Competing endogenous RNA network identifies mRNA biomarkers for overall survival of lung adenocarcinoma: two novel on-line precision medicine predictive tools

Jinsong Lin*, Shubiao Lu*, Zhijian Jiang*, Chongjing Hu* and Zhiqiao Zhang*

Department of Internal Medicine, The Affiliated Chencun Hospital of Shunde Hospital, Southern Medical University, Shunde, Guangdong, china

* These authors contributed equally to this work.

ABSTRACT

Background. Individual mortality risk predicted curve at the individual level can provide valuable information for directing individual treatment decision. The present study attempted to explore potential post-transcriptional biological regulatory mechanism related with overall survival of lung adenocarcinoma (LUAD) patients through competitive endogenous RNA (ceRNA) network and develop two precision medicine predictive tools for predicting the individual mortality risk curves for overall survival of LUAD patients.

Methods. Multivariable Cox regression analyses were performed to explore the potential prognostic indicators, which were used to construct a prognostic model for overall survival of LUAD patients. Time-dependent receiver operating characteristic (ROC) curves were used to assess the predictive performance of prognostic model.

Results. There were 494 LUAD patients in model cohort and 233 LUAD patients in validation cohort. Differentially expressed mRNAs, miRNAs, and lncRNAs were identified between LUAD tissues and normal tissues. A ceRNA regulatory network was constructed on previous differentially expressed mRNAs, miRNAs, and lncRNAs. Fourteen mRNA biomarkers were identified as independent risk factors by multivariate Cox regression and used to develop a prognostic model for overall survival of LUAD patients. The C-indexes of prognostic model in model group were 0.786 (95% CI [0.744–0.828]), 0.736 (95% CI [0.694–0.778]) and 0.766 (95% CI [0.724–0.808]) for one year, two year and three year overall survival respectively. Two precision medicine predicted tools were developed for predicting individual mortality risk curves for LUAD patients.

Conclusion. The current study explored potential post-transcriptional biological regulatory mechanism and prognostic biomarkers for overall survival of LUAD patients. Two on-line precision medicine predictive tools were helpful to predict the individual mortality risk predicted curves for overall survival of LUAD patients. Smart Cancer Survival Predictive System could be used at https://zhangzhiqiao2.shinyapps.io/Smart_cancer_predictive_system_9_LUAD_E1002/.

Submitted 8 September 2020
Accepted 15 April 2021
Published 7 May 2021

Corresponding author
Zhiqiao Zhang, sdgrxjbk@smu.edu.cn

Academic editor
Debmalya Barh

Additional Information and Declarations can be found on page 15

DOI **10.7717/peerj.11412**

How to cite this article Lin J, Lu S, Jiang Z, Hu C, Zhang Z. 2021. Competing endogenous RNA network identifies mRNA biomarkers for overall survival of lung adenocarcinoma: two novel on-line precision medicine predictive tools. *PeerJ* **9**:e11412
http://doi.org/10.7717/peerj.11412

Subjects Bioinformatics, Oncology, Respiratory Medicine, Medical Genetics
Keywords Competitive endogenous RNA, Lung adenocarcinoma, mRNA, Overall survival, Prognostic model

Background

Lung cancer is the most common cause of malignant tumours and tumour-related death (*Bray et al., 2018*). There were approximately 2.1 million newly diagnosed lung cancer patients and 1.8 million lung cancer deaths worldwide in 2018 (*Bray et al., 2018*). As the most prevalent type of non-small-cell lung cancer, lung adenocarcinoma (LUAD) is the leading cause of lung cancer-related mortality (*Bray et al., 2018*). The overall survival of patients with LUAD is extremely poor, with a 5-year overall survival (OS) rate of less than 20% (*Lin et al., 2016*). Although great progress has been made in cancer diagnosis and targeted therapy, the 5-year overall survival rate of patients with LUAD remains low (*Qi et al., 2016*). Therefore, establishing a reliable prognostic model for screening high-risk patients with poor overall survival is of great clinical significance for optimizing individualized treatments and improving the management of patients with cancer.

Several studies have reported potential molecular biological regulation mechanisms for different tumours (*Huang et al., 2018*; *Shi et al., 2018*; *Zeng et al., 2018*; *Zhong et al., 2018*). Salmena et al. proposed a posttranscriptional molecular biological regulation mechanism named competing endogenous RNA (ceRNA) (*Salmena et al., 2011*). Long noncoding RNAs (lncRNAs), as endogenous molecular sponges, can upregulate the expression of mRNA through competitive binding with miRNA response elements. (*Thomson & Dinger, 2016*). Some researchers have used a ceRNA regulatory network to explore the potential molecular biological regulation mechanism for the prognosis of patients with LUAD. (*Li et al., 2018*; *Sui et al., 2016*; *Wang et al., 2018*; *Wang et al., 2019*). *Li et al. (2018)* developed a prognostic signature (without external validation) for the prognosis of lung cancer by using 29 mRNAs and three lncRNAs . However, this prognostic model is too complex to calculate and be used for clinical application.

A nomogram is a graphical tool that can directly display the results, with the advantages of easy calculation and good interpretability (*Cheng, 2018*; *Song, Miao & Chen, 2018*). Based on the prognostic nomogram, our research team has further developed precision medicine predictive tools for providing individual mortality risk prediction curves for different cancers (*Cheng et al., 2019*; *Zhang et al., 2019b*; *Zhang et al., 2019c*). Therefore, the present study attempted to construct two precision medicine predictive tools for providing individualized mortality risk prediction curves for the OS of LUAD patients by using prognostic mRNA biomarkers through a ceRNA regulatory network.

MATERIAL AND METHODS

Protocol approval

Data were collected as described in our previous studies (*Zhang et al., 2019a*; *Zhang et al., 2019b*). The study data in the current study were downloaded from The Cancer Genome Atlas (TCGA) database (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga), UCSC Xena database (https://xena.ucsc.edu/), cBioPortal

database (http://www.cbioportal.org/), and Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The current study did not require ethics approval, as all study datasets were downloaded and analysed in accordance with the corresponding data policies of previous databases.

The gene expression information of model cohort

Data were collected as described in our previous studies (*Zhang et al., 2020*; *Zhang et al., 2019b*). The current study downloaded the original RNA expression information from TCGA database (https://tcga-data.nci.nih.gov/docs/publications/tcga/) on the Illumina HiSeq 2000 RNA Sequencing platform. The RNA symbols were determined according to GENCODE Version 29 (https://www.gencodegenes.org/human/). The RNA expression dataset contained 8,848 lncRNAs and 21,138 mRNAs from 535 lung adenocarcinoma samples and 59 normal samples. The miRNA expression dataset was obtained from the UCSC Xena database (https://xena.ucsc.edu/). The miRNA expression dataset involved 1,881 miRNAs from 518 lung adenocarcinoma and 46 normal samples. According to the median of the original gene expression value, the original expression values were converted into "0" for lower expression and "1" for higher expression.

Differentially expressed analyses

Differentially expressed analyses between tumour samples and normal samples were performed by the edgeR package with R software (version 3.5.2) as described in our previous studies (*Zhang et al., 2019b*; *Zhang et al., 2018*). For differentially expressed analyses, *a* false discovery rate (FDR) less than 0.05 and a ratio value of 1.5 between tumour tissues and normal tissues were defined as the thresholds.

The clinical information of model cohort

Data were collected as described in our previous studies (*Zhang et al., 2019b*; *Zhang et al., 2018*). The clinical information was obtained from the cBioPortal database. Patients with LUAD with inadequate survival information or OS times less than one month (for living patients) were removed ($n = 19$). Therefore, 503 LUAD patients were identified with adequate OS information. There were 494 patients with adequate gene detection information and survival information when performing intersection between clinical datasets on the gene detection dataset.

The corresponding information of validation cohort

Data were collected as described in our previous studies (*Zhang et al., 2019b*; *Zhang et al., 2018*). The current study downloaded GSE37745 and GSE50081 as external validation datasets from the GEO database. Gene expression information was detected on the GPL570 platform (Affymetrix Genome U133 Plus 2.0 Array). The verification group contained 233 LUAD patients and 22,850 RNA expression count values.

Statistical analysis

The statistical analyses were carried out as described in our previous studies (*Zhang et al., 2020*; *Zhang et al., 2019b*): continuous variables are presented as the mean ± standard deviation (SD) or median (first percentile, third percentile) as appropriate. Continuous

variables were compared by t-test or Mann–Whitney U test as appropriate. Categorical variables were compared through the chi-squared test or Fisher's exact test as appropriate. Time-dependent receiver operating characteristic (ROC) curves were used to assess the predictive performance of prognostic models. Decision curve analysis (DCA) is an assessment method to compare the predictive performance of prognostic models (*Localio & Goodman, 2012*; *Vickers et al., 2008*; *Vickers & Elkin, 2006*). The statistical analyses were carried out using SPSS Statistics 19.0 (SPSS Inc., USA) and R software (version 3.5.2) with the following packages: "GOplot", "timeROC", "rms", "pROC", "survival", "clusterProfler" and "glmnet", as described in our previous studies (*Zhang et al., 2020*; *Zhang et al., 2019b*). A P value < 0.05 was the threshold value for statistical significance.

RESULTS

Study cohorts

In the model group (Doc S1), 182 patients (36.8%) out of 494 patients died during the follow-up period, while 127 patients (54.5%) out of 233 patients died in the verification group (Doc S2). The basic features of the two groups are compared in Table 1.

Differentially expressed analyses

Based on the given threshold, 3,310 upregulated lncRNAs, 675 downregulated lncRNAs, 95 upregulated miRNAs, 30 downregulated miRNAs, 4,913 upregulated mRNAs, and 1,921 downregulated mRNAs were identified as differentially expressed RNAs between LUAD tissues and normal tissues. Based on previous differentially expressed mRNAs, 2,982 mRNAs were identified as prognostic mRNA indicators by univariate Cox analyses.

Construction of a competing endogenous RNA network

To obtain the lncRNA-miRNA pairs, the miRcode database (http://www.mircode.org/) was searched according to the differentially expressed lncRNAs. To obtain the miRNA-mRNA pairs, TargetScan (http://www.targetscan.org/), miRDB (http://mirdb.org/), and miRTarBase (https://bio.tools/mirtarbase) were searched according to the previous lncRNA-targeted miRNAs. The miRNA-predicted mRNAs that could be searched in three databases were defined as miRNA-targeted mRNAs. Then, we determined the interaction between miRNA-targeted mRNAs and prognostic mRNAs to identify prognosis-associated miRNA-targeted mRNAs. Finally, twenty-two lncRNAs, twenty-nine miRNAs, and seventy-three mRNAs were used to build the ceRNA network for the overall survival of Patients with LUAD. The ceRNA regulatory network was depicted by Cytoscape software (Fig. 1).

Functional enrichment analyses

To explore the biological functions of prognostic mRNAs, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed on prognostic mRNAs in the ceRNA network. The bar plot and dot plot of prognostic mRNAs are depicted in Figs. 2A and 2B. GO terms and KEGG pathways of prognostic mRNAs are depicted in Figs. 3A and 3B.

Table 1 The clinical features of lung adenocarcinoma patients in the model group and validation group.

	Model group ($n = 494$)	Validation group ($n = 233$)	*P* value
Death [n(%)]	182(36.8)	127(54.5)	<0.001
Survival time for living patients(mean ± SD, month)	22.2(15.6,37.6)	69.1(59.3,88.0)	<0.001
Survival time for dead patients (mean ± SD, month)	20.3(9.8,36.8)	25.3(12.4,48.4)	0.006
Age (mean ± SD, year)	65.3 ± 10.0	66.1 ± 9.9	0.318
Male [(n)%]	229(46.4)	111(47.6)	0.746
AJCC Stage (IV)	25	4	<0.001
AJCC Stage (III)	80	13	
AJCC Stage (II)	116	54	
AJCC Stage (I)	273	162	
AJCC Stage (NA)	0	0	
AJCC PT (T1)	229	NA	NA
AJCC PT (T0)	265	NA	
AJCC PT (NA)	0	NA	
AJCC PN (N4)	18	NA	NA
AJCC PN (N3)	44	NA	
AJCC PN (N2)	264	NA	
AJCC PN (N1)	165	NA	
AJCC PN (N0)	3	NA	
AJCC PN (NA)	0	NA	
AJCC PM (M3)	2	NA	NA
AJCC PM (M2)	69	NA	
AJCC PM (M1)	92	NA	
AJCC PM (M0)	302	NA	
AJCC PM (NA)	11	NA	

Note
Continuous variables were compared by *t*-test or Mann-Whitney *U* test as appropriate; Categorical variables were compared by chi-squared test or Fisher's exact test as appropriate.
NA, missing data; SD, standard deviation; AJCC, American Joint Committee on Cancer.

Development of a prognostic nomogram

The previous prognostic mRNAs were used to construct a prognostic model for overall survival. The information on these mRNAs is summarized in Table 2. The risk score of the prognostic model was calculated by using the following formula: risk score = (−0.7305*DNAJC27) + (0.4776*NPAS2) + (0.3941*PHKA1) + (−0.5537*CDADC1) + (0.4792*PTGFRN) + (0.3980*DDIT4) + (−0.4133*SCAMP5) + (−0.4367*SPRY2) + (0.3850*CSE1L) + (−0.3940*ELAVL4) + (0.3293*TRIM29) + (0.3685*LPGAT1) + (−0.3279*TRPC3) + (0.3257*DCBLD2). A prognostic nomogram chart is presented in Fig. 4. Kaplan–Meier survival curves (Fig. S1) demonstrated that these 14 prognostic mRNAs were significantly correlated with OS ($P < 0.05$). The results of differential expression analysis of 14 enrolled mRNAs are listed in Table 3.

Predictive performance of the prognostic model

Based on the median value, LUAD patients in the model group could be classified into a high-risk subgroup and a low-risk subgroup. The overall survival rate (Fig. 5A) in the

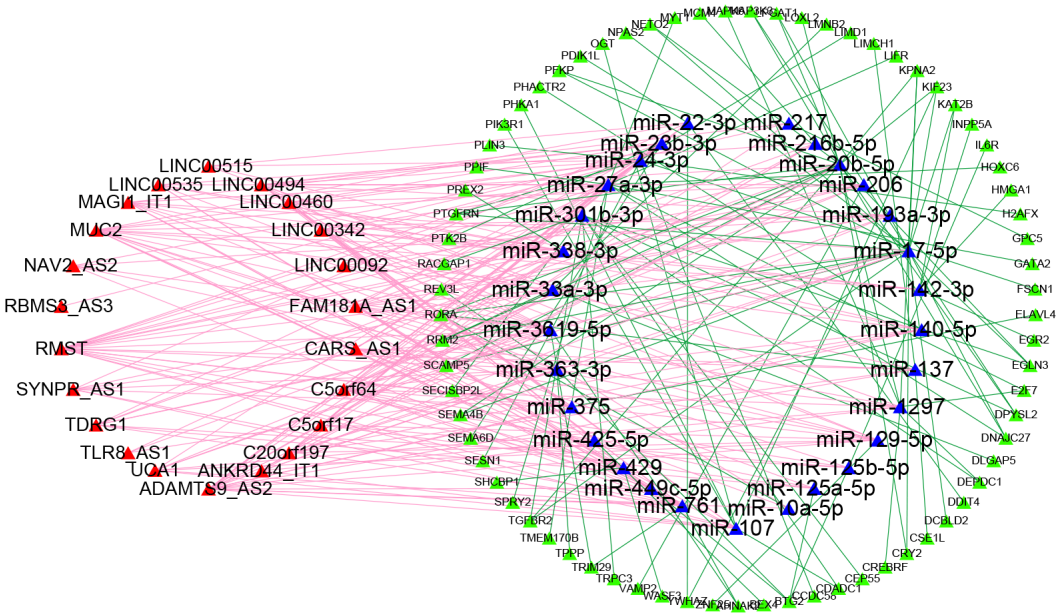

Figure 1 **Competitive endogenous RNA network chart: the red triangles represent 22 lncRNAs; the blue triangles represent 29 miRNAs; the green circles represent 73 mRNAs.**

Full-size DOI: 10.7717/peerj.11412/fig-1

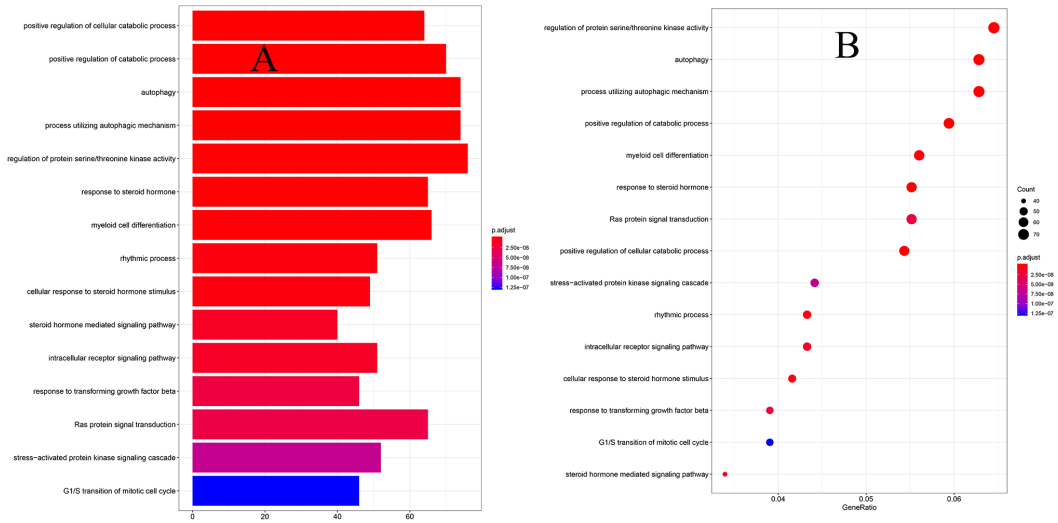

Figure 2 **(A) The barplot of GO terms for prognostic mRNAs; (B) the dotplot of GO terms for prognostic mRNAs.**

Full-size DOI: 10.7717/peerj.11412/fig-2

high-risk group was significantly poorer than that in the low-risk group ($P < 0.001$). Harrell's concordance indexes (C-indexes) of the prognostic signature for overall survival in the model group were 0.786 (95% CI [0.744–0.828]), 0.736 (95% CI [0.694–0.778]), and 0.766 (95% CI [0.724–0.808]) for 1-year, 2-year, and 3-year OS, respectively (Fig. 5B). The scatter plot and the interaction distribution scatter plot are presented in Fig. 5C and

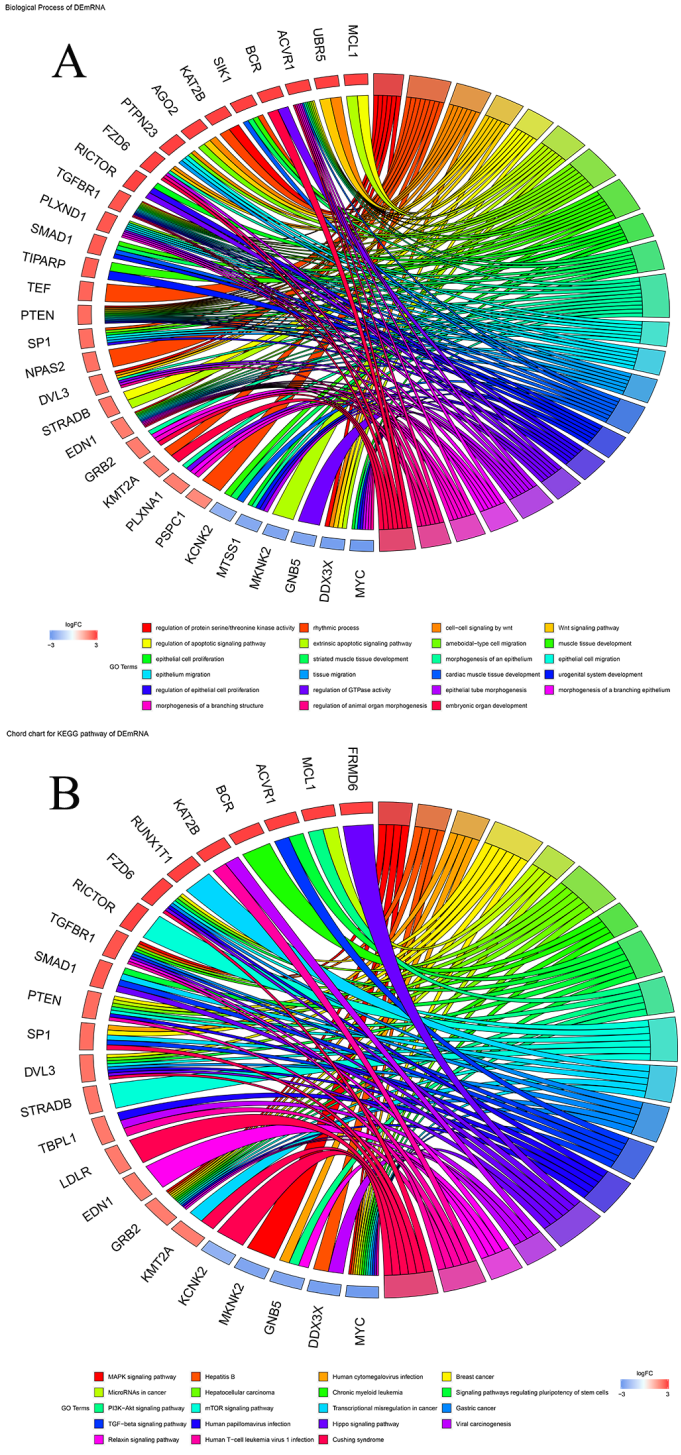

Figure 3 **(A) Chord chart of GO terms for prognostic mRNAs; (B) KEGG pathways for prognostic mRNAs.**

Full-size DOI: 10.7717/peerj.11412/fig-3

Table 2 The model information of prognostic mRNAs in univariate and multivariable Cox regression analyses.

	Univariate analysis			Multivariate analysis			
Gene	**HR**	**95% CI**	***P*-value**	**coefficient**	**HR**	**95% CI**	***P*-value**
DNAJC27 (High/Low)	0.594	0.441–0.800	<0.001	−0.731	0.482	0.336–0.691	<0.001
NPAS2 (High/Low)	1.613	1.202-2.164	<0.001	0.478	1.612	1.166–2.229	0.004
PHKA1 (High/Low)	1.367	1.020–1.832	0.036	0.394	1.483	1.032–2.131	0.033
CDADC1 (High/Low)	0.699	0.521–0.936	0.016	−0.554	0.575	0.394–0.839	0.004
PTGFRN (High/Low)	1.356	1.012–1.818	0.042	0.479	1.615	1.120–2.329	0.010
DDIT4 (High/Low)	1.634	1.217–2.193	<0.001	0.398	1.489	1.086–2.042	0.013
SCAMP5 (High/Low)	0.737	0.550–0.988	0.041	−0.413	0.662	0.474–0.924	0.015
SPRY2 (High/Low)	0.714	0.533–0.957	0.024	−0.437	0.646	0.455–0.918	0.015
CSE1L (High/Low)	1.342	1.001–1.798	0.049	0.385	1.470	1.024–2.110	0.037
ELAVL4 (High/Low)	0.707	0.527–0.950	0.021	−0.394	0.674	0.496–0.918	0.012
TRIM29 (High/Low)	1.391	1.037–1.864	0.028	0.329	1.390	1.010–1.912	0.043
LPGAT1 (High/Low)	1.410	1.052–1.890	0.022	0.369	1.446	1.017–2.054	0.040
TRPC3 (High/Low)	0.681	0.506–0.915	0.011	−0.328	0.720	0.525–0.988	0.042
DCBLD2 (High/Low)	1.346	1.005–1.802	0.046	0.326	1.385	1.001–1.915	0.049

Notes.

HR, hazard ratio; CI, confidence interval.

The medians of mRNA expression values were used as cut-off values to stratify mRNA expression values into high expression group (as value 1) and low expression group (as value 0).

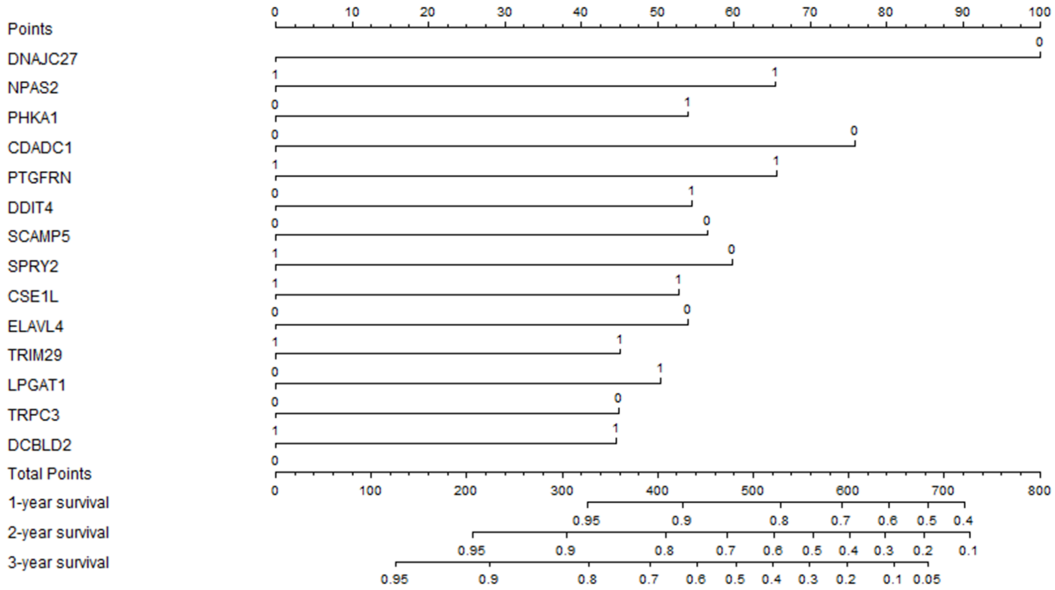

Figure 4 The prognostic nomogram for overall survival.

Full-size DOI: 10.7717/peerj.11412/fig-4

Fig. 5D. Calibration curves are depicted in Fig. S2A for 1 year, Fig. S2B for 2 years, and Fig. S2C for 3-year overall survival.

Table 3 Results of differential expression analysis of the enrolled mRNAs. .

Symbol	F	*P* value	FC	logCPM	FDR
PHKA1	207.115	<0.001	3.252	4.717	<0.001
CDADC1	125.913	<0.001	0.666	2.854	<0.001
LPGAT1	116.687	<0.001	2.218	6.843	<0.001
PTGFRN	104.576	<0.001	2.304	6.604	<0.001
DNAJC27	95.382	<0.001	0.621	2.815	<0.001
SCAMP5	66.600	<0.001	2.384	4.159	<0.001
CSE1L	63.349	<0.001	1.694	6.671	<0.001
NPAS2	52.396	<0.001	2.416	4.793	<0.001
TRPC3	44.057	<0.001	0.429	−0.488	<0.001
SPRY2	42.334	<0.001	0.629	5.171	<0.001
ELAVL4	41.979	<0.001	3.749	−0.076	<0.001
DDIT4	34.794	<0.001	2.249	7.135	<0.001
DCBLD2	30.157	<0.001	2.262	7.216	<0.001
TRIM29	7.324	0.007	2.108	5.400	0.010

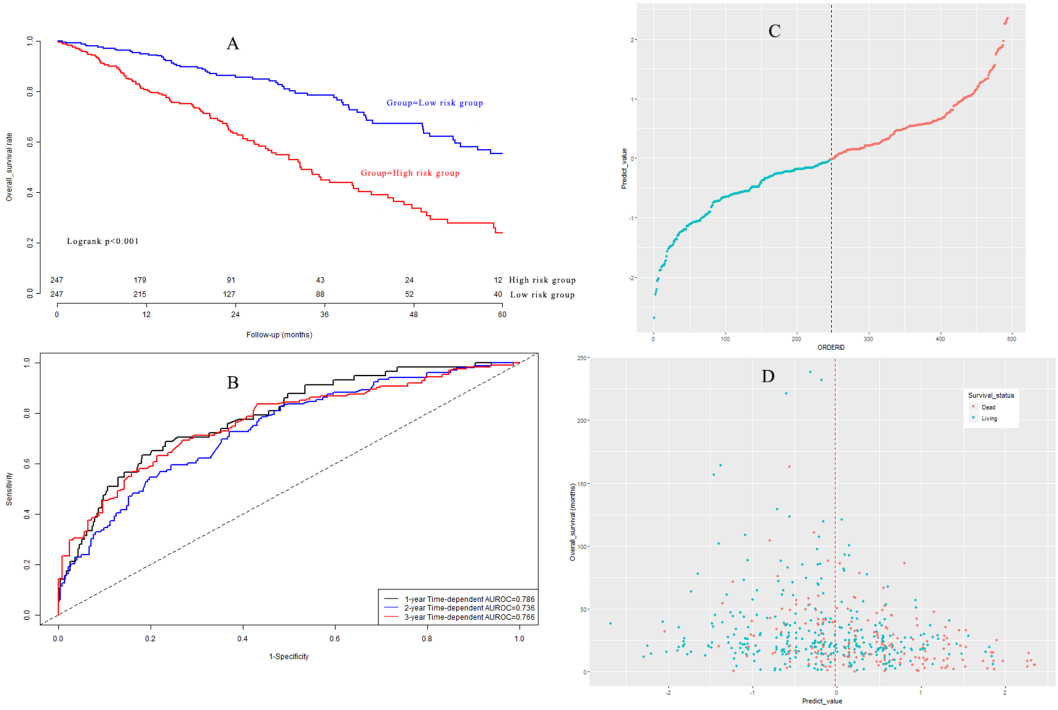

Figure 5 (A) Survival curves in model group; (B) time-dependent receiver operating characteristic curves in model group; (C) the distribution of prognostic model score in model group; (D) the overall survival status and overall survival time in model group.

Full-size DOI: 10.7717/peerj.11412/fig-5

External validation of the prognostic model

Kaplan–Meier analysis (Fig. 6A) indicated that there was a significant difference in OS between the low-risk group and the high-risk group in the validation dataset ($P < 0.001$).

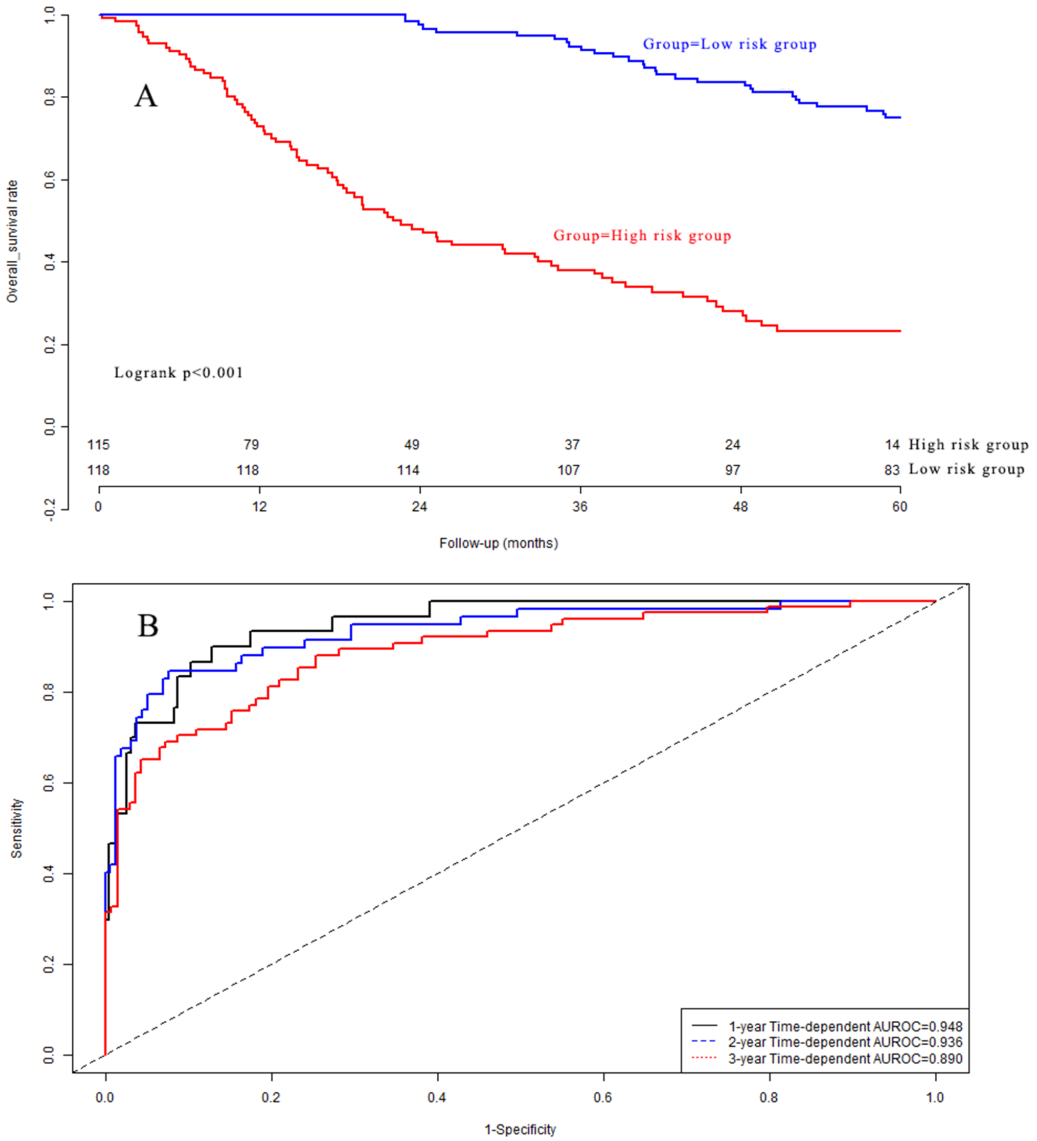

Figure 6 **(A) Survival curves in validation group; (B) time-dependent receiver operating characteristic curves in validation group.**

Full-size DOI: 10.7717/peerj.11412/fig-6

The C-indexes of the prognostic model for overall survival in the validation dataset were 0.948 (95% CI [0.912–0.984]), 0.936 (95% CI [0.900–0.972]), and 0.890 (95% CI [0.854–0.926]) for 1-year, 2-year, and 3-year overall survival, respectively (Fig. 6B). Calibration curves are depicted in Fig. S3A for 1 year, Fig. S3B for 2 years, and Fig. S3C for 3-year overall survival.

Independence assessment of the prognostic model

Table 4 shows that the prognostic model was an independent influencing factor for OS after adjustment for confounding effects in the model dataset. In the validation dataset, multivariate Cox regression analyses demonstrated that the prognostic model was an

Table 4 Univariate and multivariable Cox regression analyses for independence assessment of prognostic model.

	Univariate analysis			Multivariate analysis			
	HR	95% CI	*P*-value	coefficient	HR	95% CI	*P*-value
Model group ($n = 494$)							
Age(≥60/<60 years)	1.003	0.722–1.392	0.987	0.078	1.082	0.776–1.507	0.643
Gender (Male/Female)	1.048	0.783–1.403	0.754	0.077	1.080	0.805–1.449	0.609
AJCC stage (IV,III/II,I)	2.599	1.906–3.545	<0.001	0.800	2.225	1.622–3.051	<0.001
Prognostic model (High/Low)	2.705	1.986–3.684	<0.001	0.895	2.447	1.786–3.352	<0.001
Validation group($n = 233$)							
Age(≥60/<60 years)	0.876	0.595–1.289	0.501	−0.271	0.762	0.512–1.136	0.183
Gender (Male/Female)	1.308	0.921–1.857	0.134	0.268	1.308	0.911–1.877	0.146
AJCC stage (IV,III/II,I)	0.988	0.517–1.888	0.974	0.354	1.424	0.738–2.750	0.292
Prognostic model (High/Low)	5.109	3.428–7.348	<0.001	1.651	5.214	3.531–7.701	<0.001

Notes.
AJCC, the American Joint Committee on Cancer; HR, hazard ratio; CI, confidence interval.
The median of Prognostic model scores was used as the cut-off value to stratify gastric cancer patients into high risk group and low risk group.

independent influencing factor for OS. DCA is shown in Fig. S4A for 1-year OS, Fig. S4B for 2-year OS, and Fig. S4C for 3-year OS. The clinical impact curve is presented in Fig. S4D.

Smart Cancer Survival Predictive System

A precision medicine predictive tool named the Smart Cancer Survival Predictive System was developed for predicting the prognosis of LUAD patients. The Smart Cancer Survival Predictive System (Fig. 7) is available at https://zhangzhiqiao2.shinyapps.io/Smart_cancer_predictive_system_9_LUAD_E1002/.

The Smart Cancer Survival Predictive System could predict full-time mortality risk prediction curves for one specific patient (Fig. 7A). Figure 7B, demonstrates the mortality rate predicted percentage and 95% confidence interval at different user-defined time points.

Gene Survival Analysis Screen System

A second precision medicine predictive tool named the Gene Survival Analysis Screen System (Fig. 8) was developed to explore the survival features of prognostic mRNAs. The Gene Survival Analysis Screen System is available at https://zhangzhiqiao7.shinyapps.io/Gene_Survival_Analysis_Screening_System_9_LUAD_E1002/. Figure 8A depicts and compares the survival curves between two defined subgroups. Figure 8B displays the results of univariate survival analysis for selected variables.

DISCUSSION

The current ceRNA regulatory network depicts lncRNA-miRNA-mRNA regulatory pathways, which are helpful for understanding the biological regulatory mechanisms of OS in LUAD. The current study constructed and verified a fourteen-mRNA prognostic nomogram for OS. This fourteen-mRNA prognostic nomogram was suitable to screen LUAD patients with poor OS. Based on this fourteen-mRNA prognostic nomogram, we

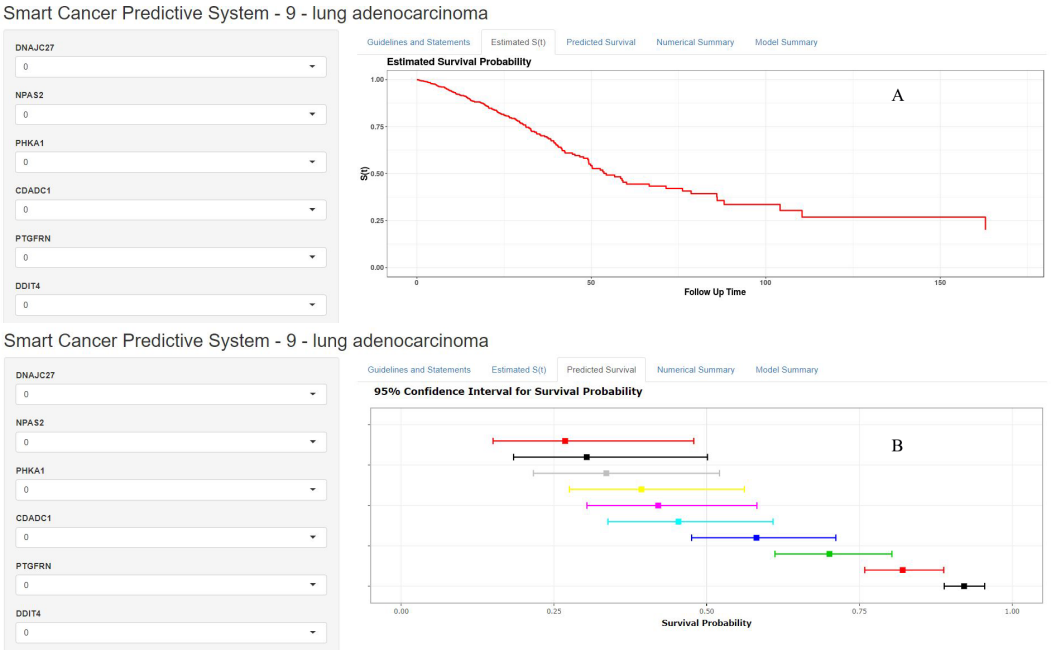

Figure 7 **Home page of Smart Cancer Survival Predictive System: (A) individual mortality risk predictive curve display page; (B) different time-point individual mortality risk prediction display page.**
Full-size DOI: 10.7717/peerj.11412/fig-7

developed an online precision medicine predictive tool named the Smart Cancer Survival Predictive System, which can generate full-time mortality risk prediction curves for one specific patient at the individual level.

Several prognostic models have been built for predicting the prognosis of lung cancer patients (*Xie & Xie, 2019*; *Yan et al., 2018*; *Zuo et al., 2019*). However, these previous prognostic models could only provide prognostic information for a special subgroup at the group level and failed to provide individual mortality risk prediction at the individual level. Our Smart Cancer Survival Predictive System was superior to the previous prognostic models for its special predictive ability in predicting individual mortality risk curves at the individual level. Meanwhile, the Smart Cancer Survival Predictive System could further provide the mortality rate predicted percentage and 95% confidence interval at different user-defined time points. These special predictive functions in smart cancer survival predictive systems are of great significance for improving individual treatment decisions.

NPAS2 was associated with a favourable prognosis of LUAD patients (*Qiu et al., 2019*). CDADC1 was found to be associated with survival in bladder cancer (*Shivakumar et al., 2017*). DDIT4 promotes gastric cancer proliferation and tumorigenesis (*Du et al., 2018*). The knockdown or overexpression of SPRY2 promoted or suppressed the proliferation of prostate cancer cells (*Gao et al., 2018*). CSE1L was correlated with overall survival in patients with hepatocellular carcinoma (*Zhang et al., 2019c*). Ectopic expression of TRIM29 potentially contributes to metastasis and poor prognosis in patients with osteosarcoma (*Zeng et al., 2017*). Increased activity of TRPC3 channels is necessary for the development of ovarian cancer (*Yang et al., 2009*). DCBLD2 correlated with the overall survival in

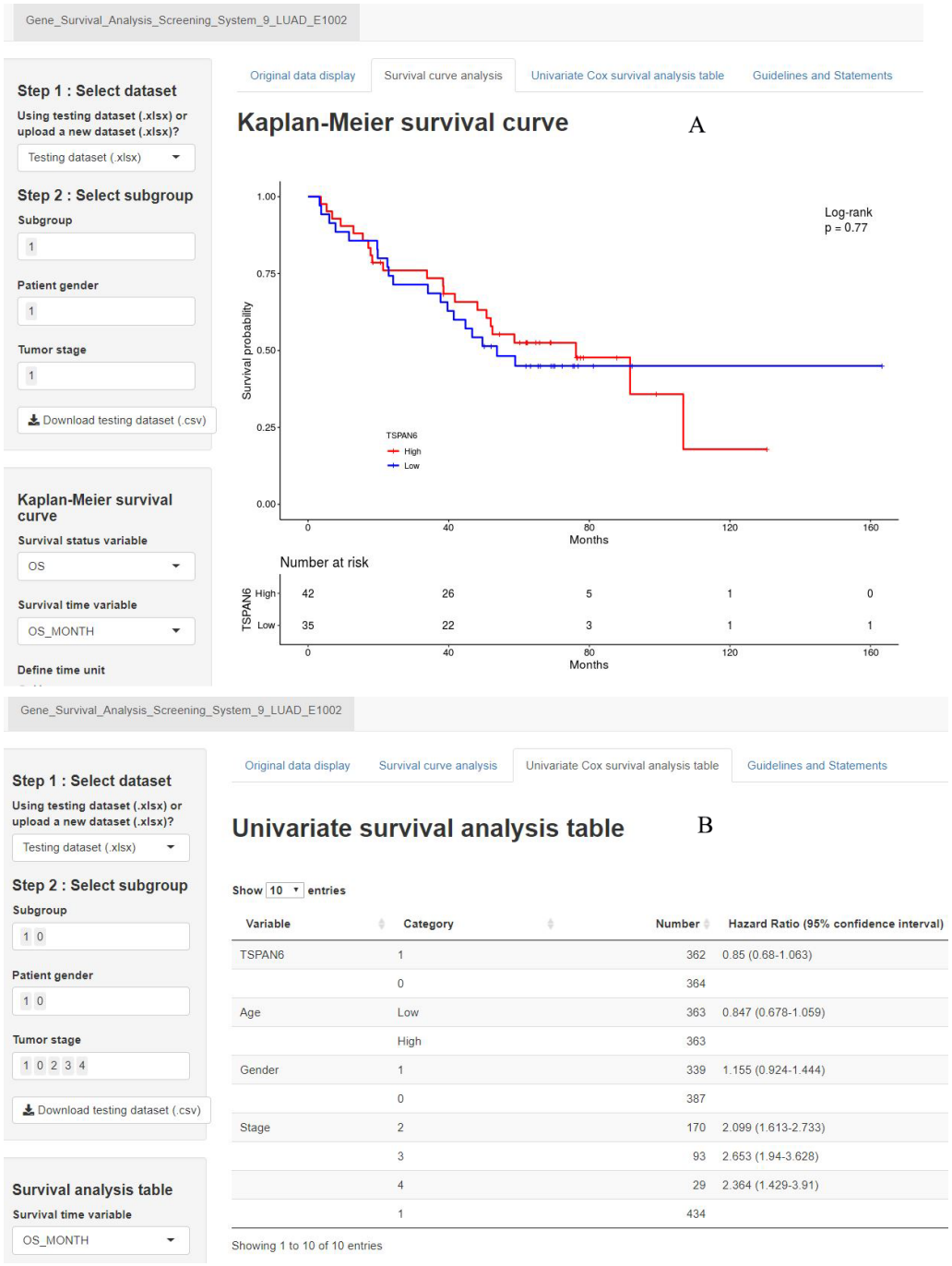

Figure 8 **Home page of Gene Survival Analysis Screen System: (A) survival curve plot display page; (B) univariate analysis table display page.**

Full-size DOI: 10.7717/peerj.11412/fig-8

pancreatic ductal adenocarcinoma (*Raman et al., 2018*). The results of the current study were in good agreement with those of previous studies.

There are four main advantages of the current study: First, all operations are computed automatically in the background without manual calculation, which is convenient for use; Second, our tool can provide full-time individual mortality risk prediction through an individualized survival curve for a particular patient; Third, our tool can provide individual mortality risk prediction percentages and 95% CIs at specific time points (such as 12, 24, 36, 48, 60, and 72 months) through line graphs and tables; Fourth, the Gene Survival Analysis Screen System allows users to define different subgroups by themselves. Meanwhile, users are free to download, upload, and select the dataset for gene survival analysis. To the best of our knowledge, this online full-time individualized risk calculator is the first to provide full-time individual mortality risk prediction through a web calculator based on gene expression data for patients with lung cancer.

The shortcomings of our study were as follows: First, the genetic detection platforms of the model group and validation group were different and need to be taken into account when interpreting the results of the current study; Second, the performance of the fourteen-mRNA prognostic nomogram in the verification group was better than that in the model group, which could be explained by the longer follow-up period and higher mortality in the verification group. The median survival times were 20.3 and 25.3 months ($P = 0.006$) for patients who died in the model group and validation group, respectively, whereas the median survival times were 22.2 and 69.1 months ($P < 0.001$) for living patients in the model group and validation group, respectively; Third, some important prognostic factors, such as surgical procedures and adjuvant therapies, were not included in the survival analysis. It is necessary to carry out multicentre, prospective, and large-sample clinical research to further study the clinical application value of fourteen-mRNA prognostic nomograms in different populations.

CONCLUSIONS

In conclusion, the current study explored potential posttranscriptional biological regulatory mechanisms and prognostic biomarkers for the overall survival of LUAD patients. Two online precision medicine predictive tools were developed and are helpful to predict the individual mortality risk prediction curves for the overall survival of LUAD patients. The Smart Cancer Survival Predictive System can be used at https://zhangzhiqiao2.shinyapps.io/Smart_cancer_predictive_system_9_LUAD_E1002/.

Abbreviations

LUAD	lung adenocarcinoma
TCGA	The Cancer Genome Atlas
GEO	Gene Expression Omnibus
ROC	receiver operating characteristic
OS	overall survival
lncRNA	long noncoding RNA
miRNA	microRNA
mRNA	messenger RNA
HR	hazard ratio

CI	confidence interval
AJCC	American Joint Committee on Cancer
SD	standard deviation
DCA	decision curve analysis
ceRNA	competitive endogenous RNA

ACKNOWLEDGEMENTS

We would like to express sincere thanks to Dr. Gary S Collins (University of Oxford), Dr Manali Rupji (Emory University), and Mrs Qingmei Liu for inspirations, suggestions, and assistance in the development of our precision medicine tools.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Guangdong Provincial Health Department (A2016450 and B2018237). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Guangdong Provincial Health Department: A2016450, B2018237.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Jinsong Lin, Shubiao Lu, Zhijian Jiang, Chongjing Hu and Zhiqiao Zhang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The current study did not require ethics approval as all study datasets were downloaded and analyzed in accordance with the corresponding data policies of previous databases.

Data Availability

The following information was supplied regarding data availability:

The study data is available in the Supplemental Files. The study data in the current study were downloaded from The Cancer Genome Atlas (TCGA) database (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga), UCSC Xena database (https://xena.ucsc.edu/), cBioPortal database (http://www.cbioportal.org/), and the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/).

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.11412#supplemental-information.

REFERENCES

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68**:394–424.

Cheng C, Wang Q, Zhu M, Liu K, Zhang Z. 2019. Integrated analysis reveals potential long non-coding RNA biomarkers and their potential biological functions for disease free survival in gastric cancer patients. *Cancer Cell International* **19**:123 DOI 10.1186/s12935-019-0846-6.

Cheng P. 2018. A prognostic 3-long noncoding RNA signature for patients with gastric cancer. *Journal of Cellular Biochemistry* **119**:9261–9269 DOI 10.1002/jcb.27195.

Du F, Sun L, Chu Y, Li T, Lei C, Wang X, Jiang M, Min Y, Lu Y, Zhao X, Nie Y, Fan D. 2018. DDIT4 promotes gastric cancer proliferation and tumorigenesis through the p53 and MAPK pathways. *Cancer Commun* **38**:45.

Gao W, Hong Z, Huang H, Zhu A, Lin S, Cheng C, Zhang X, Zou G, Shi Z. 2018. miR-27a in serum acts as biomarker for prostate cancer detection and promotes cell proliferation by t argeting Sprouty2. *Oncology Letters* **16**:5291–5298.

Huang Y, Xiang B, Liu Y, Wang Y, Kan H. 2018. LncRNA CDKN2B-AS1 promotes tumor growth and metastasis of human hepatocellular carcinoma by targeting let-7c-5p/NAP1L1 axis. *Cancer Letters* **437**:56–66 DOI 10.1016/j.canlet.2018.08.024.

Li X, Li B, Ran P, Wang L. 2018. Identification of ceRNA network based on a RNA-seq shows prognostic lncRNA biomarkers in human lung adenocarcinoma. *Oncology Letters* **16**:5697–5708 DOI 10.3892/ol.2018.9336.

Lin JJ, Cardarella S, Lydon CA, Dahlberg SE, Jackman DM, Janne PA, Johnson BE. 2016. Five-year survival in EGFR-mutant metastatic lung adenocarcinoma treated with EGFR-TKIs. *J Thorac Oncol* **11**:556–565 DOI 10.1016/j.jtho.2015.12.103.

Localio AR, Goodman S. 2012. Beyond the usual prediction accuracy metrics: reporting results for clinical decision making. *Annals of Internal Medicine* **157**:294–295 DOI 10.7326/0003-4819-157-4-201208210-00014.

Qi L, Li Y, Qin Y, Shi G, Li T, Wang J, Chen L, Gu Y, Zhao W, Guo Z. 2016. An individualised signature for predicting response with concordant survival benefit for lung adenocarcinoma patients receiving platinum-based chemotherapy. *British Journal of Cancer* **115**:1513–1519 DOI 10.1038/bjc.2016.370.

Qiu M, Chen YB, Jin S, Fang XF, He XX, Xiong ZF, Yang SL. 2019. Research on circadian clock genes in non-small-cell lung carcinoma. *Chronobiology International* **36**:739–750 DOI 10.1080/07420528.2018.1509080.

Raman P, Maddipati R, Lim KH, Tozeren A. 2018. Pancreatic cancer survival analysis defines a signature that predicts outcome. *PLOS ONE* **13**:e0201751 DOI 10.1371/journal.pone.0201751.

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.

Shi X, Zhao Y, He R, Zhou M, Pan S, Yu S, Xie Y, Li X, Wang M, Guo X, Qin R. 2018. Three-lncRNA signature is a potential prognostic biomarker for pancreatic adenocarcinoma. *Oncotarget* **9**:24248–24259 DOI 10.18632/oncotarget.24443.

Shivakumar M, Lee Y, Bang L, Garg T, Sohn KA, Kim D. 2017. Identification of epigenetic interactions between miRNA and DNA methylation associated with gene expr ession as potential prognostic markers in bladder cancer. *BMC Medical Genomics* **10**:30.

Song W, Miao DL, Chen L. 2018. Nomogram for predicting survival in patients with pancreatic cancer. *OncoTargets and Therapy* **11**:539–545 DOI 10.2147/ott.s154599.

Sui J, Li YH, Zhang YQ, Li CY, Shen X, Yao WZ, Peng H, Hong WW, Yin LH, Pu YP, Liang GY. 2016. Integrated analysis of long non-coding RNAassociated ceRNA network reveals potential lncRNA biomarkers in human lung adenocarcinoma. *International Journal of Oncology* **49**:2023–2036 DOI 10.3892/ijo.2016.3716.

Thomson DW, Dinger ME. 2016. Endogenous microRNA sponges: evidence and controversy. *Nature Reviews Genetics* **17**:272–283 DOI 10.1038/nrg.2016.20.

Vickers AJ, Cronin AM, Elkin EB, Gonen M. 2008. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Medical Informatics and Decision Making* **8**:53 DOI 10.1186/1472-6947-8-53.

Vickers AJ, Elkin EB. 2006. Decision curve analysis: a novel method for evaluating prediction models. *Medical Decision Making* **26**:565–574 DOI 10.1177/0272989x06295361.

Wang X, Ding Y, Da B, Fei Y, Feng G. 2018. Identification of potential prognostic long noncoding RNA signatures based on a competing endogenous RNA network in lung adenocarcinoma. *Oncology Reports* **40**:3199–3212 DOI 10.3892/or.2018.6719.

Wang Y, Lu T, Wo Y, Sun X, Li S, Miao S, Dong Y, Leng X, Jiao W. 2019. Identification of a putative competitive endogenous RNA network for lung adenocarcinoma using TCGA datasets. *PeerJ* **7**:e6809 DOI 10.7717/peerj.6809.

Xie H, Xie C. 2019. A six-gene signature predicts survival of adenocarcinoma type of non-small-cell lung cancer patients: a comprehensive study based on integrated analysis and weighted gene coexpression network. *BioMed Research International* **2019**:4250613 DOI 10.1155/2019/4250613.

Yan H, Xin S, Ma J, Wang H, Zhang H, Liu J. 2018. A three microRNA-based prognostic signature for small cell lung cancer overall survival. *Journal of Cellular Biochemistry* **120**:8723–8730 DOI 10.1002/jcb.28159.

Yang SL, Cao Q, Zhou KC, Feng YJ, Wang YZ. 2009. Transient receptor potential channel C3 contributes to the progression of human ovarian cancer. *Oncogene* **28**:1320–1328 DOI 10.1038/onc.2008.475.

Zeng J, Cai X, Hao X, Huang F, He Z, Sun H, Lu Y, Lei J, Zeng W, Liu Y, Luo R. 2018. LncRNA FUNDC2P4 down-regulation promotes epithelial-mesenchymal

transition by reducing E-cadherin exp ression in residual hepatocellular carcinoma after insufficient radiofrequency ablation. *International Journal of Hyperthermia* **34**:802–811 DOI 10.1080/02656736.2017.1422030.

Zeng SX, Cai QC, Guo CH, Zhi LQ, Dai X, Zhang DF, Ma W. 2017. High expression of TRIM29?(ATDC) contributes to poor prognosis and tumor metastasis by inducing epith elial?mesenchymal transition in osteosarcoma. *Oncology Reports* **38**:1645–1654 DOI 10.3892/or.2017.5842.

Zhang Z, He T, Huang L, Ouyang Y, Li J, Huang Y, Wang P, Ding J. 2019a. Two precision medicine predictive tools for six malignant solid tumors: from gene-based research to clinical application. *Journal of Translational Medicine* **17**:405 DOI 10.1186/s12967-019-02151-8.

Zhang Z, Li J, He T, Ding J. 2020. Bioinformatics identified 17 immune genes as prognostic biomarkers for breast cancer: application study based on artificial intelligence algorithms. *Frontiers in Oncology* **10**:330 DOI 10.3389/fonc.2020.00330.

Zhang Z, Li J, He T, Ouyang Y, Huang Y, Liu Q, Wang P, Ding J. 2019b. The competitive endogenous RNA regulatory network reveals potential prognostic biomarkers for overall survival in hepatocellular carcinoma. *Cancer Science* **110**:2905–2923 DOI 10.1111/cas.14138.

Zhang Z, Liu Q, Wang P, Li J, He T, Ouyang Y, Huang Y, Wang W. 2018. Development and internal validation of a nine-lncRNA prognostic signature for prediction of overall survival in colorectal cancer patients. *PeerJ* **6**:e6061 DOI 10.7717/peerj.6061.

Zhang Z, Ouyang Y, Huang Y, Wang P, Li J, He T, Liu Q. 2019c. Comprehensive bioinformatics analysis reveals potential lncRNA biomarkers for overall survival in patients with hepatocellular carcinoma: an on-line individual risk calculator based on TCGA cohort. *Cancer Cell Int* **19**:174.

Zhong X, Long Z, Wu S, Xiao M, Hu W. 2018. LncRNA-SNHG7 regulates proliferation, apoptosis and invasion of bladder cancer cells assurance guidelines. *J BUON* **23**:776–781.

Zuo S, Wei M, Zhang H, Chen A, Wu J, Wei J, Dong J. 2019. A robust six-gene prognostic signature for prediction of both disease-free and overall survival in non-small cell lung cancer. *Journal of Translational Medicine* **17**:152 DOI 10.1186/s12967-019-1899-y.