

A three-gene prognostic signature based on circular RNA-associated competing endogenous RNA network for patients with lung adenocarcinoma

Yang Li

Department of Central Laboratory, Xuzhou Central Hospital, Xuzhou Clinical School of Xuzhou Medical University, Xuzhou 221009, China

Rui Li

Department of Central Laboratory, Xuzhou Central Hospital, Xuzhou Clinical School of Xuzhou Medical University, Xuzhou 221009, China

Yonggang Chen

Department of Clinical Pharmacy, Xuzhou Central Hospital, Clinical School of Xuzhou Medical University, Xuzhou 221009, China

He Du (✉ duhe2016@126.com)

Tongji University Affiliated Shanghai Pulmonary Hospital <https://orcid.org/0000-0003-3468-3869>

Rongrong Sun

Department of Medical Oncology, Xuzhou Central Hospital, Clinical School of Xuzhou Medical University, Xuzhou 221009, China

Research

Keywords: lung adenocarcinoma, circular RNAs, competitive endogenous RNA network, signature; prognosis

Posted Date: September 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-65387/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Evidence is increasingly indicating that circular RNAs (circRNAs) are closely involved in tumorigenesis and cancer progression. However, functions of circRNAs in lung adenocarcinoma (LUAD) are still unknown. It is necessary to investigate the regulatory mechanism of circRNAs based on competing endogenous RNA (ceRNA) network in LUAD progression and further construct a prognostic signature for predicting overall survival of LUAD patients.

Methods: Differentially expressed circRNAs (DEcircRNAs), differentially expressed miRNAs (DEmiRNAs) and differentially expressed mRNAs (DEmRNAs) were selected to construct the ceRNA network based on TargetScan prediction tool and Pearson correlation coefficient. Functional and pathway enrichment analysis were performed using DAVID database. A PPI network was constructed and then visualized by Cytoscape software. Finally, we constructed a prognostic signature for LUAD patients using LASSO method and assessed the prognostic performance in the validation cohort.

Results: A total of 38 DEcircRNAs, 56 DEmiRNAs, and 960 DEmRNAs were identified. Based on the interactions predicted by TargetScan, we constructed a circRNA-associated ceRNA network including 11 DEcircRNAs, 8 DEmiRNAs and 49 DEmRNAs. GO and KEGG pathway analysis indicated that the circRNA-associated ceRNA network might be involved in regulation of GTPase activity and endothelial cell differentiation. After removing the discrete points, a PPI network containing 12 DEmRNAs was constructed. Univariate cox regression analysis showed that three DEmRNAs were significantly associated with overall survival. Therefore, we constructed a three-gene prognostic signature for LUAD patients using LASSO method. By applying the signature, patients in the training cohort could be categorized into high-risk or low-risk subgroup with significant survival difference (HR: 1.62, 95% CI: 1.12-2.35, log-rank $p = 0.009$). The prognostic performance was confirmed in an independent GEO cohort (HR: 2.59, 95% CI: 1.32-5.10, log-rank $p = 0.004$). Multivariate cox regression analysis proved that the three-gene signature was an independent prognostic factor for LUAD.

Conclusions: Our findings provided a deeper understanding of the circRNA-associated ceRNA regulatory mechanism in LUAD pathogenesis and constructed a prognostic signature that could be a useful guide for personalized treatment of LUAD patients.

Introduction

Lung cancer is the leading cause of malignancy-related death worldwide. Lung adenocarcinoma (LUAD) accounts for ~ 50% of all types of lung cancer and has been on the rise year after year, especially in women and young adults. Despite recent advances in diagnostic and therapeutic approaches, the 5-year survival rate for LUAD is < 20% [1]. In addition, the underlying molecular mechanisms of LUAD still remain unclear. Therefore, identifying the key molecular mechanisms of LUAD and establishing effective prognostic models for individualized treatment are urgent.

As the boom of sequencing technology and bioinformatics methods, increasing noncoding RNAs (ncRNAs) were identified, one of which is circular RNA (circRNA). CircRNA, featuring stable structure and high tissue-specific expression widely expressed, has a covalently closed loop structure in which the 3' and 5' ends are linked in a non-collinear way by a process termed “back-splicing” [2–4]. In recent years, thousands of circRNAs have been identified in various organisms and found being closely associated with many diseases, particularly with cancer [5–7]. For example, Tian et al. reported that has_circ_0003159 expression was significantly down-regulated in gastric cancer and associated with lymphatic node metastasis and distal metastasis [8]. Gao et al. showed that has_circ_0018289 was up-regulated in cervical cancer and promotes proliferation, migration and invasion of tumor cells [9]. Such evidences strongly support that circRNAs play critical roles in tumor progression. Hence, exploring the regulatory mechanism of circRNAs in tumor is necessary and may be beneficial for tumor therapy.

There is increasing evidence that circRNA contains microRNA response elements (MREs) and inhibits the function of microRNA (miRNA) as miRNA sponges through the competing endogenous RNA (ceRNA) network [10]. The ceRNAs are implicated in many biological processes and the equilibrium of ceRNAs and miRNAs can be critical for promotion of diseases. For example, Sang et al. reported that hsa_circ_0025202 could act as a miRNA sponge for miR-182-5p and further regulate the expression and activity of FOXO3, regulating tamoxifen sensitivity and tumor progression in breast cancer [11]. Fang et al. showed that circRNA-100290 could promote colorectal cancer progression through miR-516b-induced downregulation of FZD4 expression and Wnt/ β -catenin signaling [12]. In addition, Chen et al. found that has_circ_100395 regulates miR-1228/TCF21 pathway to inhibit lung cancer progression [13]. Collectively, these findings show that dysregulation of important circRNAs in the ceRNA network disrupt the miRNA-mediated circRNA/mRNA interactions and therefore contribute to cancer initiation and progression. However, systematic analysis of circRNA-associated ceRNA network in LUAD remains insufficient and requires further exploration.

In this study, based on the public cohorts downloaded from Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA), we aimed to detect the aberrant key circRNAs and explore the potential regulatory mechanism by constructing a circRNA-associated ceRNA network for LUAD patients. Furthermore, using LASSO method, we established an mRNA-based prognostic signature derived from ceRNA network and validated the prognostic performance in an independent cohort. This new approach of constructing prognosis signature based on ceRNA networks can elucidate the circRNA-mediated regulatory mechanisms in LUAD progression and serve as a guide toward more effective individualized treatment decisions for LUAD patients.

Material And Method

Data processing

Two circRNA expression profiles (GSE101586 [14] and GSE101684 [15]) were obtained from the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) database, including 5 and 4 pairs of LUAD and adjacent normal

lung tissues respectively. These raw data were processed by background correction and normalization using the "affy" package of R/Bioconductor. The expression data of mRNA (513 LUAD tissues and 59 normal tissues) and miRNA (513 LUAD tissues and 46 normal tissues) were retrieved from TCGA (<https://portal.gdc.cancer.gov/>). The normalized count values of level 3 gene expression data derived from Illumina HiSeqV2 were extracted as gene expression measurements. Clinical information of 513 LUAD patients was obtained. Forty-seven LUAD patients were excluded because of unknown survival time, age, and tumor stage. Ultimately, 466 patients were retained in our study. An independent cohort (GSE42127 [16]) collected from GEO contained 133 LUAD patients, which was used to test the prognostic ability. The Robust Multi-array Average algorithm was used for preprocessing the raw data. All cohorts and clinical information were described in Table 1 and Supplementary Table S1.

Differentially expressed circRNAs (DEcircRNAs) were identified by student t-test with $p < 0.05$ between LUAD and adjacent normal lung tissues. The common DEcircRNAs overlapped between two cohorts were selected to construct ceRNA network. The edgeR package was used to identify the differentially expressed miRNAs (DEmiRNAs) and differentially expressed mRNAs (DEmRNAs). The cutoff values were set at the $FDR < 0.05$ and $|\log_2FC| > 2$.

Table 1
Cohorts analyzed in this study

	circRNA		miRNA	mRNA	Clinical data	
	Expression data		Expression data	Expression data		
	GSE101586	GSE101684	TCGA-LUAD	TCGA-LUAD	TCGA-LUAD	GSE42127
Normal	5	4	59	46	-	-
Tumor	5	4	513	513	466	133
Platform	Agilent Arraystar CircRNA microarray		Illumina HiSeqV2	Illumina HiSeqV2		

Constructing the circRNA-associated ceRNA Network

A circRNA-associated ceRNA network was constructed based on the interactions between DEcircRNAs, DEmiRNAs and DEmRNAs. TargetScan prediction tool was used to identify interactions between DEcircRNAs with the target miRNAs. To obtain high-quality circRNAs acting as miRNA targets and distinguish those circRNAs acting as miRNA decoys, the circRNAs that had perfect nucleotide pairing between the 2nd and 8th positions of the 5' end of miRNA sequences were selected [17]. The miRNAs within circRNA-miRNA interactions were further screened by DEmiRNAs identified from TCGA. The remained DEmiRNAs were further used to screen the interactions between miRNA and mRNA predicted by Targetscan. The mRNAs targeted by DEmiRNAs were intersected with the DEmRNAs identified from

TCGA. The correlation between DEmiRNAs and DEmRNAs was further calculated by Pearson Correlation Coefficient. Only the interactions with significant negative correlation were retained. Finally, removing the nodes that could not form a circRNA-miRNA-mRNA axis, a circRNA-associated ceRNA network was established and further visualized by Cytoscape software (version 3.7.0; www.cytoscape.org).

Functional Enrichment

In order to explore the biological functions which the circRNA-associated ceRNA network might be involved in, we selected the DEmRNAs within ceRNA network to make the functional enrichment analysis using Database for Annotation, Visualization and Integrated Discovery (DAVID; <http://www.david.abcc.ncifcrf.gov/>). Gene ontology involved three categories: biological processes, molecular function and cellular components. Pathway enrichment was carried out using the Kyoto Encyclopedia of Genes and Genomes (KEGG, <https://www.kegg.jp/>), and it contains information about genomes, chemical substances, biological pathways and diseases.

Construction of the PPI Network

PPI information can be evaluated by an online tool, The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (<https://string-db.org/>). This database has a comprehensive score for each PPI relationship pair that is distributed between 0 and 1; the higher the score, the more reliable the PPI relationship. In this study, we applied STRING to explore potential correlation between DEmRNAs within ceRNA network with a medium confidence criterion (confidence score ≥ 0.4). The Cytoscape software was used to visualize the PPI network.

Establishment of prognostic signature

Based on DEmRNAs within PPI network, univariate cox regression analysis was utilized to test the association between gene expression and overall survival (OS). Genes with significant association were selected to construct the prognostic signature by LASSO method. The model was applied to the expression matrix of candidate genes and the optimal value of the penalty parameter λ was selected to calculate the coefficient of each gene constituting prognostic signature. Using the combination of weighted genes expression values, a risk scoring model was established and the risk scores were calculated as shown in the following equation: Risk score = expression of Gene 1 * β_1 + expression of Gene 2 * β_2 + ... expression of Gene n * β_n . β_i is the regression coefficient of Gene i, which represents the contribution of Gene i to the prognostic risk score. Using the median risk score as the cutoff point, patients in each cohort were divided into low-risk or high-risk group correspondingly.

Statistical analysis

T-test was used to observe the gene expression difference between metastasis samples and non-metastasis patients. The multivariate Cox proportional-hazards regression model was used to evaluate independent association between prognostic signature and patient survival after adjusting for stage, age and gender. Hazard ratios (HRs) and 95% confidence intervals (CIs) were computed based on the Cox regression analysis. Survival curves were estimated using the Kaplan–Meier method and were compared using the log-rank test. Fisher's exact test was used to observe the differences in mortality rate and lymph

node metastasis rate between different risk groups. The significance was defined as a P value of < 0.05 . All statistical analysis were performed using the R3.4.0.

Results

Identification of differentially expressed RNAs

The flowchart for this study is shown in Fig. 1. We initially performed differential expression analysis by comparing circRNA expression between LUAD and adjacent normal lung tissues in the two cohorts from GEO. With cut-off criteria of $p < 0.05$, a total of 38 DEcircRNAs (including 31 up-regulated circRNAs and 7 down-regulated circRNAs) were concordance between GSE101586 and GSE101684 (Fig. 2A). The expression data of miRNAs and mRNAs download from TCGA were analyzed using the “edgeR” package in R. With cut-off criteria of $FDR < 0.05$ and $|\log_2FC| > 2.0$, 56 differentially expressed miRNAs (including 37 up-regulated miRNAs and 19 down-regulated miRNAs) and 960 differentially expressed mRNAs (including 653 up-regulated mRNAs and 307 down-regulated mRNAs) were identified respectively (Fig. 2B, 2C).

Construction of the circRNA-associated ceRNA Network

Using the TargetScan prediction tool to identify DEcircRNAs with the target miRNAs, we identified 494 circRNA-miRNA pairs based on 38 DEcircRNA and 221 miRNAs. After screened by the DEmiRNAs identified from TCGA-LUAD cohort, only 39 circRNA-miRNA pairs remained, including 18 DEcircRNAs and 18 DEmiRNAs. We further identified mRNAs targeted by these DEmiRNAs from TargetScan database and selected those that were overlapped with the DEmRNAs identified from TCGA-LUAD cohort. A total of 305 miRNA-mRNA pairs were identified, including 11 DEmiRNAs and 191 DEmRNAs. After calculating the expression correlation between miRNA and mRNA by Pearson correlation coefficient, only 53 miRNA-mRNA pairs with significant correlation were retained. Finally, combining the circRNA-miRNA pairs with miRNA-mRNA pairs, a circRNA-miRNA-mRNA network was constructed, including 11 DEcircRNAs, 8 DEmiRNAs and 49 DEmRNAs, and subsequently visualized using Cytoscape (Fig. 3).

Functional enrichment of DEmRNAs

To further investigate the biological processes which the circRNA-associated ceRNA network might be involved in, we performed GO annotation and KEGG pathway enrichment using DAVID database. Functional enrichment results were shown in Fig. 4. For GO annotation, ten GO terms were significantly enriched. For example, positive regulation of GTPase activity, as the most significantly enriched GO term, had been reported in multiple articles related to the progression and metastasis of cancer [18–20]. Besides, endothelial cell differentiation and regulation of transcription from RNA polymerase II promote, were also related to cancer progression and chemoresistance validated in many researches [21–24]. KEGG pathway enrichment analysis indicated that two pathways, including pathways in cancer and focal adhesion, were significantly enriched. Focal adhesion, regulating cell proliferation, survival, migration, and invasion, plays an essential role in tumour invasiveness and metastasis [25, 26]. Such results showed

that DEmRNAs within network played key roles in multiple cancer-related pathways, and further indicated that the circRNA-associated ceRNA network might be involved in LUAD carcinogenesis and progression.

Construction of the PPI network and evaluation of relationship with clinical characteristics

To explore the interactions between the 49 DEmRNAs, we constructed a PPI network using the STRING database. After removing unconnected nodes, the PPI network comprised 12 nodes and 27 edges (Fig. 5A). Furthermore, we investigated the correlation between the gene expression and clinical variables including lymphatic node metastasis, distant metastasis and overall survival. Univariate cox regression analysis results showed that the high expression of three genes (including CEP55, KIF14 and PRR11) predicted poor prognosis in LUAD patients (Table S2). Notably, we found that CEP55 and PRR11 were significantly up-regulated in patients with lymphatic node metastasis (Fig. 5B, 5D). The expression of KIF14 also had an upward trend in lymphatic node metastasis patients (Fig. 5C). Although no statistical significance was observed, the average expression levels of all three genes increased in distant metastasis patients (Fig. 5E-5G). These results indicated that these three genes might play key roles in the progression and prognosis of LUAD.

Construction and validation of the prognostic signature

Based on the importance of three genes in the progression and prognosis of LUAD, we tried to construct a prognostic signature for LUAD patients using LASSO method. Combining the regression coefficients with gene expression values, the risk-score formula was created as follows: Risk score = (0.058*expression level of CEP55) + (0.065*expression level of KIF14) + (0.202*expression level of PRR11). We then calculated the risk score for each patient and ranked them based on increasing score, after which patients were classified into a high-risk (n = 233) or a low-risk (n = 233) group based on the median risk score. We observed the overall survival between two risk groups with significantly different survival rate (log-rank p = 0.009; Fig. 6A). Patients with high risk score had significantly shorter OS than patients with low risk score. Mortality rate was 30.5% (71/233) in the high-risk group, as compared to 20.6% (48/233) in the low-risk group (p = 0.021, Fisher exact test, Fig. 6B). The rate of lymphatic node metastasis in high-risk group was significantly higher than that in low-risk group (p = 0.005, Fisher exact test, Fig. 6C). The risk score distribution, survival status, and expression profile of the three prognostic genes are shown in Fig. 6D. Taking into the patients' clinical features, including age, gender and stage, the multivariate Cox regression analysis showed that the three-gene signature also had statistical significance as an independent prognostic factor in the training cohort (HR: 1.45, 95% CI: 1.01–2.13, P = 0.050) (Table 2).

We test the prognostic performance of the three-gene signature in an independent cohort. Similar to the training cohort findings, patients in high-risk group had a shorter survival time than patients in low-risk group (log-rank p = 0.004, Fig. 7A). The risk score distribution, survival status, and expression profile of the three prognostic genes are shown in Fig. 7B and 7C. Multivariate Cox regression analysis showed that, after adjusting for age, gender and stage, the prognostic signature remained significantly associated with patient OS (HR: 2.45, 95% CI: 1.24–4.84, P = 0.010) (Table 2).

Table 2
Multivariate analysis of prognostic factors by Cox proportional hazard model

Variables	TCGA-LUAD		GSE42127	
	HR (95%CI)	P value	HR (95%CI)	P value
Age	1.28 (0.84–1.96)	0.25	2.03 (0.89–4.63)	0.09
Prognostic signature				
High-risk vs. low-risk	1.45 (1.01–2.13)	0.05	2.45 (1.24–4.84)	0.01
Gender				
Male vs. Female	0.80 (0.55–1.16)	0.24	1.57 (0.82–3.03)	0.18
TNM stage				
III + IV vs. I + II	2.65 (1.81–3.90)	< 0.001	1.33 (0.61–2.89)	0.45

Discussion

LUAD is a major lung cancer that is in a locally advanced or metastatic stage at the time of diagnosis, which leaves no time for early detection or treatment [27]. The accurate diagnosis and prognosis may warrant timely treatment to potentially decrease the mortality. Therefore, it is essential to explore the molecular mechanisms of LUAD progression and identify the effective biomarkers contributing to better treatment and better overall prognosis for LUAD patients.

Growing experimental evidences have shown that circRNAs play important roles in many complicated human diseases, including malignant tumors [28–30]. Lately, lots of studies indicated that circRNAs could function as a tumor regulator in lung cancer [31–33]. As a type of high efficiency ceRNA, it could inhibit the binding of miRNAs to target genes and regulate the expression level of target genes by exerting a miRNA sequestering effect [34]. However, the expression pattern and biological function of circRNAs in LUAD remain largely elusive. In present study, we constructed a circRNA-miRNA-mRNA ceRNA network to explore the regulatory mechanism of circRNAs involved in tumor progression. An integrated circRNA-associated ceRNA network, including 11 circRNAs, 8 miRNAs and 49 mRNAs, was established. The 11

circRNAs identified in the ceRNA network were hsa_circ_0002191, hsa_circ_0002727, hsa_circ_0049271, hsa_circ_0050395, hsa_circ_0001974, hsa_circ_0004006, hsa_circ_0000641, hsa_circ_0079929, hsa_circ_0007788, hsa_circ_0001936, hsa_circ_0015278. We found that two of eleven circRNAs had been reported to be related to disease progression and pathogenesis. Peng et al. reported that hsa_circ_0015278 was significantly down-regulated in papillary thyroid carcinoma, showing interactive potential with two cancer-related miRNAs [35]. Furthermore, several promising cancer-related genes that may be targets of the dysregulated hsa_circRNA_0015278/miR-141-3p/miR-200a-3p axis were identified to explore the pathogenesis of papillary thyroid carcinoma. Another circRNA, named hsa_circ_0079929, has been reported by Zou et al [36]. They found that elevated expression of hsa_circRNA_0079929 might inhibit the expression of hsa-miR-26a-3p to increase aortic smooth muscle cells phenotype or apoptosis in thoracic aortic dissection. Although other circRNAs have not been reported previously, we could detect the interactive potential with miRNAs and target genes based on the constructed ceRNA network. For example, in circRNA-associated ceRNA network constructed in present study, hsa_circ_0001936 interacted with has-miR-142-5p and further might regulate its target genes, such as CAV2, FGD5 and S1PR1, which were involved in positive regulation of GTPase activity process. Such results indicated that hsa_circ_0001936 might play key roles in LUAD progression by regulating GTPase activity. Thus, the regulatory mechanism of these circRNAs speculated in ceRNA network requires further experimental investigation in future studies.

Limited by small sample size or lack of corresponding clinical data, it is not feasible to construct diagnostic or prognostic signatures for clinical application based on circRNAs at the current stage. However, extensive transcriptome data with corresponding clinical data provide us the possibility to construct clinically available diagnostic or prognostic signatures. In present study, we identified the key mRNA participating in the circRNA-associated ceRNA network based on STRING database and further constructed a three-gene prognostic signature for LUAD patients. The three-gene prognostic signature is composed of centrosomal protein 55 (CEP55), kinesin family member 14 (KIF14) and proline rich 11 (PRR11). CEP55 localizes to the centrosome of interphase cells and to the midbody during cytokinesis [37]. Many studies have demonstrated that CEP55 was highly expressed in a variety of cancers and could be used as a diagnostic and prognostic marker for several cancers [38–40]. For LUAD, Fu et al. had reported that CEP55 was significantly up-regulated in LUAD patients and could be an independent prognostic factor [41]. As a mitotic kinesin, KIF14 has been reported to serve oncogenic roles through the regulation of the cell cycle, DNA replication and DNA repair biological processes in a variety of malignancies, such as colorectal cancer [42], ovarian cancer [43], gastric cancer [44], prostate cancer [45], and so on. Zhang et al. had confirmed that KIF14 was notably up-regulated in tumor tissues of LUAD and the expression levels of the KIF14 exhibited a strongly correlation with OS [46]. PRR11, located on chromosome 17q22, has been reported to be closely associated with cell cycle progression and was also demonstrated to participate in various biological processes in tumor cells, including cell invasion, migration and proliferation, by acting as an oncogene [47–50]. Ji et al. found that PRR11 was periodically expressed in a cell cycle-dependent manner. RNAi-mediated silencing of PRR11 caused S phase arrest and suppressed cellular proliferation, colony formation ability in lung cancer cells,

demonstrating that PRR11 had a critical role in both cell cycle progression and tumorigenesis [51]. All three genes are associated with tumor progression and prognosis, indicating that the three-gene signature could be served as clinically available prognostic signature to provide potential novel targets and promote individualized treatment. In addition, we found that all three genes were regulated by has-miR-144-3p interacting with hsa_circ_0002191 and hsa_circ_0002727, indicating that hsa_circ_0002191 and hsa_circ_0002727 might serve as potential regulators to affect gene expression and further influenced the progression and prognosis of LUAD patients. Thus, further experimental investigation needs to be implemented to explore the regulatory mechanism and prognostic ability of hsa_circ_0002191 and hsa_circ_0002727.

Conclusion

In summary, the present study constructed and analyzed a circRNA-associated ceRNA network via comprehensive bioinformatics analysis, which may provide meaningful evidences to future studies focused on the molecular mechanisms of LUAD. Furthermore, a three-gene signature derived from circRNA-associated ceRNA network was constructed and could be applied for clinical prognostic evaluation, which provide a useful guide for personalized treatment of LUAD patients.

Abbreviations

LUAD: Lung adenocarcinoma; GEO: Gene Expression Omnibus; TCGA: The Cancer Genome Atlas; GO: gene ontology; KEGG: Kyoto Encyclopaedia of Genes and Genomes; circRNA: circular RNA; ceRNA: competing endogenous RNA; LASSO: The least absolute shrinkage and selection operator; OS: overall survival; CI: Confidence interval; HR: Hazard ratios.

Declarations

Acknowledgments

None

Authors' contributions

YL, HD and RL participated in the methodology, statistical analysis and drafted the manuscript. RRS edited the manuscript. YL and YGC conceived and designed the experiments. All authors read and approved the final manuscript.

Funding

This work was supported by the Natural Science Foundation of Jiangsu Province (grant number 20180179).

Availability of data and materials

The datasets generated and analyzed during the present study are available in the TCGA (<https://portal.gdc.cancer.gov/>) and the GEO (<https://www.ncbi.nlm.nih.gov/geo/>).

Consent for publication

All the authors agree to the publication clause.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2019. *CA Cancer J Clin* 2019, 69(1):7-34.
2. Jeck WR, Sharpless NE: Detecting and characterizing circular RNAs. *Nat Biotechnol* 2014, 32(5):453-461.
3. Qu S, Yang X, Li X, Wang J, Gao Y, Shang R, Sun W, Dou K, Li H: Circular RNA: A new star of noncoding RNAs. *Cancer Lett* 2015, 365(2):141-148.
4. Petkovic S, Muller S: RNA circularization strategies in vivo and in vitro. *Nucleic Acids Res* 2015, 43(4):2454-2465.
5. Yao JT, Zhao SH, Liu QP, Lv MQ, Zhou DX, Liao ZJ, Nan KJ: Over-expression of CircRNA_100876 in non-small cell lung cancer and its prognostic value. *Pathol Res Pract* 2017, 213(5):453-456.
6. Zhang HD, Jiang LH, Sun DW, Hou JC, Ji ZL: CircRNA: a novel type of biomarker for cancer. *Breast Cancer* 2018, 25(1):1-7.
7. Zhu M, Wang H, Cui J, Li W, An G, Pan Y, Zhang Q, Xing R, Lu Y: Calcium-binding protein S100A14 induces differentiation and suppresses metastasis in gastric cancer. *Cell Death Dis* 2017, 8(7):e2938.
8. Tian M, Chen R, Li T, Xiao B: Reduced expression of circRNA hsa_circ_0003159 in gastric cancer and its clinical significance. *J Clin Lab Anal* 2018, 32(3).
9. Gao YL, Zhang MY, Xu B, Han LJ, Lan SF, Chen J, Dong YJ, Cao LL: Circular RNA expression profiles reveal that hsa_circ_0018289 is up-regulated in cervical cancer and promotes the tumorigenesis. *Oncotarget* 2017, 8(49):86625-86633.
10. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J: Natural RNA circles function as efficient microRNA sponges. *Nature* 2013, 495(7441):384-388.
11. Sang Y, Chen B, Song X, Li Y, Liang Y, Han D, Zhang N, Zhang H, Liu Y, Chen T et al: circRNA_0025202 Regulates Tamoxifen Sensitivity and Tumor Progression via Regulating the miR-182-5p/FOXO3a Axis in Breast Cancer. *Mol Ther* 2019, 27(9):1638-1652.

12. Fang G, Ye BL, Hu BR, Ruan XJ, Shi YX: CircRNA_100290 promotes colorectal cancer progression through miR-516b-induced downregulation of FZD4 expression and Wnt/beta-catenin signaling. *Biochem Biophys Res Commun* 2018, 504(1):184-189.
13. Chen D, Ma W, Ke Z, Xie F: CircRNA hsa_circ_100395 regulates miR-1228/TCF21 pathway to inhibit lung cancer progression. *Cell Cycle* 2018, 17(16):2080-2090.
14. Chen T, Yang Z, Liu C, Wang L, Yang J, Chen L, Li W: Circ_0078767 suppresses non-small-cell lung cancer by protecting RASSF1A expression via sponging miR-330-3p. *Cell Prolif* 2019, 52(2):e12548.
15. Zhao J, Li L, Wang Q, Han H, Zhan Q, Xu M: CircRNA Expression Profile in Early-Stage Lung Adenocarcinoma Patients. *Cell Physiol Biochem* 2017, 44(6):2138-2146.
16. Tang H, Xiao G, Behrens C, Schiller J, Allen J, Chow CW, Suraokar M, Corvalan A, Mao J, White MA et al: A 12-gene set predicts survival benefits from adjuvant chemotherapy in non-small cell lung cancer patients. *Clin Cancer Res* 2013, 19(6):1577-1586.
17. Yi Y, Liu Y, Wu W, Wu K, Zhang W: Reconstruction and analysis of circRNAmiRNAmRNA network in the pathology of cervical cancer. *Oncol Rep* 2019, 41(4):2209-2225.
18. Deryugina EI, Quigley JP: Matrix metalloproteinases and tumor metastasis. *Cancer Metastasis Rev* 2006, 25(1):9-34.
19. Haga RB, Ridley AJ: Rho GTPases: Regulation and roles in cancer cell biology. *Small GTPases* 2016, 7(4):207-221.
20. Huh YH, Oh S, Yeo YR, Chae IH, Kim SH, Lee JS, Yun SJ, Choi KY, Ryu JH, Jun CD et al: Swiprosin-1 stimulates cancer invasion and metastasis by increasing the Rho family of GTPase signaling. *Oncotarget* 2015, 6(15):13060-13071.
21. Dellino GI, Palluzzi F, Chiariello AM, Piccioni R, Bianco S, Furia L, De Conti G, Bouwman BAM, Melloni G, Guido D et al: Release of paused RNA polymerase II at specific loci favors DNA double-strand-break formation and promotes cancer translocations. *Nat Genet* 2019, 51(6):1011-1023.
22. He XD, Hu SL, Shen ZJ, Tao SN, Dong L, Zhu YY, Li M: [Chiral selectivity in differentiation of lung cancer A549 cell to vascular endothelial cells after drug resistance induced by D-or L-methotrexate enantiomers]. *Zhonghua Yi Xue Za Zhi* 2009, 89(10):690-694.
23. Meng H, Bartholomew B: Emerging roles of transcriptional enhancers in chromatin looping and promoter-proximal pausing of RNA polymerase II. *J Biol Chem* 2018, 293(36):13786-13794.
24. Zhou X, Zhou R, Zhou H, Li Q, Hong J, Meng R, Zhu F, Zhang S, Dai X, Peng G et al: ETS-1 Induces Endothelial-Like Differentiation and Promotes Metastasis in Non-Small Cell Lung Cancer. *Cell Physiol Biochem* 2018, 45(5):1827-1839.
25. Aboubakar Nana F, Lecocq M, Ladjemi MZ, Detry B, Dupasquier S, Feron O, Massion PP, Sibille Y, Pilette C, Ocak S: Therapeutic Potential of Focal Adhesion Kinase Inhibition in Small Cell Lung Cancer. *Mol Cancer Ther* 2019, 18(1):17-27.
26. Zhou B, Wang GZ, Wen ZS, Zhou YC, Huang YC, Chen Y, Zhou GB: Somatic Mutations and Splicing Variants of Focal Adhesion Kinase in Non-Small Cell Lung Cancer. *J Natl Cancer Inst* 2018, 110(2).

27. Morgensztern D, Ng SH, Gao F, Govindan R: Trends in stage distribution for patients with non-small cell lung cancer: a National Cancer Database survey. *J Thorac Oncol* 2010, 5(1):29-33.
28. Bi W, Huang J, Nie C, Liu B, He G, Han J, Pang R, Ding Z, Xu J, Zhang J: CircRNA circRNA_102171 promotes papillary thyroid cancer progression through modulating CTNNBIP1-dependent activation of beta-catenin pathway. *J Exp Clin Cancer Res* 2018, 37(1):275.
29. Li P, Yang X, Yuan W, Yang C, Zhang X, Han J, Wang J, Deng X, Yang H, Li P et al: CircRNA-Cdr1as Exerts Anti-Oncogenic Functions in Bladder Cancer by Sponging MicroRNA-135a. *Cell Physiol Biochem* 2018, 46(4):1606-1616.
30. Tang W, Fu K, Sun H, Rong D, Wang H, Cao H: CircRNA microarray profiling identifies a novel circulating biomarker for detection of gastric cancer. *Mol Cancer* 2018, 17(1):137.
31. Qu D, Yan B, Xin R, Ma T: A novel circular RNA hsa_circ_0020123 exerts oncogenic properties through suppression of miR-144 in non-small cell lung cancer. *Am J Cancer Res* 2018, 8(8):1387-1402.
32. Yang L, Wang J, Fan Y, Yu K, Jiao B, Su X: Hsa_circ_0046264 up-regulated BRCA2 to suppress lung cancer through targeting hsa-miR-1245. *Respir Res* 2018, 19(1):115.
33. Zhu X, Wang X, Wei S, Chen Y, Chen Y, Fan X, Han S, Wu G: hsa_circ_0013958: a circular RNA and potential novel biomarker for lung adenocarcinoma. *FEBS J* 2017, 284(14):2170-2182.
34. Lasda E, Parker R: Circular RNAs: diversity of form and function. *RNA* 2014, 20(12):1829-1842.
35. Peng N, Shi L, Zhang Q, Hu Y, Wang N, Ye H: Microarray profiling of circular RNAs in human papillary thyroid carcinoma. *PLoS One* 2017, 12(3):e0170287.
36. Zou M, Huang C, Li X, He X, Chen Y, Liao W, Liao Y, Sun J, Liu Z, Zhong L et al: Circular RNA expression profile and potential function of hsa_circRNA_101238 in human thoracic aortic dissection. *Oncotarget* 2017, 8(47):81825-81837.
37. Fabbro M, Zhou BB, Takahashi M, Sarcevic B, Lal P, Graham ME, Gabrielli BG, Robinson PJ, Nigg EA, Ono Y et al: Cdk1/Erk2- and Plk1-dependent phosphorylation of a centrosome protein, Cep55, is required for its recruitment to midbody and cytokinesis. *Dev Cell* 2005, 9(4):477-488.
38. Inoda S, Hirohashi Y, Torigoe T, Nakatsugawa M, Kiriyaama K, Nakazawa E, Harada K, Takasu H, Tamura Y, Kamiguchi K et al: Cep55/c10orf3, a tumor antigen derived from a centrosome residing protein in breast carcinoma. *J Immunother* 2009, 32(5):474-485.
39. Jones J, Otu H, Spentzos D, Kolia S, Inan M, Beecken WD, Fellbaum C, Gu X, Joseph M, Pantuck AJ et al: Gene signatures of progression and metastasis in renal cell cancer. *Clin Cancer Res* 2005, 11(16):5730-5739.
40. Shiraishi T, Terada N, Zeng Y, Suyama T, Luo J, Trock B, Kulkarni P, Getzenberg RH: Cancer/Testis Antigens as potential predictors of biochemical recurrence of prostate cancer following radical prostatectomy. *J Transl Med* 2011, 9:153.
41. Fu L, Wang H, Wei D, Wang B, Zhang C, Zhu T, Ma Z, Li Z, Wu Y, Yu G: The value of CEP55 gene as a diagnostic biomarker and independent prognostic factor in LUAD and LUSC. *PLoS One* 2020, 15(5):e0233283.

42. Wang ZZ, Yang J, Jiang BH, Di JB, Gao P, Peng L, Su XQ: KIF14 promotes cell proliferation via activation of Akt and is directly targeted by miR-200c in colorectal cancer. *Int J Oncol* 2018, 53(5):1939-1952.
43. Qiu HL, Deng SZ, Li C, Tian ZN, Song XQ, Yao GD, Geng JS: High expression of KIF14 is associated with poor prognosis in patients with epithelial ovarian cancer. *Eur Rev Med Pharmacol Sci* 2017, 21(2):239-245.
44. Yang Z, Li C, Yan C, Li J, Yan M, Liu B, Zhu Z, Wu Y, Gu Q: KIF14 promotes tumor progression and metastasis and is an independent predictor of poor prognosis in human gastric cancer. *Biochim Biophys Acta Mol Basis Dis* 2019, 1865(1):181-192.
45. Zhang Y, Yuan Y, Liang P, Zhang Z, Guo X, Xia L, Zhao Y, Shu XS, Sun S, Ying Y et al: Overexpression of a novel candidate oncogene KIF14 correlates with tumor progression and poor prognosis in prostate cancer. *Oncotarget* 2017, 8(28):45459-45469.
46. Zhang L, Zhu G, Wang X, Liao X, Huang R, Huang C, Huang P, Zhang J, Wang P: Genomewide investigation of the clinical significance and prospective molecular mechanisms of kinesin family member genes in patients with lung adenocarcinoma. *Oncol Rep* 2019, 42(3):1017-1034.
47. Song Z, Liu W, Xiao Y, Zhang M, Luo Y, Yuan W, Xu Y, Yu G, Hu Y: PRR11 Is a Prognostic Marker and Potential Oncogene in Patients with Gastric Cancer. *PLoS One* 2015, 10(8):e0128943.
48. Tan S, Jiang Z, Hou A, Wang J, Zhang J, Dai L: Expression of PRR11 protein and its correlation with pancreatic cancer and effect on survival. *Oncol Lett* 2017, 13(6):4117-4122.
49. Wang Y, Zhang C, Mai L, Niu Y, Wang Y, Bu Y: PRR11 and SKA2 gene pair is overexpressed and regulated by p53 in breast cancer. *BMB Rep* 2019, 52(2):157-162.
50. Zhang C, Zhang Y, Li Y, Zhu H, Wang Y, Cai W, Zhu J, Ozaki T, Bu Y: PRR11 regulates late-S to G2/M phase progression and induces premature chromatin condensation (PCC). *Biochem Biophys Res Commun* 2015, 458(3):501-508.
51. Ji Y, Xie M, Lan H, Zhang Y, Long Y, Weng H, Li D, Cai W, Zhu H, Niu Y et al: PRR11 is a novel gene implicated in cell cycle progression and lung cancer. *Int J Biochem Cell Biol* 2013, 45(3):645-656.

Figures

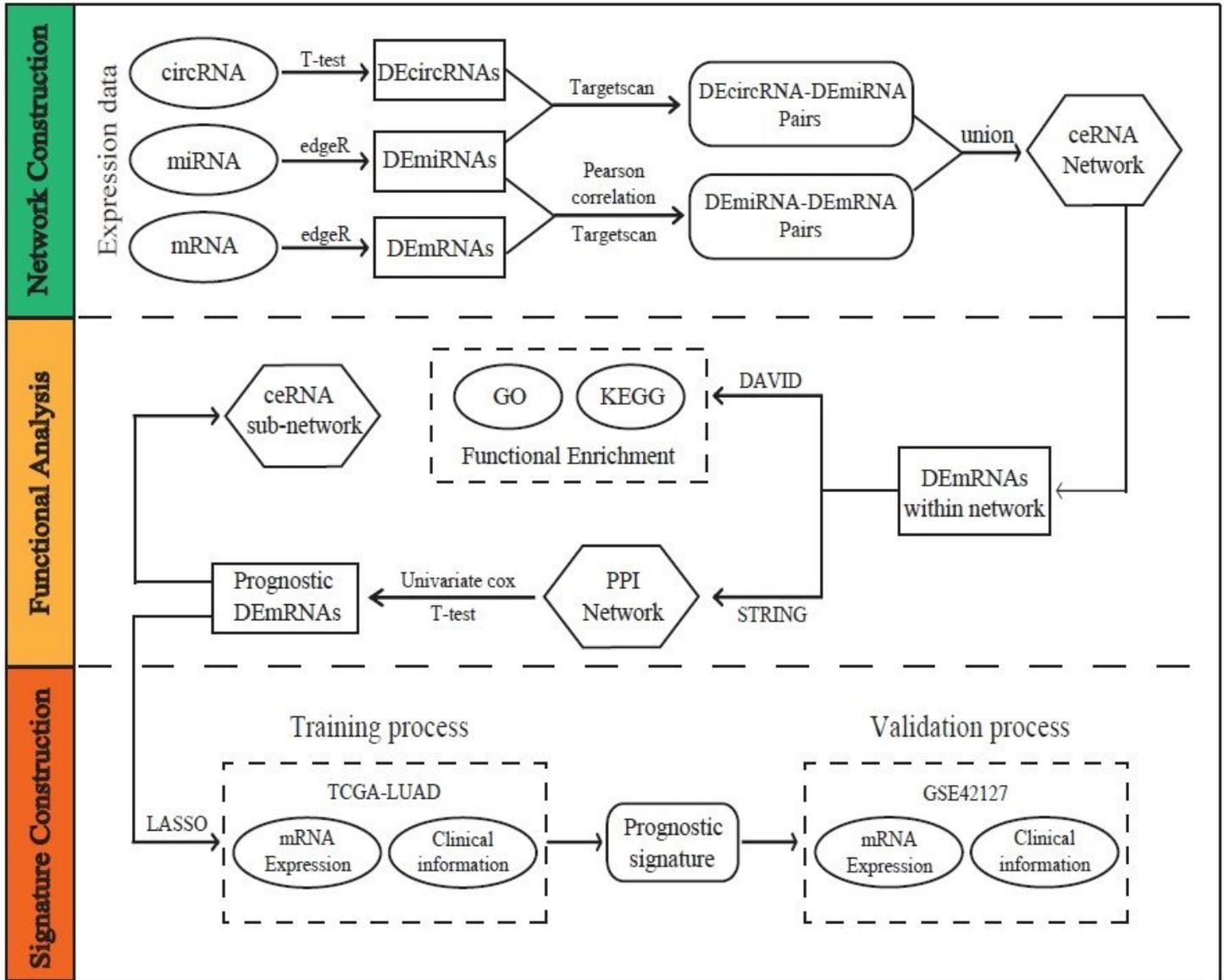


Figure 1

The flowchart for this study.

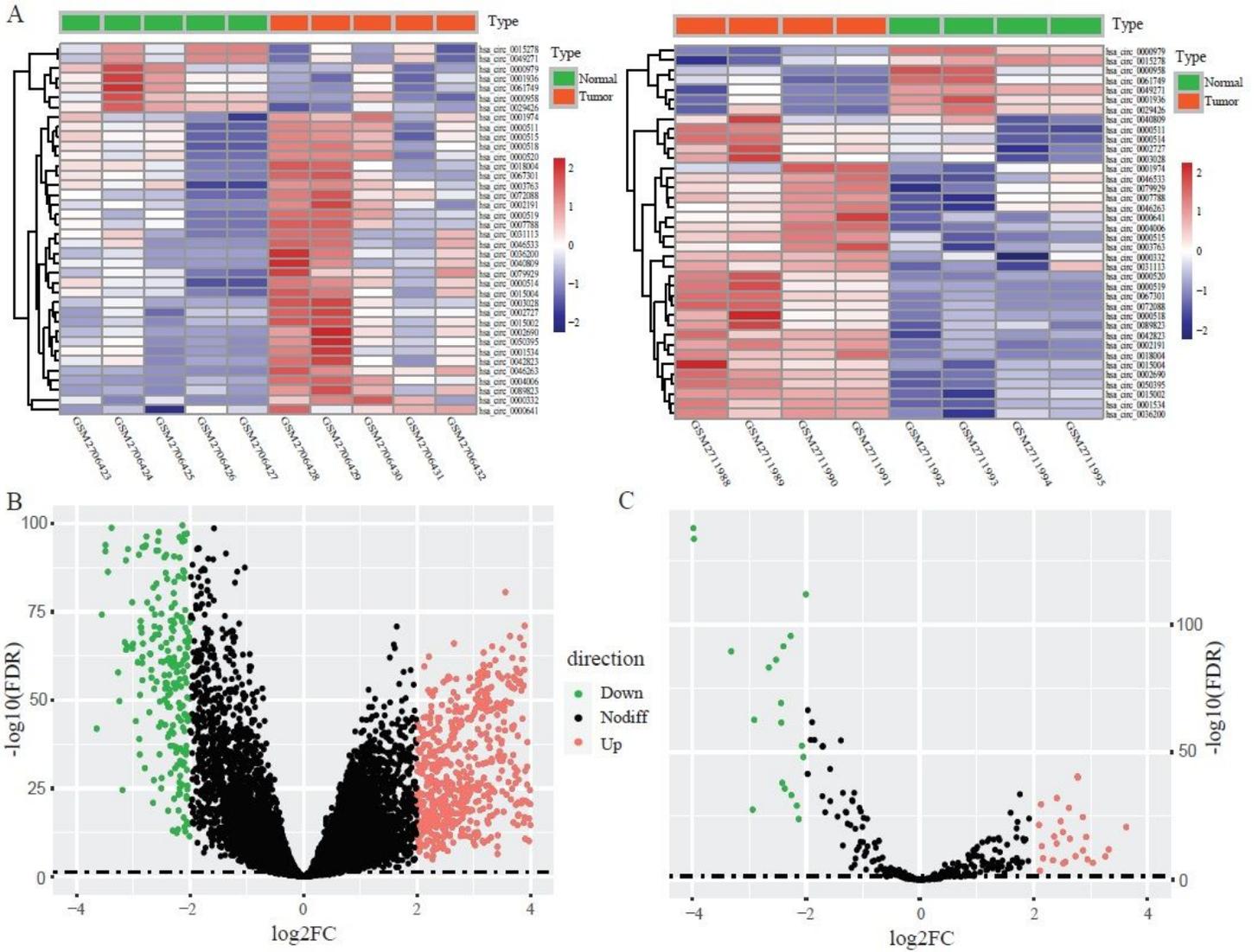


Figure 2

Heat maps of the 38 differentially expressed circular RNAs (A) and volcano plots of differentially expressed mRNA (B) and differentially expressed miRNA (C). Red and green dots represent significantly up-regulated and down-regulated RNAs, respectively ($FDR < 0.05$ and $|\log_2FC| > 2.0$).

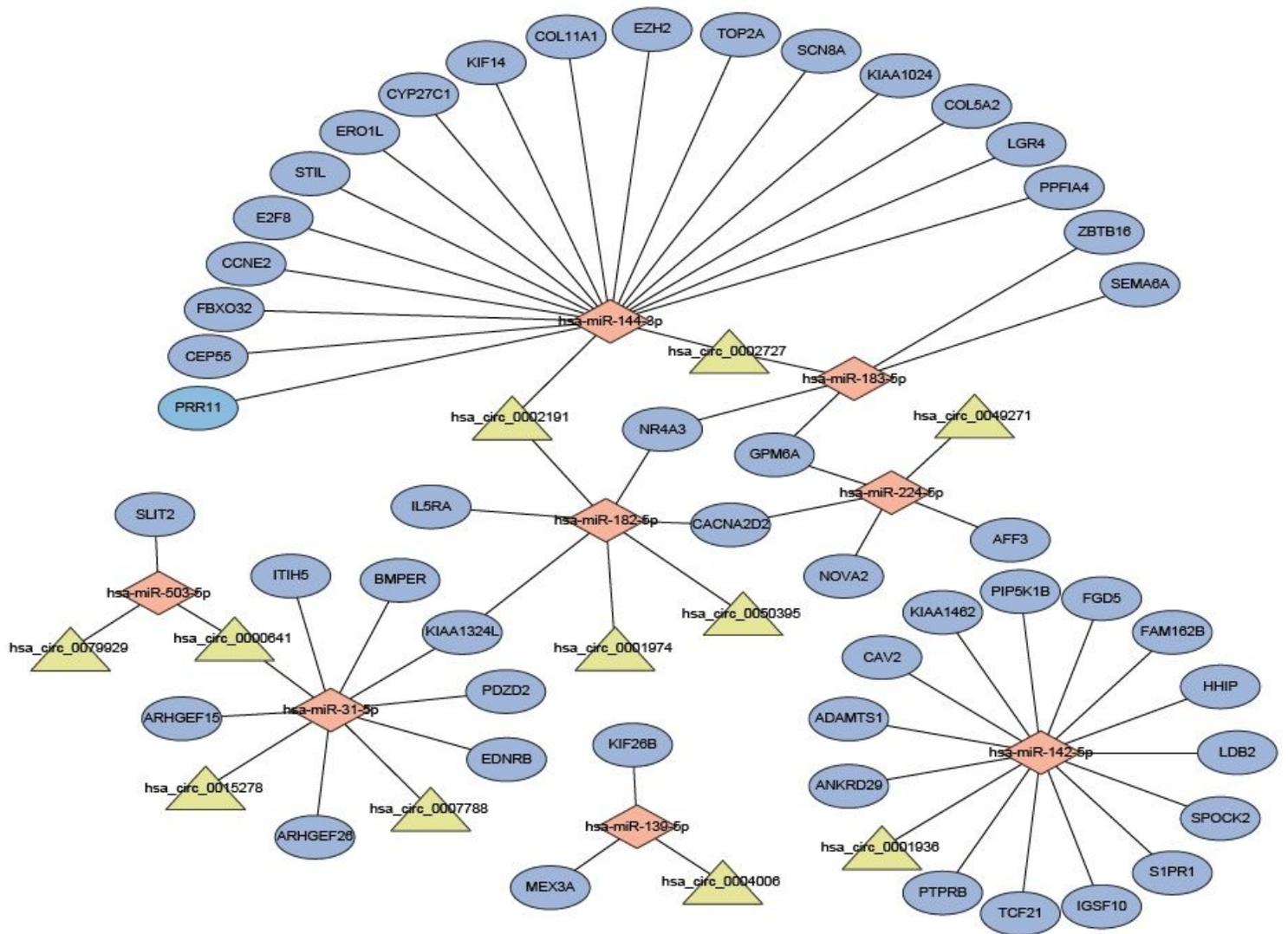


Figure 3

The circRNA-associated competing endogenous network in lung adenocarcinoma. Triangles represent circRNAs, diamonds represent miRNAs, ellipses represent mRNAs, and black lines represent circRNA-miRNA-mRNA interactions.

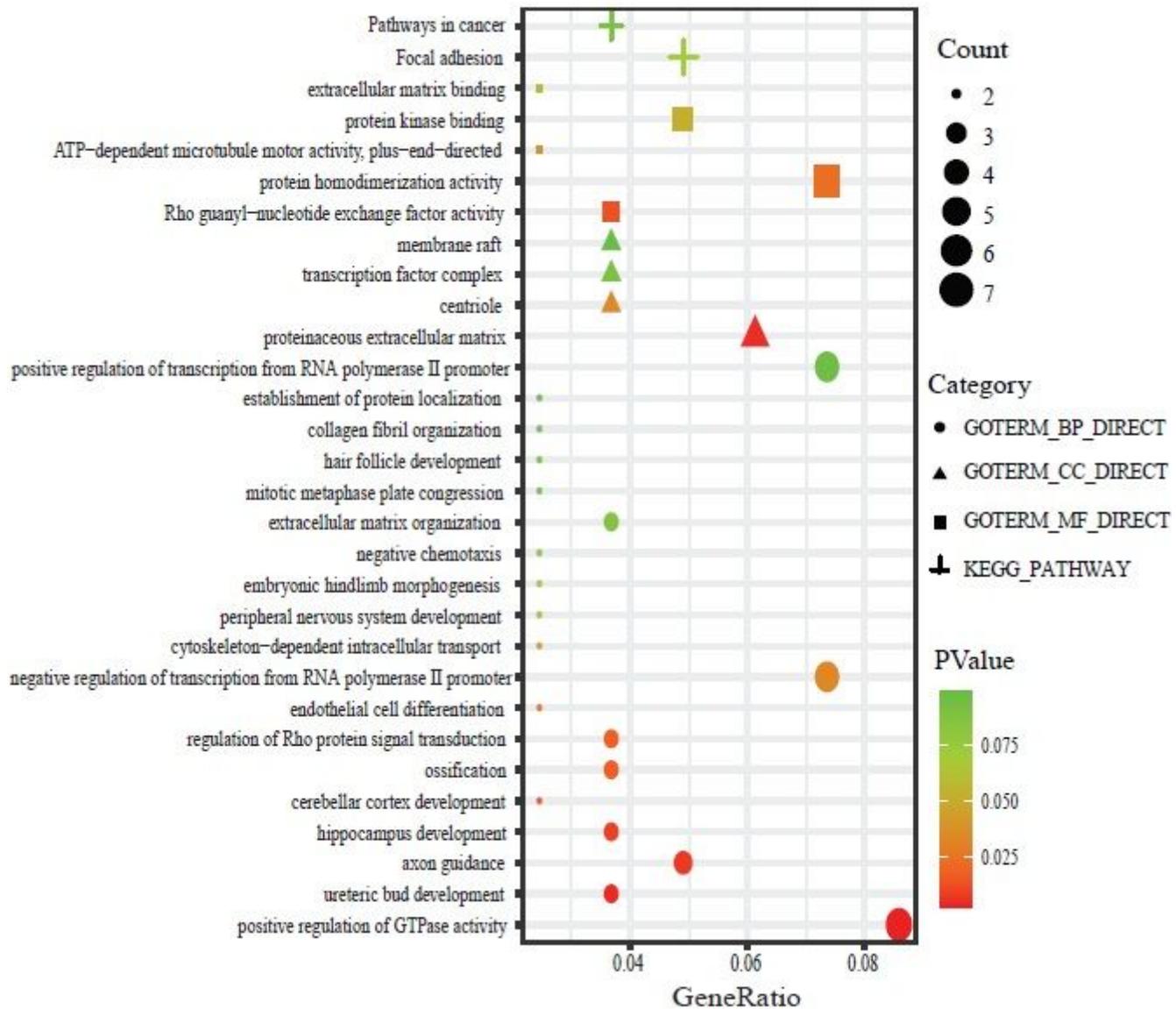


Figure 4

Enrichment of Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of differentially expressed mRNAs in the competing endogenous RNA network.

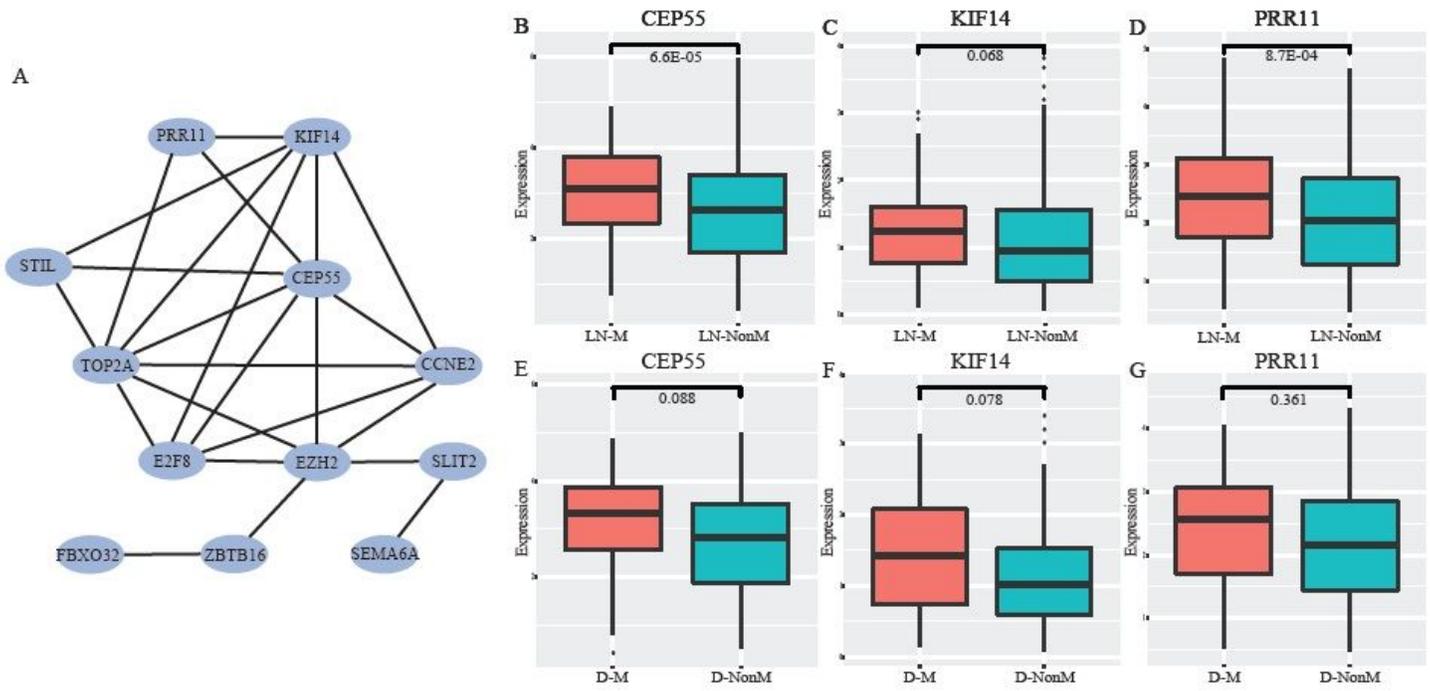


Figure 5

Clinical correlation analysis of genes within protein-protein interaction network. (A) protein-protein interaction network consisting of 12 nodes and 27 edges. (B-D) Expression difference of CEP55, KIF14 and PRR11 between lymphatic node metastasis (LN-M) and non-metastasis (LN-NonM) samples respectively. (E-G) Expression difference of CEP55, KIF14 and PRR11 between distant metastasis (D-M) and non-metastasis (D-NonM) samples respectively.

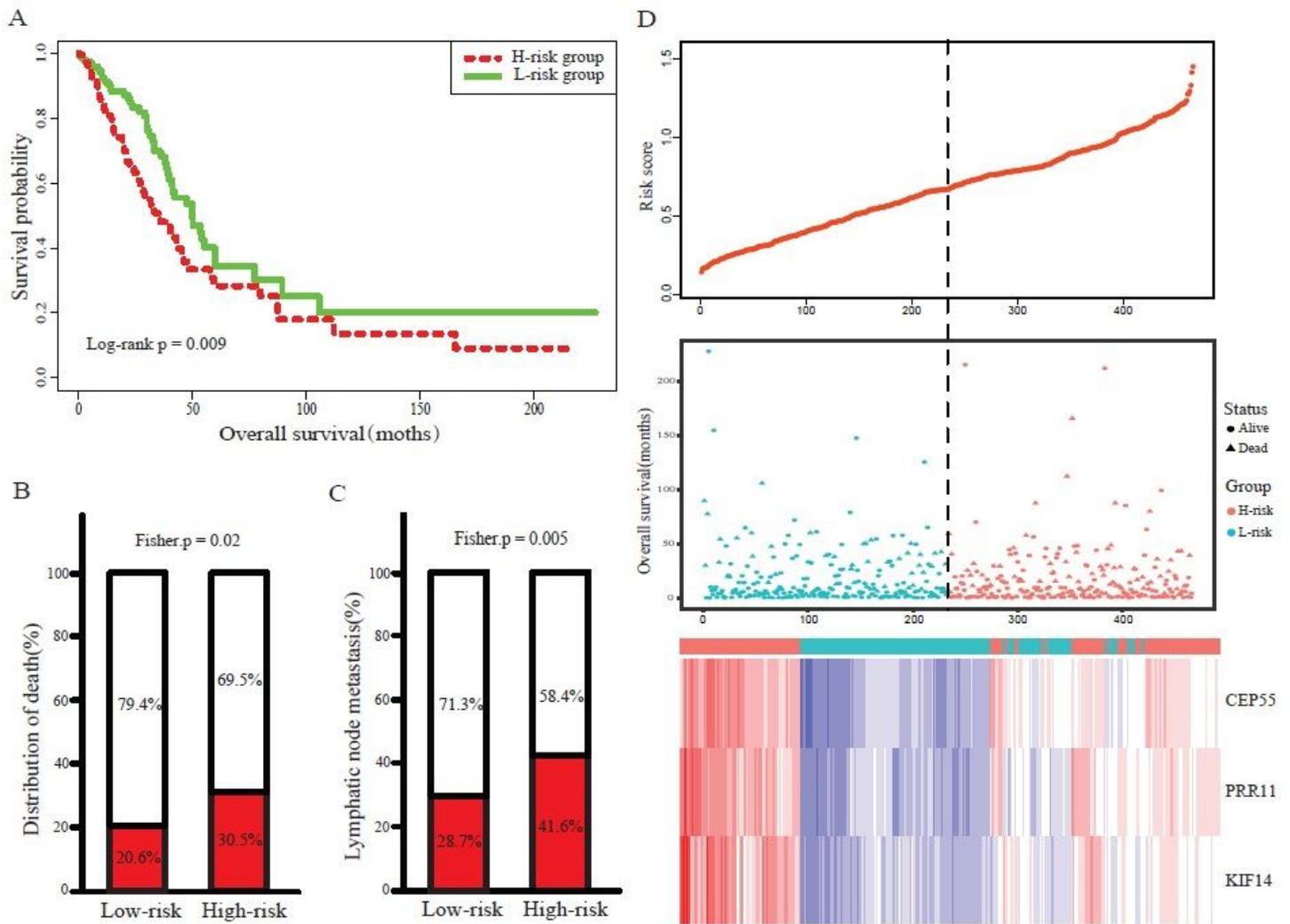


Figure 6

Construction of the three-gene prognostic signature. Patients in TCGA-LUAD cohort were stratified into high-risk group and low-risk group based on median of risk score. (A) Kaplan-Meier curve of the overall survival for high-risk and low-risk scores ranking by the three-gene prognostic signature. (B) The distribution of death in high-risk and low-risk groups. (C) The distribution of lymphatic node metastasis in high-risk and low-risk groups. (D) Risk score distribution, survival status and expression heat map of three genes corresponding to each sample above.

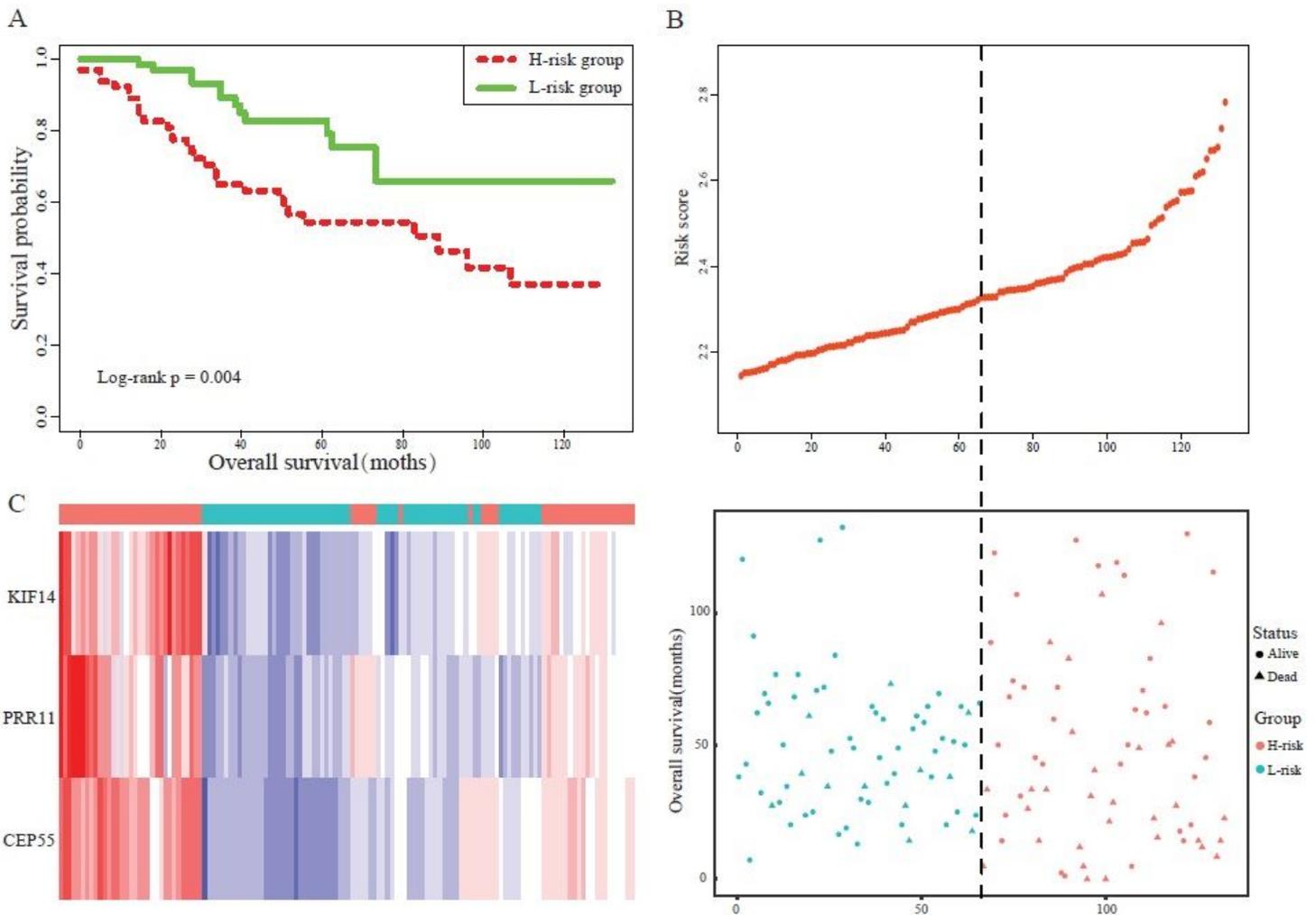


Figure 7

Validation of the three-gene prognostic signature in an independent cohort. Patients in GSE84437 cohort were stratified into high-risk group and low-risk group based on median of risk score. (A) Kaplan-Meier curve of the overall survival for high-risk and low-risk scores ranking by the three-gene prognostic signature. (B) Risk score distribution and survival status for each patient. (C) Expression heat map of three genes corresponding to each sample above.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [TableS2.docx](#)