis in clear cell renal cell carcinoma. *Tumour Biology* **36**:2947–2955 DOI 10.1007/s13277-014-2925-6.

Zubakov D, Stupar Z, Kovacs G. 2006. Differential expression of a new isoform of DLG2 in renal oncocytoma. *BMC Cancer* **6**:106 DOI 10.1186/1471-2407-6-106.