A Seven-Long Non-coding RNA Signature Improves Prognosis Prediction of Lung Adenocarcinoma: An Integrated Competing Endogenous RNA Network Analysis

Rang Li††, Kedong Han††, Dehua Xu†, Xiaolin Chen†, Shujin Lan†, Yuanjun Liao†, Shengnan Sun† and Shaoqi Rao*†

1 Institute of Medical Systems Biology, School of Public Health, Guangdong Medical University, Dongguan, China, 2 Department of Cardiology, Maoming People’s Hospital, Maoming, China

Early and precise prediction is an important way to reduce the poor prognosis of lung adenocarcinoma (LUAD) patients. Nevertheless, the widely used tumor, node, and metastasis (TNM) staging system based on anatomical information only often could not achieve adequate performance on foreseeing the prognosis of LUAD patients. This study thus aimed to examine whether the long non-coding RNAs (lncRNAs), known highly involved in the tumorigenesis of LUAD through the competing endogenous RNAs (ceRNAs) mechanism, could provide additional information to improve prognosis prediction of LUAD patients. To prove the hypothesis, a dataset consisting of both RNA sequencing data and clinical pathological data, obtained from The Cancer Genome Atlas (TCGA) database, was analyzed. Then, differentially expressed RNAs (DElncRNAs, DEMiRNAs, and DEmRNAs) were identified and a lncRNA–miRNA–mRNA ceRNA network was constructed based on those differentially expressed RNAs. Functional enrichment analysis revealed that this ceRNA network was highly enriched in some cancer-associated signaling pathways. Next, lasso-Cox model was run 1,000 times to recognize the potential survival-related combinations of the candidate IncRNAs in the ceRNA network, followed by the “best subset selection” to further optimize these lncRNA-based combinations, and a seven-lncRNA prognostic signature with the best performance was determined. Based on the median risk score, LUAD patients could be well distinguished into high-/low-risk subgroups. The Kaplan–Meier survival curve showed that LUAD patients in the high-risk group had significantly shorter overall survival than those in the low-risk group (log-rank test $P = 4.52 \times 10^{-9}$). The ROC curve indicated that the clinical genomic model including both the TNM staging system and the signature had a superior performance in predicting the patients’ overall survival compared to the clinical model with the TNM staging system only. Further stratification analysis suggested that the signature could work well in the different strata of the stage, gender, or age, rendering it to be a wide application. Finally, a ceRNA subnetwork related to the signature was extracted, demonstrating its high involvement in the
Long non-coding RNAs (lncRNAs) are defined as any ncRNA that is 200 nucleotides to 100 kb in length (Yin et al., 2018). Many studies reported that lncRNA plays an important role in the pathogenesis of cancer and has significant clinical value in prognosis and diagnosis (Knoll et al., 2015; Evans et al., 2016; Huang et al., 2017; Tripathi et al., 2018). It was also demonstrated that lncRNA can act as a “sponge” to regulate the targeted gene expression by competitively binding with miRNA (Zhang et al., 2016). This novel model of gene regulation is a part of the competing endogenous RNA (ceRNA) hypothesis, which was first proposed in 2011 (Salmena et al., 2011). ceRNAs (including lncRNA, circRNA, and mRNA) competitively bind with microRNA via sharing microRNA response elements (MREs) to weaken the inhibition effect for the target gene. The regulatory relations among lncRNAs, miRNA, and mRNA form a complex ceRNA network, and the abnormal expression of lncRNA would destroy the balance of the ceRNA network to lead to the initiation and progression of cancer (Karreth and Pandolfi, 2013).

Owing to the heterogeneity and polygenic mutation in lung cancer, a single genomic mutation is difficult to explain the various phenotypes and the variable risks of complex disease (Andreassen et al., 2014). Compared with a single gene and single factor, a biomolecular network(s) including multiple disease-related factors, which perform their dysfunctions through physical and biochemical interactions in a network (Zhao and Liu, 2019), represents various molecular relationships underlying complex diseases and depicts a clear global picture of interactions among disease-related factors (Jinawath et al., 2016). As a biomolecular network, the ceRNA network, describing post-transcriptional interactions between ceRNAs and miRNAs, had great value in prognosis, diagnosis, and therapy of cancers (Lin et al., 2018; Zhang et al., 2018; Eissa et al., 2019). In recent years, there are several successful attempts that use the ceRNA networks to identify prognostic signatures for different cancers (Hu et al., 2019; Wang et al., 2019b).

The aim of the present study was to establish a multi-lncRNA prognosis predictor. To this end, a LUAD-related lncRNA–miRNA–mRNA ceRNA network was constructed based on integrated transcriptome data from The Cancer Genome Atlas (TCGA) database. Then, by using lasso-Cox regression model, a seven-lncRNA prognostic signature was identified from the LUAD-related ceRNA network. Survival analysis and the receiver operating characteristic (ROC) curve suggested the seven-lncRNA prognostic signature is a robust and independent prognostic factor. Most importantly, our study demonstrated that the seven-lncRNA prognostic signature effectively enhanced the prognosis prediction performance over the conventional TNM staging system.
values between LUAD samples and adjacent normal tissue samples based on the edgeR package of R platform (Robinson et al., 2010). The cutoff criterion was set at $|\logFC| > 1$ and FDR < 0.05 for the screening of DE-lncRNAs, DEMiRNAs, and DE-mRNAs. Volcano plot was used to display the differentially expressed RNAs.

### Constructing the ceRNA Network

Regulatory relationships among DE-RNAs were identified by mining knowledge of several public databases. miRcode database (Jeggari et al., 2012) was used to predict the regulatory relationships between lncRNAs and miRNAs, while miRDB, miRTarBase, and TargetScan databases (Zheng et al., 2019) were used to define the regulatory relationship between miRNAs and mRNAs. According to the ceRNA theory, there should be a negative regulatory relationship between ceRNAs and miRNAs (Taulli et al., 2013). Therefore, the Pearson correlation coefficient between lncRNAs/miRNAs and miRNAs was calculated to identify negatively correlated RNA–RNA regulatory pairs ($P < 0.05$). In short, the regulatory relationships between lncRNAs/miRNAs and miRNAs were determined by three facts: (1) having a biological basis, supported by knowledge bases; (2) being a negative relationship, which agrees with the competing endogenous RNA theory; and (3) achieving the significance level in the Pearson correlation analysis based on their expression data. Finally, based on the shared miRNAs among these regulatory pairs, the lncRNA–miRNA–mRNA ceRNA network was built by connecting negative lncRNA–miRNA and miRNA–mRNA regulatory pairs. Cytoscape v3.7.1 was used for network visualization (Shannon et al., 2003).

To reveal the biological function(s) that ceRNA regulatory network involved, the Kyoto Encyclopedia of Genes and Genomes (KEGG)-based enrichment analysis was conducted to assess the ceRNA regulatory network using clusterProfiler package in R (Yu et al., 2012). The enriched KEGG pathway(s) with a FDR less than 0.05 was considered as statistically significant.

### Defining the Prognostic Signature

In order to identify the lncRNAs and optimal subset(s) related to the overall survival of LUAD patients, lasso-Cox model (R package glmnet) was run 1,000 times to recognize the potential survival-related combinations of the candidate lncRNAs in the ceRNA network, followed by the “best subset selection” (the area under the ROC curve, AUC > 0.70 with the minimal set size) to further optimize these lncRNA-based combinations (i.e., the survival-related signatures). Finally, in order to evaluate the joint effect of the best signature, a risk score was calculated based on a linear combination of the expression levels of the included lncRNAs weighted by their regression coefficients derived from the multivariate Cox regression analysis.

The risk score formula was defined as following:

$$\text{Riskscore} = \sum_{i=1}^{n} \exp \beta_i$$

Here, $n$ represents the number of lncRNAs in the model, $\exp$ represents the expression level of lncRNA $i$, and $\beta_i$ is the regression coefficient of lncRNA $i$ in multivariate Cox regression model.

### Assessing the Prognosis Value of the Newly Identified Signature

First, in order to evaluate its potential to classify the LUAD patients, according to the median risk score of the signature, LUAD patients were divided into low- and high-risk groups. The Kaplan–Meier method was used to display the difference in survival time between low-risk and high-risk LUAD groups. The statistical significance of the difference between the survival profiles of the two groups was determined by using the log-rank test. ROC curve was used to estimate its sensitivity and specificity. Second, to assess whether combining the lncRNA-based signature with the TNM stages could improve the prognosis prediction for LUAD, a clinical genomic model with the TNM stages and the lncRNA-based signature combined was constructed, and AUC was compared to the model with the TNM stages only. Third, in order to explore its applicability, a stratification analysis by the TNM stages, gender, or age was performed. Finally, to explore its biological role(s), a core ceRNA network was constructed and its functional involvements were identified by a KEGG-based enrichment analysis.

The detailed workflows of the proposed strategies for identifying and assessing the survival-related lncRNA-based signature are illustrated in Figure 1.

### RESULTS

#### Aberrantly Expressed lncRNAs, miRNAs, and mRNAs

In total, RNA-seq data for 535 LUAD samples and 59 adjacent normal samples, miRNA-seq data for 519 LUAD samples and 48 adjacent normal samples, and the corresponding clinical data for 504 LUAD patients were obtained. According to the cutoff criteria ($|\logFC| > 1$ and FDR < 0.05), 5,537 miRNAs (3,721 upregulated and 1,816 downregulated), 352 miRNAs (273 upregulated and 79 downregulated), and 3,939 lncRNAs (3,202 upregulated and 737 downregulated) were found differentially expressed (named DE-lncRNAs, DE-miRNAs, and DE-mRNAs, respectively). Their volcano plots are shown in Figure 2.

#### The ceRNA Network for LUAD

Among these DE-RNAs, 475 lncRNA–miRNA pairs between 197 lncRNAs and 39 miRNAs and 198 miRNA–mRNA pairs between 39 miRNAs and 140 mRNAs were found showing a significant negative correlation ($P < 0.05$), after excluding positively correlated pairs. Based on 39 shared miRNAs among these regulatory pairs, the lncRNA–miRNA–mRNA (ceRNA) network was established (Figure 3A). The top 15 KEGG pathways ($P$-value < 0.05) that the network was involved are shown in Figure 3B, indicating that the ceRNA network for LUAD was closely related to some cancer-associated pathways, such as
microRNAs in cancer, transcriptional misregulation in cancer, cellular senescence, cell cycle, p53 signaling pathway, small cell lung cancer, and so on.

The lncRNA-Based Prognostic Signature for LUAD

Among 1,000 lncRNA sets, constructed by using lasso-Cox regression analysis of 197 lncRNAs included in the ceRNA network for LUAD, five had AUC > 0.7. The optimal set with minimal size was selected as the lncRNA-based prognostic signature for LUAD, which was consisted of seven lncRNAs (SNHG12, DLEU7_AS1, FAM41C, FAM181A AS1, AC022148.1, CCDC13 AS1, and LINC00319). Figures 4A,B shows the convergence of the lasso-based variable selection (or called feature shrinkage) with the log of the penalty parameter lambda (λ) as well as the changes of model fitting statistics (partial likelihood deviance).

Then, a predictive model for the lncRNA set was constructed according to their lncRNA expression values and their corresponding coefficients derived from the multivariate Cox regression analysis. The risk score based on the lncRNA set was defined as following:

\[
\text{Riskscore} = (-0.2196 \times SNHG12) + (-0.14458 \times DLEU7\_AS1) + (-0.12641 \times FAM41C) + (-0.08658 \times FAM181A\_AS1) + (-0.11915 \times AC022148.1) + (-0.15204 \times CCDC13\_AS1) + (0.12557 \times LINC00319)
\]

Based on the median value of the risk score, 504 patients were subdivided into high-risk and low-risk groups. The expression patterns (heat maps) for patients of two groups (low-risk group and high-risk group) are shown in Figure 4C,
while their risk score distribution, as well as their survival, data are shown in Figure 4D, which clearly indicated that with the increase of the risk score, LUAD patients tended to have a shorter survival time. Kaplan–Meier survival analysis showed that the patients in the low-risk group had a better prognosis than those with high-risk scores \((P = 4.52e^{-9})\) (Figure 4E). ROC analysis estimated that the AUC values of the lncRNA-based prognostic signature for survival were 0.721 for the third year and 0.72 for the fifth year, respectively (Figure 4F).
Comparison Between the Clinical Genomic Model With Both the TNM Stages and the IncRNA Signature Included and the Clinical Model With the TNM Stages Only

In order to verify whether combining the IncRNA-based signature with the TNM stages could improve the prognosis prediction for LUAD, a clinical genomic model with the TNM stages and the IncRNA-based signature combined was constructed, and its area under the ROC curve (AUC) was compared to the model with the TNM stages only. As shown in Figure 5A, the AUC values for the model with the TNM stages only were 0.688 and 0.684 for 3- and 5-year survival, respectively, which were markedly lower than the estimates for the clinical genomic model with both the TNM stages and the IncRNA-based signature (0.751 and 0.782 for the two periods, respectively) (Figure 5B). A comparison of the two models for prognosis prediction over the two periods separately all demonstrated that the clinical genomic model performed significantly better than the conventional model with the TNM stage (all \( P < 0.01 \)) (Figures 5C,D). These results suggest that IncRNAs could provide additional information on the prognosis prediction for LUAD, and more importantly, this fact may render an earlier prognosis prediction for LUAD to be practical.

Applicability of the IncRNA-Based Prognostic Signature

In order to explore its applicability, a stratification analysis by the TNM stages (overall rating, divided into four ranks), gender, or age was performed. The results indicated that the IncRNA-based prognostic signature could work well in most strata of the TNM stages, gender, or age (Figures 6A–C,E–H), i.e., having a good capacity in separating the LUAD patients into high-risk and low-risk groups. For the TNM stages, stage IV was the only exception where no statistical significance was found between the low- and high-risk groups defined by the IncRNA-based signature (Figure 6D), which might be due to the small sample size (\( n = 26 \)). These results suggested that the IncRNA-based prognostic signature was largely independent of tumor stage, gender, and age, which was agreed well with the multivariate Cox regression analysis where the IncRNA-based prognostic signature remained to be the most significant factor (\( P < 0.001 \)) even after adjusting for several demographic and clinical factors (Figure 6I).

The ceRNA Subnetwork Related to the IncRNA-Based Signature

On basis of the seven IncRNAs contained in the prognostic signature for LUAD, a core network (Figure 7A) was extracted from the primary ceRNA network. KEGG-based functional enrichment analysis demonstrated that this IncRNA-mediated ceRNA subnetwork was highly involved in several cancer-associated signaling pathways (Figure 7B), which implicated that the seven IncRNAs play vital roles in the tumorigenesis mechanism of LUAD via regulating related gene expressions by competitively sponging several miRNAs.

DISCUSSION

As the most common malignancy, LUAD had an unfavorable 5-year survival rate at an advanced stage. Early detection and diagnosis was an important way to improve the LUAD patient’s prognosis. Although it had been confirmed as an effective prognosis predictor for LUAD patients (Carter et al., 2018), the TNM staging system, which was founded on the anatomical information only, could not perfectly perform prognosis prediction. An increasing amount of evidence demonstrated that genetic disorders and alterations were of significance in tumorigenesis and the progression of LUAD, suggesting that molecular markers had great value in the prediction of overall survival of LUAD patients (Trimarchi et al., 2014; Timmer et al., 2016).

With the development of high-throughput sequencing technology, the roles of long non-coding RNAs (IncRNAs) in human cancers had received more and more attention. Previous studies demonstrated that IncRNAs play an important role in tumor proliferation, migration, and invasion (Zhang et al., 2018) and have potential value in applications to early...
FIGURE 6 | Applicability of the lncRNA-based prognostic signature. (A–D) survival curves of the LUAD patients in different risk groups, stratified by four tumor stages [stage I (A), stage II (B), stage III (C), and stage IV (D)]. (E,F) Survival curves of the LUAD patients in different risk groups, stratified by two ages [age ≤60 (E) and age <60 (F)]. (G,H) Survival curves of the LUAD patients in different risk groups, stratified by gender [male (G) and female (H)]. (I) Multivariate Cox regression with the lncRNA-based signature (risk score) and all available demographic and clinical factors included.

FIGURE 7 | The ceRNA subnetwork mediated by the lncRNA-based signature. (A) The ceRNA subnetwork depicting the relationships between seven lncRNAs and their counterparts. Diamonds represent lncRNAs, triangles miRNAs, and inverse-triangles mRNAs. The red color depicts upregulation and green downregulation. (B) The top 15 significantly enriched KEGG pathways.

prognosis and diagnosis for cancers (Takahashi et al., 2014). Zhai et al. (2017) demonstrated that the lower expression of lncRNA-SARCIC influenced downstream genes such as K-RAS, MMP-13, AKT, and P-ERK expression by suppressing miR-143-3p expression, which could enhance RCC cell invasion, migration, and proliferation. Wu et al. (2017) first reported that lncRNA CASC9, as an oncogene, promoted ESCC cell growth by negatively regulating PDCD4 expression via recruiting EZH2, which could be a potential diagnosis and prognosis biomarker for ESCC. All these facts highlight the importance of the large
number of lncRNAs to be the molecular biomarkers for early prognosis prediction or early diagnosis of cancers.

Compared with the coding RNAs, it was more complex to study the functional meanings of lncRNAs. The ceRNA hypothesis provided a new solution for achieving better functional studies of lncRNAs. In the ceRNA theory, lncRNAs regulate the expression of the targeted genes by competitively absorbing miRNAs at the post-transcriptional level, forming a huge ceRNA regulatory network (Calloni and Bonatto, 2019). In most scenarios, "communications" between ceRNAs and miRNAs were in dynamic balance (Cai and Wan, 2018). However, an abnormal expression of lncRNA destroyed the balance of the ceRNA network, which was closely related to tumorigenesis (Shao et al., 2015). In this study, a total of 3,939 DE-lncRNAs, 352 DE-miRNA, and 5,537 DE-mRNAs were identified. According to the ceRNA theory, the negatively correlated RNA–RNA regulatory pairs were built among DE-lncRNAs, DE-miRNA, and DE-mRNA. Subsequently, a lncRNA–miRNA–mRNA ceRNA network including 197 DE-lncRNAs, 39 DE-miRNAs, and 140 DE-mRNAs was constructed by connecting the negatively correlated RNA–RNA regulatory pairs. Further functional enrichment analysis showed that the ceRNA network was mainly involved in some cancer-related pathways including "microRNAs in cancer," "transcriptional misregulation in cancer," "cell cycle," "p53 signaling pathway," "colorectal cancer," "small cell lung cancer," etc., which was not surprising at all to us because more and more evidence indicated that at the molecular levels various cancers were interconnected. In brief, the ceRNA regulatory network including complex molecular regulatory relationships not only had potential value to mine prognosis-related biomarkers (Bai et al., 2019) but also provided a new avenue to broaden our knowledge on massive lncRNAs and their functional involvements in the pathogenic mechanisms for cancers like LUAD.

Most of the preceding studies focused on single lncRNAs related to lung cancer. Nie et al. (2016) found that LncRNA-UCA1 upregulated ERBB4 by sponging miR-193a-3p to exert oncogenic functions. Guan et al. (2019) demonstrated that LINCC00673-v4 enhanced cancer cell invasion, migration, and metastasis by overactivating WNT/β-catenin signaling and could be a candidate for the therapeutic target of LUAD patients. Nevertheless, as a complex disease, LUAD was thought to be a series of biological cascades resulting from the perturbations of intracellular and intercellular elements (Fiscon et al., 2018), and it was impossible to have a global picture about the sophisticated pathogenic mechanism of LUAD by studying only a single biomarker. Therefore, compared with a single biomarker, a prognostic signature integrating multiple biomarkers could achieve more power in prognosis prediction for LUAD (Kratz et al., 2019). This study demonstrated that the newly identified lncRNA-based signature with seven lncRNAs could provide >9% improvement in prognosis prediction over various periods for LUAD and was deemed to a robust complement for the conventional TNM staging system.

Among the seven lncRNAs containing the prognostic signature, four (AC022148.1, DLEU7_AS1, LINCC00319, and SNHG12) were found to be involved in tumorigenesis, migration, and metastasis in cancers. Qi et al. (2019) demonstrated that a seven-lncRNA prognostic model including AC022148.1 was a robust indicator to assess the prognosis risk of lung squamous cell carcinoma patients. Liu et al. (2018) found that DLEU7_AS1, as an adverse prognosis factor, was closely associated with colorectal cancer (CRC) staging, lymph node metastasis, and distant metastasis and may regulate the Wnt/β-catenin pathway to promote the occurrence and development of CRC. Zhou et al. (2017) demonstrated that LINCC00319 strengthened proliferation and invasion of lung cancer cells by downregulating the expression of miR-32 and upregulating the expression levels of miR-32 target genes. Wang et al. (2019a) revealed that knockdown of SNHG12 inhibited migration and invasion of NSCLC cells via the Slug/zinc finger E-box-binding homeobox 2 EMT signaling pathway by upregulating the expression of miR-218 and could be a potential prognostic marker and therapeutic target for NSCLC. Up to date, there is dearth of information about the roles of the remaining three lncRNAs in tumors, waiting for further studies to clarify.

Compared with several previous studies, our approach has the following improvements. First, for constructing the ceRNA network, we used only negatively correlated lncRNAs/miRNAs–miRNAs regulatory pairs, which fits well the definition of the ceRNA network compared to the previous method without this distinction (Li et al., 2018; Fan et al., 2020; Wu et al., 2020). Second, we applied 1,000 rounds of lasso-Cox regression model fittings to identify the optimal lncRNA-based set, which is deemed to be more robust than the conventional Cox model utilized in a previous study (Zheng et al., 2017). Third, consequently, our study achieved better performance on predicting the survival of the LUAD patients by using a seven-lncRNA-based signature (AUC = 0.72 and 0.721 for 3- and 5-year survival, respectively) than a previous study by (Yao et al., 2020), who identified an eight-lncRNA signature for LUAD (AUC = 0.702 and 0.671 for 3- and 5-year survival). Up to date, only the clinical genomic model built by the present study has achieved adequate improvement over the conventional TNM staging system (P = 0.003961 and 9.16% for the third year; P = 0.002448 and 14.33% for the fifth year), compared to a previous similar study (Zheng et al., 2017), who reported a non-significant improvement (P > 0.05 and 4.24% for the fifth year).

In conclusion, we applied an integrated ceRNA network analysis to identify a lncRNA-based signature for predicting the prognosis of LUAD patients. The established molecular signature with seven lncRNAs, derived from the ceRNA network, was demonstrated to be a robust and independent factor for the survival prediction of LUAD patients and, hence, could be an important complement for the conventional TNM staging system.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ supplementary material.
AUTHOR CONTRIBUTIONS
SR, RL, and KH conceived and designed the study. RL, KH, SR, YL, SS, and SL performed data analysis. RL, XC, and DX contributed in the software and programming. RL, SR, KH, XC, and DX wrote, reviewed, and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING
This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 31071166 and 81373085 to SR) and the High-level Hospital Construction Research Project of Maoming People’s Hospital (ZX2020013 to KH).

REFERENCES
Andreasen, O. A., Zuber, V., Thompson, W. K., Schork, A. J., Bettella, F., Djurovic, S., et al. (2014). Shared common variants in prostate cancer and blood lipids. *Int. J. Epidemiol.* 43, 1205–1214. doi: 10.1093/ije/dyu090
Bai, Y., Long, J., Liu, Z., Lin, J., Huang, H., Wang, D., et al. (2019). Comprehensive analysis of a ceRNA network reveals potential prognostic cytoplasmic IncRNAs involved in HCC progression. *J. Cell. Physiol.* 234, 18837–18848.doi: 10.1002/jcp.28522
Bai, D. (2019). TNM in non-small cell lung cancer: A staging system for all oncologists or just for surgeons? *Ann. Transl. Med.* 7:5103. doi: 10.21037/atm.2019.04.84
Cai, Y., and Wan, J. (2018). Competing endogenous RNA regulations in neurodegenerative disorders: current challenges and emerging insights. *Front. Mol. Neurosci.* 11:370. doi: 10.3389/fnmol.2018.00370
Calloni, R., and Bonatto, D. (2019). Characteristics of the competition among IncRNAs for the binding of shared miRNAs. *Eur. J. Cell Biol.* 98, 94–102.doi: 10.1016/j.ejcb.2019.04.001
Carter, B. W., Lichtenberger, J. R., Benveniste, M. K., de Groot, P. M., Wu, C. C., et al. (2014). Shared common variants in prostate cancer and blood lipids. *Ann. Transl. Med.* 2:S103. doi: 10.21037/atm.2014.01.017
Fan, F., Ping, Y., Yang, L., Duan, X., Resegofetse, M. N., Li, B., et al. (2020). Genomic-wide analysis of dysregulated long non-coding RNAs/microRNAs/mRNAs in metastasis of lung adenocarcinoma. *J. Transl. Med.* 18:372. doi: 10.1186/s12967-018-1732-z
Li, P., Wen, D. Y., Li, Q., He, Y., Yang, H., and Chen, G. (2018). Genome-wide analysis of prognostic IncRNAs, miRNAs, and mRNAs forming a competing endogenous RNA network in hepatocellular carcinoma. *Cell. Physiol. Biochem.* 48, 1953–1967.doi: 10.1007/s00294-018-04925-9
Liu, X. B., Han, C., and Sun, C. Z. (2018). Long non-coding RNA DLEU7-AS1 promotes the occurrence and development of colorectal cancer via Wnt/beta-catenin pathway. *Eur. Rev. Med. Pharmacol. Sci.* 22, 110–117. doi: 10.26355/errv.201801.14107
Pontius, L. N., Oyekunle, T. O., Thomas, S. M., Stang, M. T., Scheri, R. P., Roman, S. A., et al. (2017). Projecting survival in papillary thyroid cancer: a comparison of the seventh and eighth editions of the American joint committee on cancer staging system in breast cancer. *J. Clin. Oncol.* 35, 1119–1127.doi: 10.1200/JCO.2014.57.2958
Nie, W., Ge, H. J., Yang, X. Q., Sun, X., Huang, H., Tao, X., et al. (2016). lncRNA-UCA1 exerts oncogenic functions in non-small cell lung cancer by targeting miR-193a-3p. *Cancer Lett.* 371, 99–106.doi: 10.1016/j.canlet.2015.11.024
Pontius, L. N., Oyekunle, T. O., Thomas, S. M., Stang, M. T., Scheri, R. P., Roman, S. A., et al. (2017). Projecting survival in papillary thyroid cancer: a comparison of the seventh and eighth editions of the American joint committee on cancer/union for international cancer control staging systems in two contemporary national patient cohorts. *Thyroid* 27, 1408–1416.doi: 10.1089/thy.2017.0306
Qi, L., Zhang, T., Yao, Y., Zhuang, J., Liu, C., Liu, R., et al. (2019). Identification of IncRNAs associated with lung squamous cell carcinoma prognosis in the competitive endogenous RNA network. *PeerJ* 7:e7727.doi: 10.7717/peerj.7727
Rajer, B. (2014). Pollution in the working place and social status: co-factors in lung cancer carcinogenesis. *Lung Cancer* 85, 346–350.doi: 10.1016/j.lungcan.2014.06.012
Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.doi: 10.1093/bioinformatics/btp616
Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504. doi: 10.1101/gr.129393

Shao, T., Wu, A., Chen, J., Chen, H., Lu, J., Bai, J., et al. (2015). Identification of module biomarkers from the dysregulated ceRNA-centric network interaction network in lung adenocarcinoma. Mol. Biol. Sys. 11, 3048–3058. doi: 10.1039/c5mb00364d

Srid, M., Coebergh, V. D. B. R., van de Werken, H., van Riet, J., van Galen, A., de Weerd, V., et al. (2018). Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. BMC Bioinformatics 19:236. doi: 10.1186/s12859-018-2246-7

Takahashi, K., Yan, I., Haga, H., and Patel, T. (2014). Long noncoding RNA in liver diseases. Hepatology 60, 744–753. doi: 10.1002/hep.27043

Taullli, R., Loretelli, C., and Pandolfi, P. P. (2013). From pseudo-ceRNAs to circ-ceRNAs: a tale of cross-talk and competition. Nat. Struct. Mol. Biol. 20, 541–544. doi: 10.1038/nsmb.2580

Timmer, M. R., Martinez, P., Lau, C. T., Westra, W. M., Calpe, S., Rygiel, A. M., et al. (2016). Derivation of genetic biomarkers for cancer risk stratification in Barrett's oesophagus: a prospective cohort study. Gut 65, 1602–1610. doi: 10.1136/gutjnl-2015-309642

Tripathi, M. K., Doxtater, K., Keramatnia, F., Zacheaus, C., Yallapu, M. M., Jaggi, M., et al. (2018). Role of IncRNAs in ovarian cancer: defining new biomarkers for therapeutic purposes. Drug Discov. Today 23, 1635–1643. doi: 10.1016/j.drudis.2018.04.010

Wakeam, E., Acuna, S. A., Leigl, N. R., Giuliani, M. E., Finlayson, S., Varghese, T. K., et al. (2017). Surgery versus chemotherapy and radiotherapy for early and locally advanced small cell lung cancer: a propensity-matched analysis of survival. Lung Cancer 109, 78–88. doi: 10.1016/j.lungcan.2017.04.021

Wang, Y., Liang, S., Yu, Y., Shi, Y., and Zheng, H. (2019a). Knockdown of SNHG12 suppresses tumor metastasis and epithelial-mesenchymal transition via the Slug/ZEB2 signaling pathway by targeting miR-218 in NSCLC. Oncol. Lett. 17, 2356–2364. doi: 10.3892/ol.2018.9880

Wang, Y., Liu, X., Guan, G., Xiao, Z., Zhao, W., and Zhuang, M. (2019b). Identification of a five-pseudogene signature for predicting survival and its potential conflict of interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Copyright © 2021 Li, Han, Xu, Chen, Lan, Liao, Sun and Rao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.