Characterization of lncRNA-Associated ceRNA Network to Reveal Potential Prognostic Biomarkers in Lung Adenocarcinoma

*Yang Wang, Ruyi He and Lixin Ma**

State Key Laboratory of Biocatalysis and Enzyme Engineering, Hubei Key Laboratory of Industrial Biotechnology, Hubei University, Wuhan, China

Lung adenocarcinoma (LUAD) is one of the most fatal malignant tumors harmful to human health. The complexity and behavior characteristics of long-non-coding RNA (lncRNA)-associated competing endogenous RNA (ceRNA) network in LUAD patients are still unclear. The purpose of this study was to elucidate the regulatory networks of dysregulated RNAs, view, and identify potential prognosis signatures involved in LUAD. The expression profiles of mRNAs, lncRNAs, and miRNAs were obtained from the TCGA database. In total, 2078 DEmRNAs, 257 DElncRNAs, and 101 DEmiRNAs were sorted out. A PPI network including 45 DEmRNAs was constructed. Ten hub genes in the PPI network associated with cell cycle-related pathways were identified and they played key roles in regulating cell proliferation. A total of three DEmiRNAs, seven DElncRNAs, and six DEmRNAs were enrolled in the ceRNA network. Except for certain genes without any published study reports, all the genes in the ceRNA network played an essential role in controlling tumor cell proliferation and were associated with prognosis in LUAD. Finally, based on step regression and Cox regression survival analysis, we identified four candidate biomarkers, including miR490, miR1293, LINC01740, and IGF2BP1, and established a risk model based on the four genes. Our study provided a global view and systematic dissection of the lncRNA-associated ceRNA network, and the identified four genes might be novel important prognostic factors involved in LUAD pathogenesis.

Keywords: lung adenocarcinoma, signature, ceRNA, lncRNA, prognosis

INTRODUCTION

Despite medical advances, lung cancer remains the leading cause of cancer deaths. Lung cancer is usually recognized late in its natural history and have a poor prognosis, with an overall 5-year survival rate of 10–15% (Cagle et al., 2013). The recognition of histologic subtypes of non-small cell lung carcinoma (NSCLC), namely, adenocarcinoma, squamous cell carcinoma, and large cell lung carcinoma as the most frequent subtypes, has become important as a determinant of therapy in this disease (Kerr et al., 2014). In addition, in recent years, the identification of molecular abnormalities in a large proportion of patients with lung cancer has allowed the emergence of personalized targeted therapies and has opened new horizons and created new expectations for these patients (Ezeife and Leighl, 2018). The use of predictive biomarkers to identify tumors that

could respond to targeted therapies has meant a change in the paradigm of lung cancer diagnosis (Majeed and Amir, 2018).

Currently, the rapid advancement of high-throughput technologies offers great opportunities for biomarker identification (Yu et al., 2018). Non-coding RNAs as biomarker and therapeutic targets play a significant role in human disease (Zhou et al., 2018b,c). Among which, long-non-coding RNA (lncRNAs) are a class of RNA molecules with more than 200 nucleotides in length and have no evident open reading frames (Fatica and Bozzoni, 2014). These long molecules are dysregulated among cancers (Yan et al., 2015) and play key roles in gene regulation and carcinogenesis, including proliferation, survival, migration, and genomic stability (Gutschner et al., 2013; Castro-Oropeza et al., 2018). It is believed that the clinical value of lncRNA is not confined to candidate biomarkers for diagnostic and prognostic purposes (Shi et al., 2018).

In Salmena et al. (2011) put forward a competing endogenous RNA (ceRNA) hypothesis. Subsequently, several studies also mentioned that there is an interplay between lncRNAs and miRNAs during the tumorigenic process, among which lncRNAs serve as molecular sponges for miRNAs (Liz and Esteller, 2016). For example, KCNQ1OT1 promotes cell proliferation and autophagy and inhibits cell apoptosis via regulating miR204-5p/ATG3 axis, providing a promising target for NSCLC therapy (Kang et al., 2019). Guo et al. reported that LINC00173 up-regulated Etk through functioning as a ceRNA by "sponging" miRNA-218 and led to the up-regulation of GSKIP and NDRG1 in small cell lung cancer (Zeng et al., 2019). LncRNA AGAP2-AS1 up-regulates ANXA11 expression by sponging miR16-5p and promotes proliferation and metastasis in hepatocellular carcinoma (Liu et al., 2019). Thus, the discovery of lncRNA–miRNA–mRNA networks may lead to a more comprehensive understanding of the etiology and metastasis mechanism of cancer. However, the complexity and behavior of lncRNA-associated ceRNA network remain poorly characterized in lung adenocarcinoma (LUAD).

In this study, by comprehensively integrating gene and miRNA expression data of LUAD, the LUAD-related lncRNA–miRNA–mRNA competitive network was established. We analyzed and predicted the functions of ceRNA and PPI networks and established a Cox regression model to predict the overall survival of patients with lung cancer. Finally, four predictive genes were identified, including LINC01740, mir1293, mir490, and IGF2BP1, which could contribute to LUAD. This study will contribute to understanding the molecular mechanism and provide new therapeutic targets for LUAD.

MATERIALS AND METHODS

Data Preparation and Differentially Expressed Gene Analysis

All primitive data of LUAD from The Cancer Genome Atlas (TCGA) database[1] were download through GDC Data Transfer Tool, including RNA-seq and miRNA-seq of

[1] https://portal.gdc.cancer.gov/

Transcriptome profiling and Clinical data. EdgeR package (3.3.3 version) (Robinson et al., 2010) in R software was used to analyze and identify differentially expressed RNAs (DERNAs, including DEmRNAs and DElncRNAs) and differentially expressed microRNA (DEmiRNAs) with the thresholds of | log2FoldChange| > 2.0 and FDR (adjusted p value) < 0.01. Then, biomart in R package was used to annotate DEmRNAs and DElncRNAs. The heatmap and volcano plot were constructed by the ggplot2 package in R software (Zhou et al., 2017).

Functional Enrichment Analysis

clusterProfiler (Yu et al., 2012) package in R was used to make the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) enrichment analysis, including biological process (BP), the cellular component (CC), and molecular function (MF). Pathview (Luo and Brouwer, 2013) and enrichplot packages (Ito and Murphy, 2013) were used to visualize the enrichment results. A significance level of adjusted $p < 0.05$ was set as the cutoff criteria.

Protein–Protein Interaction Analysis

The DEmRNAs were enrolled in a protein–protein interaction (PPI) network through the STRING (version 11.0) database[2] with a confidence score >0.9. Furthermore, genes with degree ≥25 were selected as hub genes, and we focused the interaction types among proteins only on physical interaction and co-expression (Sun et al., 2019). Subsequently, GO and KEGG analyses of the PPI network modules were carried out using clusterProfiler package in R.
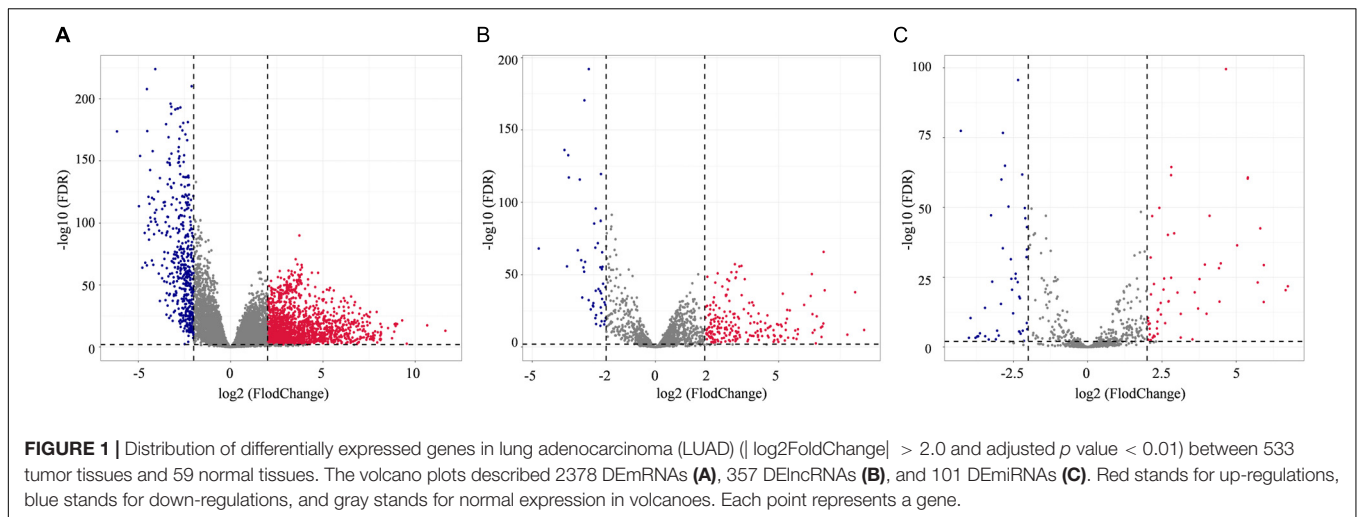
Construction of the ceRNA Network

According to the hypothesis of ceRNA, a lncRNA–miRNA–mRNA network was constructed (Zhou et al., 2018a). Relevant miRNA-target data were obtained from the miRTarBase, and the support types of targeting were only focused on experiments, including luciferase reporter assay, Western blot, Northern blot, or qRT-PCR. Only the miRNA targets that were differentially expressed between tumor and normal tissue were considered for the next analysis step. Furthermore, the candidate DElnRNA–DEmiRNA interactions were selected based on miRcode database and the following model:

$$Y_{target} =$$

$$\beta_0 + \beta_1 * miRNA + \beta_2 * lncRNA + \beta_3 * miRNA * lncRNA + \varepsilon \tag{1}$$

where $miRNA$, $lncRNA$, and Y_{target} are the gene expression of miRNA, lncRNA, and miRNA targets, respectively. β_1 and β_2 represent the effect of miRNA and lncRNA, respectively, on target by themselves alone (main effects), while β_3 represents the effect of miRNA–lncRNA interaction. If a lncRNA and miRNA interaction has effects on target expression outcomes, we expect β_3 to be non-zero. Here, all the miRNAs and lncRNAs and miRNA targets should be differentially expressed between tumor and normal tissue.

[2] https://string-db.org/

FIGURE 1 | Distribution of differentially expressed genes in lung adenocarcinoma (LUAD) (| log2FoldChange| > 2.0 and adjusted p value < 0.01) between 533 tumor tissues and 59 normal tissues. The volcano plots described 2378 DEmRNAs **(A)**, 357 DElncRNAs **(B)**, and 101 DEmiRNAs **(C)**. Red stands for up-regulations, blue stands for down-regulations, and gray stands for normal expression in volcanoes. Each point represents a gene.

Biomarkers Screening and Validation

The status and survival time of LUAD patients were extracted from the TCGA clinical dataset. Subsequently, the DEmRNAs, DElncRNAs, and DEmiRNAs identified in ceRNAs were selected for screening biomarkers. We used univariate Cox regression to screen prognostic factors (p < 0.05), and those prognostic factors whose expression levels were significantly relevant to patients' overall survival (p < 0.05) were selected as primitive biomarkers (Zhou et al., 2018b; Bao et al., 2019).

Cox Risk Regression Establishment and Validation

The lncRNAs, mRNAs, and miRNAs raw data were transformed and normalized in a log2[cpm(x) + 1] manner. Univariate cox regression was used to select prognosis-associated genes (p < 0.05) (Zhou et al., 2018a). Subsequently, we performed Cox regression analysis combined with stepwise regression to establish a Cox risk model (Zhou et al., 2018a). Finally, a validation set and Kaplan–Meier survival curves along with a logrank p test were applied to validate its accuracy (Zhou et al., 2017; Sun et al., 2019).

RESULTS

Identification of Differentially Expressed Genes

RNA expression profiles and corresponding clinical data of 533 cohort LUAD patients and 59 normal controls were downloaded from the TCGA database. Meanwhile, miRNA-seq data corresponding to 561 patients' clinical information, including 515 cohort LUAD patients and 46 normal controls, were obtained from TCGA. In total, 60,483 transcripts and 1046 miRNAs were obtained. With the cutoff criteria unified, CPM(gene) > 1, rowSum(CPM) ≥ 2, 32,495 transcripts and 613 miRNAs were selected for the differentially expressed analysis. After filtering, 5624 DERNAs and 673 DEmiRNAs were identified

with the thresholds of | log2FoldChange| > 2.0 and FDR (adjusted p value) < 0.01.

In total, 2078 DEmRNAs (1612 up-regulated and 466 down-regulated, **Figure 1A**), 257 DElncRNAs (209 up-regulated and 48 down-regulated, **Figure 1B**), and 101 DEmiRNAs (56 up-regulated and 45 down-regulated, **Figure 1C**) were sorted out.

Functional Analysis of DERNAs

Gene ontology and KEGG enrichment analyses were used to explore the potential function of DERNAs. Ten biological pathways were highly enriched within cutoff p value < 0.05. Among them, 12% DERNAs were enriched in GPCR ligand binding process, and 9.5% DERNAs were enriched in Class A/1 (Rhodopsin-like receptors) pathway, and 6.7% DERNAs were enriched in peptide ligand binding receptor pathways (**Figure 2A**). Detailed information of these enriched pathways and associated genes is summarized in **Table 1**. The GO functional enrichment analysis results of DERNAs including MF, CC, and BP were described in **Figures 2B–D**. The results show that the genes mainly focused on receptor ligand activity function, extracellular matrix, and morphogenesis of an epithelium process.

Furthermore, KEGG pathway enrichment analysis results demonstrated that the most significantly enriched pathways were neuroactive ligand–receptor interaction, alcoholism, and systemic lupus erythematosus pathways (**Figure 2E**). The pathway–pathway interaction network (PPIN) was constructed based on the DERNAs enriched in same pathway (**Figure 2F**). Four pathways were identified in the PPIN, including alcoholism, maturity onset diabetes of the young, neuroactive ligand–receptor interaction, and systemic lupus erythematosus pathway. We noticed that, all the DERNAs enriched in systemic lupus erythematosus and alcoholism pathways were up-regulated, except for gene 2354, whose gene symbol is "FOSB." Gene annotation of FOSB shows that it was a proto-oncogene, and it has been implicated as regulators of cell proliferation, differentiation, and transformation. Similarly,
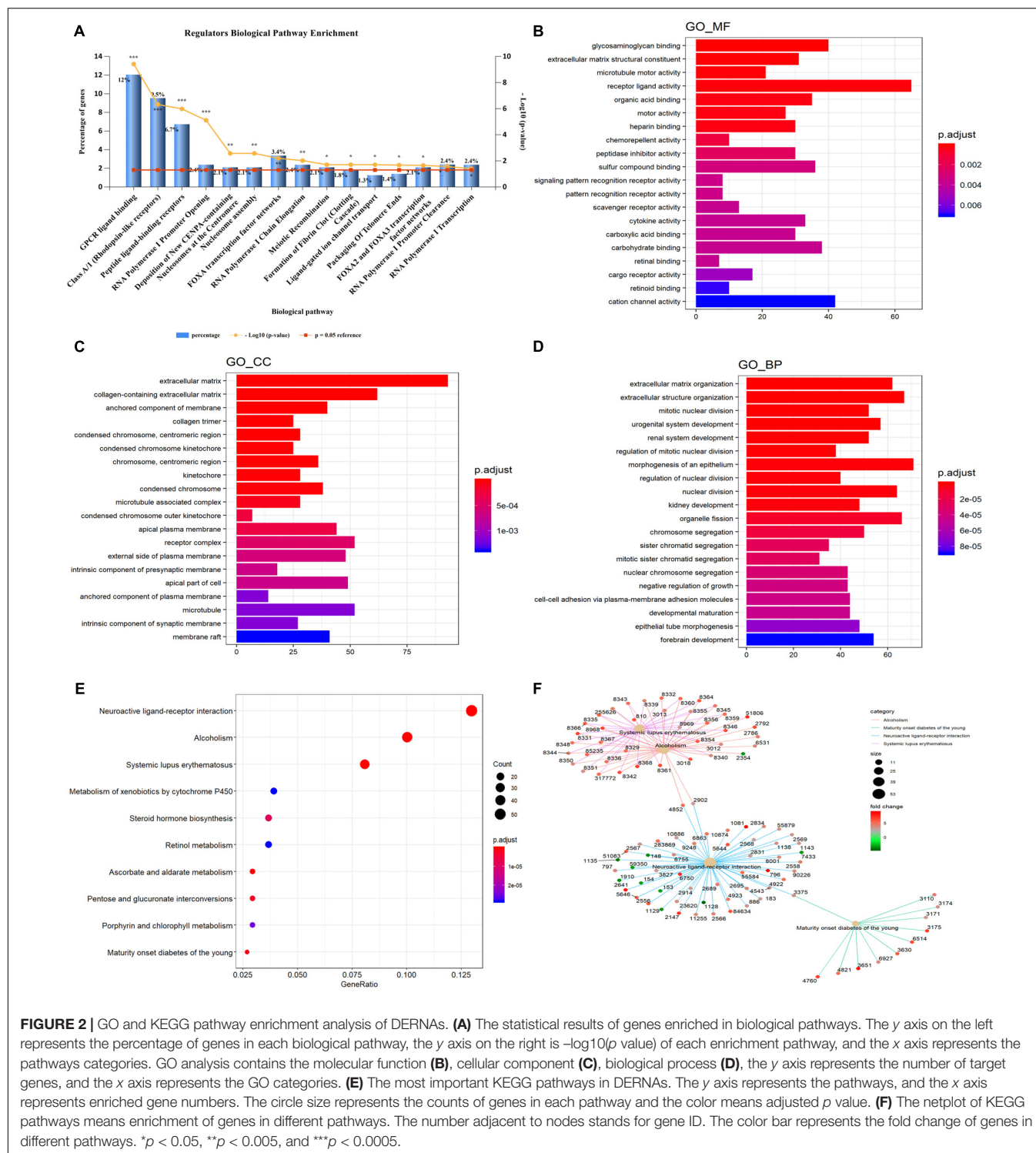
FIGURE 2 | GO and KEGG pathway enrichment analysis of DERNAs. **(A)** The statistical results of genes enriched in biological pathways. The y axis on the left represents the percentage of genes in each biological pathway, the y axis on the right is –log10(p value) of each enrichment pathway, and the x axis represents the pathways categories. GO analysis contains the molecular function **(B)**, cellular component **(C)**, biological process **(D)**, the y axis represents the number of target genes, and the x axis represents the GO categories. **(E)** The most important KEGG pathways in DERNAs. The y axis represents the pathways, and the x axis represents enriched gene numbers. The circle size represents the counts of genes in each pathway and the color means adjusted p value. **(F)** The netplot of KEGG pathways means enrichment of genes in different pathways. The number adjacent to nodes stands for gene ID. The color bar represents the fold change of genes in different pathways. *p < 0.05, **p < 0.005, and ***p < 0.0005.

all the genes enriched in maturity onset diabetes of young pathway were all up-regulated in LUAD. Furthermore, the results showed that gene 4852 and 2092 were both enriched in alcoholism pathway and neuroactive ligand–receptor interaction pathway, and played vital roles in connecting the two pathways.

PPI Network Analysis

A total of 55 proteins and 453 edges, including 45 DEmRNAs, were selected in the PPI network. A total of 10 hub genes, including CDK1, TOP2A, PBK, CDCA8, CDC20, KIF20A, DLGAP5, NDC80, NCAPG, and CCNA2, were selected from the PPI network with degree ≥25 and combined score

TABLE 1 | Enriched biological pathways and associated genes.

Biological pathway	No. of genes in DERNAs	No. of genes in background dataset	Percentage of genes	Fold enrichment	p value	p-adjusted (Bonferroni)
GPCR ligand binding	86	336	12.0	2.2	5.05E-14	8.42E-11
Class A/1 (rhodopsin-like receptors)	68	274	9.5	2.1	1.26E-10	2.1E-07
Peptide ligand-binding receptors	48	167	6.7	2.5	4.13E-10	6.89E-07
RNA polymerase I promoter opening	17	31	2.3	4.8	4.09E-09	6.83E-06
Deposition of new CENPA-containing nucleosomes at the centromere	15	35	2.1	3.7	2.11E-06	0.003
Nucleosome assembly	15	35	2.1	3.7	2.11E-06	0.003
FOXA transcription factor networks	24	81	3.3	2.6	5.93E-06	0.009
RNA polymerase I chain elongation	17	48	2.3	3.1	1.00E-05	0.016
Meiotic recombination	15	42	2.1	3.1	2.99E-05	0.049
Formation of fibrin clot (clotting cascade)	13	33	1.8	3.4	3.05E-05	0.051
Ligand-gated ion channel transport	9	17	1.2	4.6	3.11E-05	0.051
Packaging of telomere ends	10	21	1.4	4.1	3.65E-05	0.060
FOXA2 and FOXA3 transcription factor networks	15	43	2.1	3.0	4.12E-05	0.068
RNA polymerase I promoter clearance	17	54	2.3	2.7	5.79E-05	0.096
RNA polymerase I transcription	17	56	2.3	2.6	9.67E-05	0.161

>0.9 (**Figure 3A**). Furthermore, the association among these interacted proteins should be physical interaction or co-expressed with each other (**Figure 3B**). We noticed that eight RNA expression levels were significantly associated with overall survival outcomes except for CDCA8 and CDC20 (**Figures 3C,D**). Pathway enrichment analysis results of the 10 hub genes are summarized in **Table 2**.

Construction of the ceRNA Network in LUAD

A total of seven DElncRNAs, six DEmRNAs, and three DEmiRNAs were enrolled in the ceRNA network (**Figure 4**). miRTarBase was used to predict the miRNA–mRNA pairs (**Table 3**). We only focused on those miRNA–mRNA pairs whose interaction evidence was validated by experiments, including luciferase reporter assay, Western blot, Northern blot, or qRT-PCR.
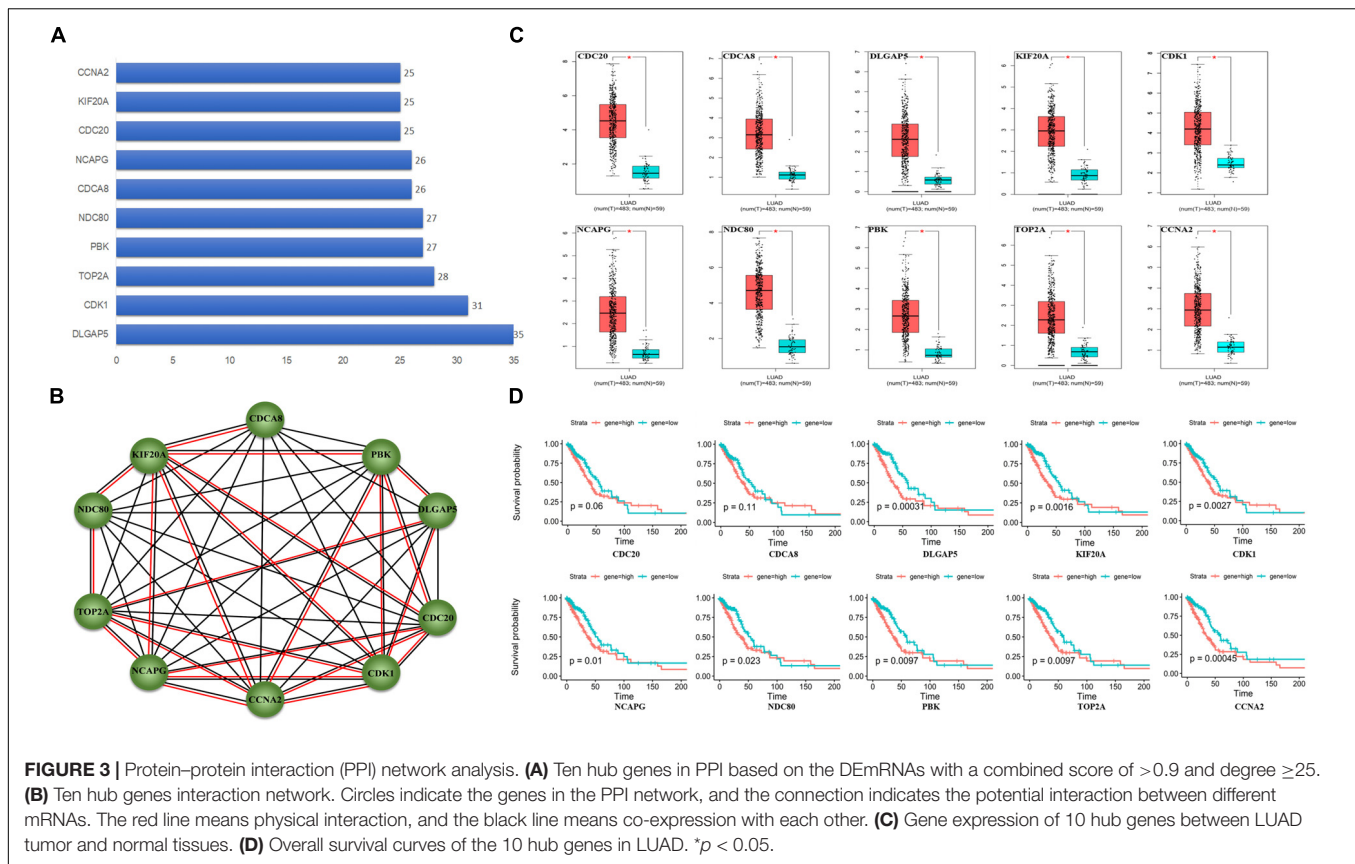
Then, we employed a simple linear regression model combined with miRcode database to predict the potential miRNA target by DElncRNAs [see Methods 2.4 model (1)]. In the model, we specified that the input of lncRNAs should be (i) differentially expressed between tumor and normal tissues; (ii) lncRNA expression is associated with overall survival outcomes (logrank p value < 0.05). Finally, 7 of 53 DElncRNAs, 6 of 340 DEmRNAs, and 3 of 9 DEmiRNAs formed the ceRNA network. Detailed information about their expression

and association with overall survival outcomes is listed in **Table 4**.

Screen Biomarkers and Construction Risk Model

Three DEmiRNAs, seven DElncRNAs, and six DEmRNAs in the ceRNA network were selected as candidate biomarkers for the following step analysis. Subsequently, combined univariate Cox regression with a logrank test analysis with p value < 0.05 and 12 variables (miR1293, miR196b, miR490, C20orf197, SCAT1, C11orf44, MALAT1, VPS9D-AS1, LINC02473, LINC01740, HOXB7, and IGF2BP1) were identified. Furthermore, a stepwise regression was performed according to the 12 variables. Consequently, four variables including miR490, miR1293, LINC01740, and IGF2BP1 were harvested in the Cox regression. Risk score = −0.455*miR490 + 0.037*miR1293 + 0.034*LINC01740 + 0.005IGF2BP1 (**Figure 5A**).

Afterward, the LUAD patients were divided into two groups based on the median value of Cox regression model. The distribution of the risk score along with the corresponding survival data and the four protective gene expression demonstrated that the high-risk LUAD patients tended to experience shorter survival time, and low-risk LUAD patients were opposite (**Figure 5B**). Results show that miR490 and LINC01740 in the high-risk group expression level were lower

FIGURE 3 | Protein–protein interaction (PPI) network analysis. **(A)** Ten hub genes in PPI based on the DEmRNAs with a combined score of >0.9 and degree ≥25. **(B)** Ten hub genes interaction network. Circles indicate the genes in the PPI network, and the connection indicates the potential interaction between different mRNAs. The red line means physical interaction, and the black line means co-expression with each other. **(C)** Gene expression of 10 hub genes between LUAD tumor and normal tissues. **(D)** Overall survival curves of the 10 hub genes in LUAD. *$p < 0.05$.

than that in the low-risk group; meanwhile, miR1293 and IGF2BP1 were opposite (**Figures 5C,D**).

DISCUSSION

In this study, a total of 2078 DEmRNAs, 257 DElncRNAs, and 101 DEmiRNAs were identified. GO analysis revealed that the function of DERNAs is mainly associated with receptor ligand activity, ligand-gated ion channel transport, morphogenesis of an epithelium process, and cell–cell adhesion, which play vital roles in tumorigenesis (Valley et al., 2015; Inoue et al., 2016). Biological pathway annotation of DERNAs showed that the GPCR ligand binding process accumulated the largest number of dysregulated genes (86 DERNAs), which indicated that the pathway may play an important role in the development and progression of tumors.

In addition, KEGG pathways analysis showed that DERNAs are mainly enriched in neuroactive ligand–receptor interaction, alcoholism, systemic lupus erythematosus, metabolism of xenobiotics by cytochrome P450, and steroid hormone biosynthesis pathways, which are related to the progression of many cancers, including lung cancer (Ashton et al., 2010; Bulk et al., 2017). Among these enriched pathways, the neuroactive ligand–receptor interaction pathway accumulated the most dysregulated genes (53 DERNAs), indicating that they were associated with lung cancer progression. Multiple DERNAs were enriched in both alcoholism pathway and systemic lupus

erythematosus pathway. To our surprise, all DERNAs enriched in these two pathways were up-regulated, except for gene 2354, whose gene symbol is "FOSB." Gene annotation of FOSB shows that it is a proto-oncogene, and it has been considered as a regulator of cell proliferation, differentiation, and transformation (Liu et al., 2018; Na and Kim, 2018; Park et al., 2019).

In this study, a total of 55 proteins (including 45 DEmRNAs) were enrolled in the PPI network, and pathway enrichment analysis was performed based on the 10 hub genes. Most of the 10 hub genes were associated to cell cycle-related pathways, including M Phase (Zhang et al., 2018; Chung et al., 2019), Cell Cycle Checkpoints (Lee et al., 2016; Wenzel and Singh, 2018), TP53 Regulates Transcription of Cell Cycle Genes (Ni et al., 2018), and Viral carcinogenesis pathways (Villa, 2013; Shibata et al., 2016), which play an important role in occurrence and development of tumors (Florez et al., 2017; Niculescu, 2019). Cell cycle disorder and cell overgrowth are common biological characteristic of tumors, leading to increased cell proliferation and decreased apoptosis (Hsiao et al., 2014; Kebsa et al., 2018). It should be noted that the cell cycle is a tightly regulated process, which is frequently aberrant in lung cancer (Shcherba et al., 2014). By inhibiting the unrestricted cell division and growth of lung cancer cells, cell cycle-related genes have emerged as new targets for the treatment of lung cancer (Girek et al., 2019).

Among the 10 hub genes, 9 were significantly associated with muscle invasive bladder cancer, including CCNA2, CDC20, CDCA8, DLGAP5, KIF20A, NCAPG, NDC80, PBK, and TOP2A

TABLE 2 | Reactome and KEGG pathway enrichment results.

Term ID	Term description	Observed counts	Background count	FDR
HSA-69278	Cell cycle, mitotic	8	483	8.61E-10
HSA-68886	M phase	6	343	2.74E-07
HSA-68877	Mitotic prometaphase	5	190	7.61E-07
HSA-69620	Cell cycle checkpoints	5	265	3.09E-06
HSA-2500257	Resolution of sister chromatid cohesion	4	118	6.59E-06
HSA-1538133	G0 and early G1	3	27	7.46E-06
HSA-170145	Phosphorylation of proteins involved in the G2/M transition by Cyclin A: Cdc2 complexes	2	3	3.96E-05
HSA-174184	Cdc20: Phospho-APC/C mediated degradation of cyclin A	3	68	8.08E-05
HSA-176408	Regulation of APC/C activators between G1/S and early anaphase	3	76	8.08E-05
HSA-176417	Phosphorylation of Emi1	2	6	8.08E-05
HSA-141444	Amplification of signal from unattached kinetochores via a MAD2 inhibitory signal	3	90	9.66E-05
HSA-2514853	Condensation of prometaphase chromosomes	2	11	0.00013
HSA-1362300	Transcription of E2F targets under negative control by p107 (RBL1) and p130 (RBL2) in complex with HDAC1	2	16	0.00023
HSA-5663220	RHO GTPases activate formins	3	131	0.00023
HSA-174048	APC/C: Cdc20 mediated degradation of cyclin B	2	22	0.00036
HSA-2467813	Separation of sister chromatids	3	178	0.00046
HSA-6804757	Regulation of TP53 degradation	2	35	0.00072
HSA-4615885	SUMOylation of DNA replication proteins	2	40	0.00087
HSA-6791312	TP53 regulates transcription of cell cycle genes	2	49	0.0012
HSA-5688426	Deubiquitination	3	279	0.0013
HSA-597592	Post-translational protein modification	5	1366	0.0013
HSA-69206	G1/S transition	2	128	0.0068
hsa04110 (KEGG)	Cell cycle	3	123	0.00045
hsa05203 (KEGG)	Viral carcinogenesis	3	183	0.00072
hsa04914 (KEGG)	Progesterone-mediated oocyte maturation	2	94	0.0052
hsa04114 (KEGG)	Oocyte meiosis	2	116	0.0059
hsa04218 (KEGG)	Cellular senescence	2	156	0.0084

(Lee et al., 2012). It indicated that there may exist relative risk between muscle invasive bladder cancer and LUAD. Furthermore, we found that all these 10 hub genes were up-regulated in LUAD tumor tissue. The PPI network showed that almost all the 10 hub genes could interact with each other, and DLGAP5, CDK1, and KIF20A play a key role in connecting the network. Among them, DLGAP5 could physically interact with PBK, TOP2A, and CDK1, and all mitosis-associated proteins correlated with poor prognosis for non-small cell lung cancer patients (Shih et al., 2012; Schneider et al., 2017). In addition, a previous study reported that CCNA2, CDC20, PBK, and TOP2A that interacted with CDK1 play vital roles in survival outcomes in human lung cancer. Loss of cytoplasmic CDK1 could predict poor survival in human lung cancer and confers chemotherapeutic resistance (Zhang et al., 2011). Hence, we concluded that these 10 hub genes play key roles in regulating cell proliferation in LUAD.

A total of three DEmiRNAs, seven DElncRNAs, and six DEmRNAs were enrolled in the ceRNA network. In ceRNA network, we found that MALAT1 as a highly conserved lncRNA whose overexpression has been shown in various cancers, such as breast, prostate, colon, and liver, especially in early stage metastasizing patients (Lin et al., 2007; Guffanti et al., 2009; Xu et al., 2011; Ren et al., 2013). In addition, Ping et al. have reported that MALAT1 can predict metastasis in early stage NSCLC (Ji et al., 2003). Consistent with Ping et al., Lars et al. verified that MALAT1 stimulates migration, invasion, and tumor growth (Schmidt et al., 2011), although the underlying mechanism is poorly understood. In our ceRNA network, the expression of miR490 is down-regulated while MALAT1 and HMGA2 expression is up-regulated in LUAD. One possible explanation is that aberrant expression of MALAT1 acts as a ceRNA for miR-490, and high-expression MALAT1 inhibits miR490 and then increased expression of HMGA2 (the target of miR490), finally accelerating to tumor progression.

Many homeobox genes, including HOXC8, HOXB7, and HOXA9, are also "members" of the ceRNA network. A previous study reported that mis-expression of homeobox genes can lead to abnormal differentiation and proliferation, leading to a change in cell identity or homeotic transformation, therefore playing an important role in carcinogenesis (Samuel and Naora, 2005). In cancer, homeobox genes function as "tumor
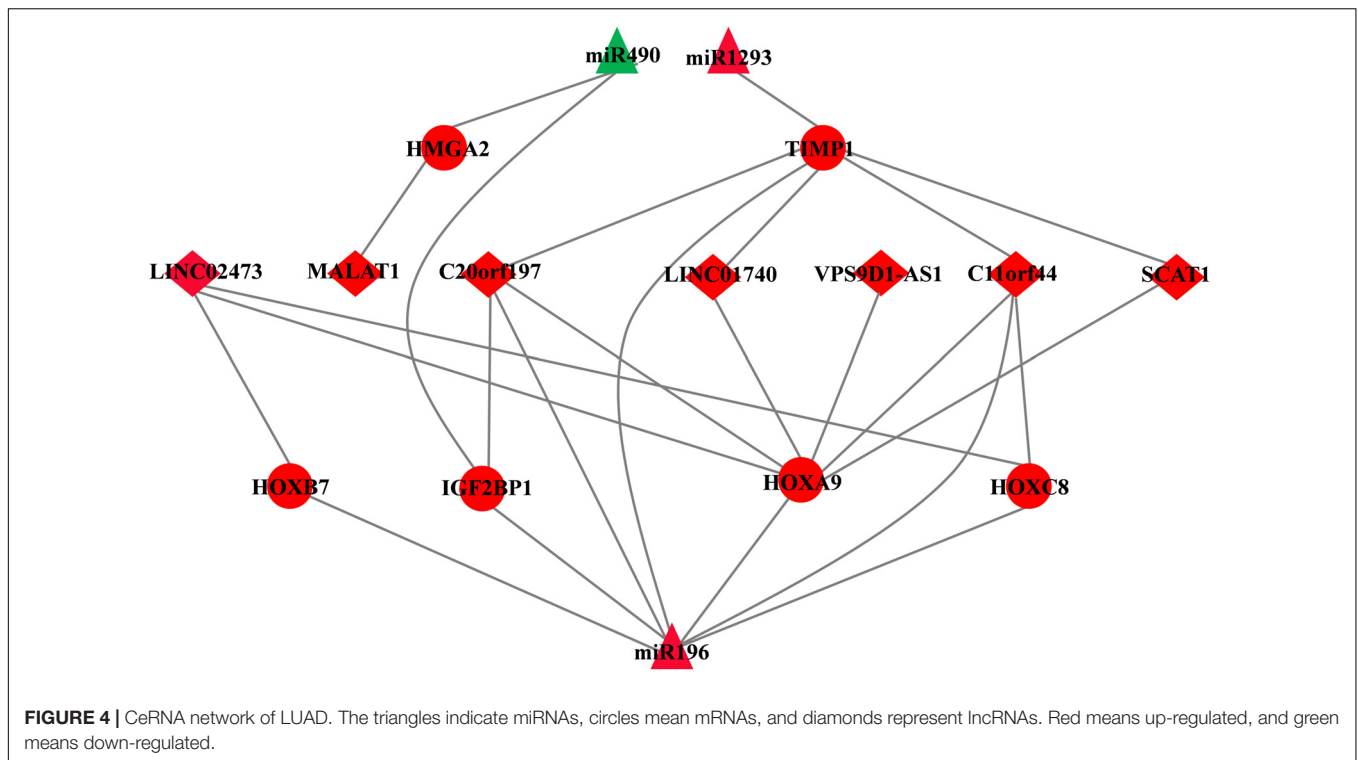
FIGURE 4 | CeRNA network of LUAD. The triangles indicate miRNAs, circles mean mRNAs, and diamonds represent lncRNAs. Red means up-regulated, and green means down-regulated.

TABLE 3 | The miRTarBase database revealed interactions between miRNA and mRNAs.

miRNA	mRNA
miR1293	TIMP1
miR196b	HOXB8, HOXC8, CD8A, HOXA9, MEIS1, FAS, ETS2, RDX, HOXB7, GATA6, TGFBR2, PIK3CG, AKT1, MTOR, FOS, IGF2BP1,
miR490	ERGIC3, FOS, SMARCD1, CCND1, PIK3CA, PAPPA, ABCC2, TGFBR1, HMGA2, TGFA, RHOA, PCBP1, TNKS2, BMPR2, HNRNPA1

modulators" as their deregulation normally involve either up-regulation of genes expressed in undifferentiated cells or down-regulation of genes expressed in differentiated tissue, thus acting either as oncogenes or tumor suppressor genes (Abate-Shen, 2002). Almost all the genes in the ceRNA network have reported that they enrolled or associated with tumor progression, except for LINC02473, LINC0170, VPS9D1-AS1, C11orf44, and SCAT1. Hence, taking all these genes in the ceRNA network into consideration, we combined step regression and Cox regression analysis and identified four genes as prognostic biomarkers in LUAD, including miR490, miR1293, LINC01740, and IGF2BP1.

By searching these genes in PubMed, we found that miR490 and IGF2BP1 have been studied for their mechanism in or association with tumor progression. Gain- and loss-of-function studies of miR490 showed that it regulates cell proliferation and is required for induction of *in vitro* migration and invasion (Zhao and Zheng, 2016). miR490 overexpression reduced proliferation, promoted G1 arrest and apoptosis,

and suppressed migration and invasion (Sun et al., 2016). In our study, miR490 expression was significantly lower in lung cancer than in normal tissues, and survival analysis result showed that the lower expression miR490 predicted poor survival in lung cancer. Opposite to miR490, IGF2BP1 expression is up-regulated, and the high expression level of IGF2BP1 showed poor overall survival outcomes in lung cancer. Studies reported that IGF2BP1 has been traditionally regarded as an oncogene and potential therapeutic target for cancers (Huang et al., 2018). It plays essential roles in embryogenesis and carcinogenesis, regulating the expression of some essential mRNA targets required for the control of tumor cell proliferation, growth, and invasion, and associating with a poor overall survival and metastasis in various types of human cancers (Gong et al., 2016). However, there is no public report on miR1293 and LINC01740 according to a PubMed search. Univariate Cox regression analysis showed that high expression of miR1293 tended to show poor survival outcomes (logrank_$p$value < 0.0001), and high expression of LINC01740 tended to show good survival outcomes (logrank_value $= 0.048$). Our results suggest that the four predictive genes may play crucial roles in the pathomechanism of LUAD and act as potential prognostic biomarkers.

Although a four-predictive gene signature was constructed and appears to be potential prognostic biomarkers in clinical application, there are some limitations. First, the prognostic value of LINC01740 is not very satisfactory. Second, the binding affinities between lncRNA and miRNA were predicted by simple linear regression model and miRcode and should be further experimentally investigated. Third, the function and mechanism

TABLE 4 | Information about differentially expressed RNAs and miRNAs in ceRNA network.

DERNAs/DEmiRNAs	Log2FC	p value	FDR	logrank_pvalue
miR1293	4.00	2.2e-13	1.20e-12	9.48e-05
miR196b	3.94	1.91e-31	2.80e-30	3.45e-02
miR490	−2.31	7.58e-07	2.63e-06	3.53e-02
C20orf197	3.06	9.22e-19	7.80e-18	0.003
SCAT1	2.87	3.25e-23	3.89e-22	0.006
C11orf44	2.32	9.35e-07	2.45e-06	0.019
MALAT1	2.09	5.96e-10	2.22e-09	0.023
VPS9D1-AS1	2.86	3.57e-38	9.82e-37	0.029
LINC02473	4.18	5.18e-25	6.98e-24	0.045
LINC01740	3.32	6.55e-07	1.74e-06	0.049
HMGA2	5.77	1.22e-22	1.39e-21	0.152
TIMP1	1.53	5.33e-25	7.16e-24	0.961
HOXB7	2.09	2.99e-17	2.24e-16	0.006
IGF2BP1	6.41	4.26e-21	4.34e-20	0.042
HOXA9	1.77	1.51e-07	4.33e-07	0.371
HOXC8	2.67	3.23e-13	1.66e-12	0.315
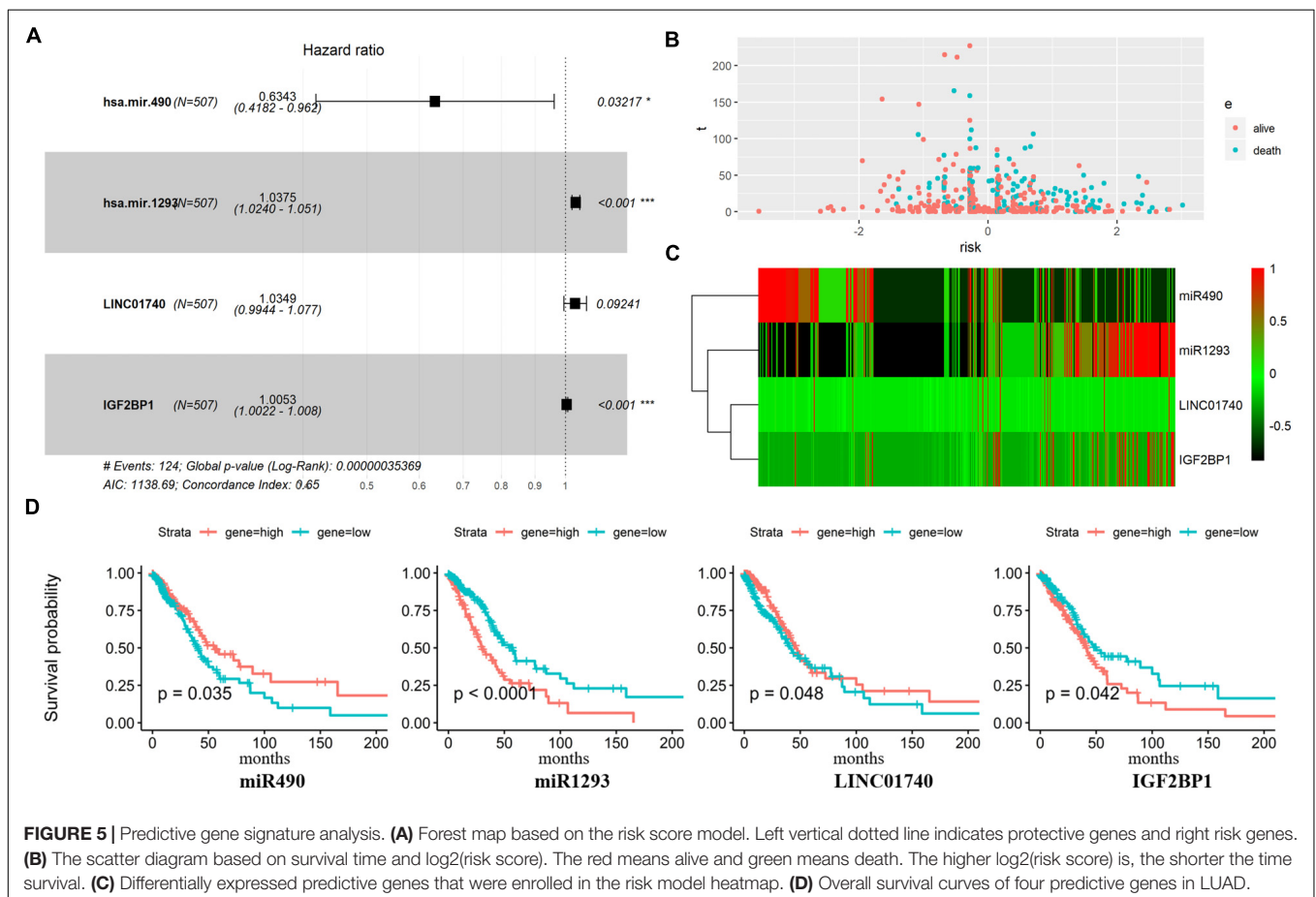
FIGURE 5 | Predictive gene signature analysis. **(A)** Forest map based on the risk score model. Left vertical dotted line indicates protective genes and right risk genes. **(B)** The scatter diagram based on survival time and log2(risk score). The red means alive and green means death. The higher log2(risk score) is, the shorter the time survival. **(C)** Differentially expressed predictive genes that were enrolled in the risk model heatmap. **(D)** Overall survival curves of four predictive genes in LUAD.

of the four predictive genes in LUAD need to be further studied by experiments.

In conclusion, we established the disordered ceRNA network, which is beneficial to understanding the relationship among lncRNA–miRNA–mRNA and provides efficient strategies for subsequent functional studies of them. In addition, we identified that miR1293, miR490, LINC01740, and IGF2BP1 might be novel important prognostic factors involved in LUAD pathogenesis, and the risk score model is helpful in studying the overall survival outcome in LUAD.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this study are publicly available from The Cancer Genome Atlas (TCGA) database, and can be accessed here: https://portal.gdc.cancer.gov/).

AUTHOR CONTRIBUTIONS

YW developed the theory and performed the computations. RH verified the analytical methods and results. LM conceived the original idea and supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

REFERENCES

Abate-Shen, C. (2002). Deregulated homeobox gene expression in cancer: cause or consequence? *Nat. Rev. Cancer* 2, 777–785. doi: 10.1038/nrc907

Ashton, K. A., Proietto, A., Otton, G., Symonds, I., McEvoy, M., Attia, J., et al. (2010). Polymorphisms in genes of the steroid hormone biosynthesis and metabolism pathways and endometrial cancer risk. *Cancer Epidemiol.* 34, 328–337. doi: 10.1016/j.canep.2010.03.005

Bao, S., Zhao, H., Yuan, J., Fan, D., Zhang, Z., Su, J., et al. (2019). Computational identification of mutator-derived lncRNA signatures of genome instability for improving the clinical outcome of cancers: a case study in breast cancer. *Brief Bioinform.* doi: 10.1093/bib/bbz118 [Epub ahead of print]

Bulk, E., Kramko, N., Liashkovich, I., Glaser, F., Schillers, H., Schnittler, H. J., et al. (2017). KCa3.1 channel inhibition leads to an ICAM-1 dependent increase of cell-cell adhesion between A549 lung cancer and HMEC-1 endothelial cells. *Oncotarget* 8, 112268–112282. doi: 10.18632/oncotarget.22735

Cagle, P. T., Allen, T. C., and Olsen, R. J. (2013). Lung cancer biomarkers: present status and future developments. *Arch. Pathol. Lab. Med.* 137, 1191–1198. doi: 10.5858/arpa.2013-0319-CR

Castro-Oropeza, R., Melendez-Zajgla, J., Maldonado, V., and Vazquez-Santillan, K. (2018). The emerging role of lncRNAs in the regulation of cancer stem cells. *Cell Oncol.* 41, 585–603. doi: 10.1007/s13402-018-0406-4

Chung, K. S., Hong, J. Y., Lee, J. H., Lee, H. J., Park, J. Y., Choi, J. H., et al. (2019). beta-Caryophyllene in the essential oil from chrysanthemum boreale induces G1 phase cell cycle arrest in human lung cancer cells. *Molecules* 24:E3754. doi: 10.3390/molecules24203754

Ezeife, D. A., and Leighl, N. B. (2018). Personalized medicine for non-small cell lung cancer: where are we now and where can we go? *Expert. Rev. Respir. Med.* 12, 81–82. doi: 10.1080/17476348.2018.1411805

Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet.* 15, 7–21. doi: 10.1038/nrg3606

Florez, M. M., Feo, H. B., da Silva, G. N., Yamatogi, R. S., Aguiar, A. J., Araujo, J. P., et al. (2017). Cell cycle kinetics, apoptosis rates and gene expressions of MDR-1, TP53, BCL-2 and BAX in transmissible venereal tumour cells and their association with therapy response. *Vet. Comp. Oncol.* 15, 793–807. doi: 10.1111/vco.12220

Girek, M., Klosinski, K., Grobelski, B., Pizzimenti, S., Cucci, M. A., Daga, M., et al. (2019). Novel tetrahydroacridine derivatives with iodobenzoic moieties induce G0/G1 cell cycle arrest and apoptosis in A549 non-small lung cancer and HT-29 colorectal cancer cells. *Mol. Cell. Biochem.* 460, 123–150. doi: 10.1007/s11010-019-03576-x

Gong, F., Ren, P., Zhang, Y., Jiang, J., and Zhang, H. (2016). MicroRNAs-491-5p suppresses cell proliferation and invasion by inhibiting IGF2BP1 in non-small cell lung cancer. *Am. J. Transl. Res.* 8, 485–495.

Guffanti, A., Iacono, M., Pelucchi, P., Kim, N., Solda, G., Croft, L. J., et al. (2009). A transcriptional sk+etch of a primary human breast cancer by 454 deep sequencing. *BMC Genomics* 10:163. doi: 10.1186/1471-2164-10-163

Gutschner, T., Hammerle, M., and Diederichs, S. (2013). MALAT1 – a paradigm for long noncoding RNA function in cancer. *J. Mol. Med.* 91, 791–801. doi: 10.1007/s00109-013-1028-y

Hsiao, C. J., Hsiao, G., Chen, W. L., Wang, S. W., Chiang, C. P., Liu, L. Y., et al. (2014). Cephalochromin induces G0/G1 cell cycle arrest and apoptosis in A549 human non-small-cell lung cancer cells by inflicting mitochondrial disruption. *J. Nat. Prod.* 77, 758–765. doi: 10.1021/np400517g

Huang, X., Zhang, H., Guo, X., Zhu, Z., Cai, H., and Kong, X. (2018). Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) in cancer. *J. Hematol. Oncol.* 11:88. doi: 10.1186/s13045-018-0628-y

Inoue, T., Adachi, K., Kawana, K., Taguchi, A., Nagamatsu, T., Fujimoto, A., et al. (2016). Cancer-associated fibroblast suppresses killing activity of natural killer cells through downregulation of poliovirus receptor (PVR/CD155), a ligand of activating NK receptor. *Int. J. Oncol.* 49, 1297–1304. doi: 10.3892/ijo.2016.3631

Ito, K., and Murphy, D. (2013). Application of ggplot2 to Pharmacometric Graphics. *CPT Pharmacometr. Syst. Pharmacol.* 2:e79. doi: 10.1038/psp.2013.56

Ji, P., Diederichs, S., Wang, W., Boing, S., Metzger, R., Schneider, P. M., et al. (2003). MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* 22, 8031–8041. doi: 10.1038/sj.onc.1206928

Kang, Y., Jia, Y., Wang, Q., Zhao, Q., Song, M., Ni, R., et al. (2019). Long Noncoding RNA KCNQ1OT1 promotes the progression of non-small cell lung cancer via regulating miR-204-5p/ATG3 Axis. *Oncol. Targets Ther.* 12, 10787–10797. doi: 10.2147/OTT.S226044

Kebsa, W., Lahouel, M., Rouibah, H., Zihlif, M., Ahram, M., Abu-Irmaileh, B., et al. (2018). Reversing multidrug resistance in chemo-resistant human lung Adenocarcinoma (A549/DOX) cells by Algerian Propolis through direct inhibiting the P-gp Efflux-pump, G0/G1 cell cycle arrest and apoptosis induction. *Anticancer Agents Med. Chem.* 18, 1330–1337. doi: 10.2174/1871520618666180808100800

Kerr, K. M., Bubendorf, L., Edelman, M. J., Marchetti, A., Mok, T., Novello, S., et al. (2014). Second ESMO consensus conference on lung cancer: pathology and molecular biomarkers for non-small-cell lung cancer. *Ann. Oncol.* 25, 1681–1690. doi: 10.1093/annonc/mdu145

Lee, S. J., Lee, E. J., Kim, S. K., Jeong, P., Cho, Y. H., Yun, S. J., et al. (2012). Identification of pro-inflammatory cytokines associated with muscle invasive bladder cancer; the roles of IL-5, IL-20, and IL-28A. *PLoS One* 7:e40267. doi: 10.1371/journal.pone.0040267

Lee, Y. S., Doonan, B. B., Wu, J. M., and Hsieh, T. C. (2016). Combined metformin and resveratrol confers protection against UVC-induced DNA damage in A549 lung cancer cells via modulation of cell cycle checkpoints and DNA repair. *Oncol. Rep.* 35, 3735–3741. doi: 10.3892/or.2016.4740

Lin, R., Maeda, S., Liu, C., Karin, M., and Edgington, T. S. (2007). A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene* 26, 851–858. doi: 10.1038/sj.onc.1209846

Liu, S., Luan, J., and Ding, Y. (2018). miR-144-3p Targets FosB Proto-oncogene, AP-1 transcription factor subunit (FOSB) to suppress proliferation, migration, and invasion of PANC-1 pancreatic cancer cells. *Oncol. Res.* 26, 683–690. doi: 10.3727/096504017X14982585511252

Liu, Z., Wang, Y., Wang, L., Yao, B., Sun, L., Liu, R., et al. (2019). Long non-coding RNA AGAP2-AS1, functioning as a competitive endogenous RNA, upregulates ANXA11 expression by sponging miR-16-5p and promotes proliferation and metastasis in hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* 38:194. doi: 10.1186/s13046-019-1188-x

Liz, J., and Esteller, M. (2016). lncRNAs and microRNAs with a role in cancer development. *Biochim. Biophys. Acta* 1859, 169–176. doi: 10.1016/j.bbagrm.2015.06.015

Luo, W., and Brouwer, C. (2013). Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* 29, 1830–1831. doi: 10.1093/bioinformatics/btt285

Majeed, H., and Amir, E. (2018). Personalized medicine in immuno-oncology: a novel prognostic index in non-small cell lung cancer. *J. Thorac. Dis.* 10(Suppl. 9), S995–S998. doi: 10.21037/jtd.2018.03.153

Na, H. H., and Kim, K. C. (2018). SETDB1-mediated FosB regulation via ERK2 is associated with an increase in cell invasiveness during anticancer drug treatment of A549 human lung cancer cells. *Biochem. Biophys. Res. Commun.* 495, 512–518. doi: 10.1016/j.bbrc.2017.10.176

Ni, Z., Wang, X., Zhang, T., Li, L., and Li, J. (2018). Comprehensive analysis of differential expression profiles reveals potential biomarkers associated with the cell cycle and regulated by p53 in human small cell lung cancer. *Exp. Ther. Med.* 15, 3273–3282. doi: 10.3892/etm.2018.5833

Niculescu, V. F. (2019). The reproductive life cycle of cancer: hypotheses of cell of origin, TP53 drivers and stem cell conversions in the light of the atavistic cancer cell theory. *Med. Hypotheses* 123, 19–23. doi: 10.1016/j.mehy.2018.12.006

Park, J. A., Na, H. H., Jin, H. O., and Kim, K. C. (2019). Increased expression of FosB through reactive oxygen species accumulation functions as pro-apoptotic protein in Piperlongumine treated MCF7 breast cancer cells. *Mol. Cells* 42, 884–892. doi: 10.14348/molcells.2019.0088

Ren, S., Liu, Y., Xu, W., Sun, Y., Lu, J., Wang, F., et al. (2013). Long noncoding RNA MALAT-1 is a new potential therapeutic target for castration resistant prostate cancer. *J. Urol.* 190, 2278–2287. doi: 10.1016/j.juro.2013.07.001

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

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

Samuel, S., and Naora, H. (2005). Homeobox gene expression in cancer: insights from developmental regulation and deregulation. *Eur. J. Cancer* 41, 2428–2437. doi: 10.1016/j.ejca.2005.08.014

Schmidt, L. H., Spieker, T., Koschmieder, S., Schaffers, S., Humberg, J., Jungen, D., et al. (2011). The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *J. Thorac. Oncol.* 6, 1984–1992. doi: 10.1097/JTO.0b013e3182307eac

Schneider, M. A., Christopoulos, P., Muley, T., Warth, A., Klingmueller, U., Thomas, M., et al. (2017). AURKA, DLGAP5, TPX2, KIF11 and CKAP5: five specific mitosis-associated genes correlate with poor prognosis for non-small cell lung cancer patients. *Int. J. Oncol.* 50, 365–372. doi: 10.3892/ijo.2017.3834

Shcherba, M., Liang, Y., Fernandes, D., Perez-Soler, R., and Cheng, H. (2014). Cell cycle inhibitors for the treatment of NSCLC. *Expert Opin. Pharmacother.* 15, 991–1004. doi: 10.1517/14656566.2014.902935

Shi, X., Tan, H., Le, X., Xian, H., Li, X., Huang, K., et al. (2018). An expression signature model to predict lung adenocarcinoma-specific survival. *Cancer Manag. Res.* 10, 3717–3732. doi: 10.2147/CMAR.S159563

Shibata, M., Minami, S., Oe, S., Miyagawa, K., Honma, Y., Onitsuka, R., et al. (2016). Recognition of infection, status of outpatient treatment, and treatment history before carcinogenesis in patients with viral hepatitis-associated liver cancer. *Nihon Shokakibyo Gakkai Zasshi* 113, 1734–1742. doi: 10.11405/nisshoshi.113.1734

Shih, M. C., Chen, J. Y., Wu, Y. C., Jan, Y. H., Yang, B. M., Lu, P. J., et al. (2012). TOPK/PBK promotes cell migration via modulation of the PI3K/PTEN/AKT pathway and is associated with poor prognosis in lung cancer. *Oncogene* 31, 2389–2400. doi: 10.1038/onc.2011.419

Sun, J., Zhao, H., Lin, S., Bao, S., Zhang, Y., Su, J., et al. (2019). Integrative analysis from multi-centre studies identifies a function-derived personalized multi-gene signature of outcome in colorectal cancer. *J. Cell Mol. Med.* 23, 5270–5281. doi: 10.1111/jcmm.14403

Sun, K. X., Chen, Y., Chen, S., Liu, B. L., Feng, M. X., Zong, Z. H., et al. (2016). The correlation between microRNA490-3p and TGFalpha in endometrial

carcinoma tumorigenesis and progression. *Oncotarget* 7, 9236–9249. doi: 10.18632/oncotarget.7061

Valley, C. C., Arndt-Jovin, D. J., Karedla, N., Steinkamp, M. P., Chizhik, A. I., Hlavacek, W. S., et al. (2015). Enhanced dimerization drives ligand-independent activity of mutant epidermal growth factor receptor in lung cancer. *Mol. Biol. Cell* 26, 4087–4099. doi: 10.1091/mbc.E15-05-0269

Villa, L. L. (2013). Viral carcinogenesis: virus implicated in cancer. *BMC Proc.* 7(Suppl. 2):K11. doi: 10.1186/1753-6561-7-S2-K11

Wenzel, E. S., and Singh, A. T. K. (2018). Cell-cycle checkpoints and aneuploidy on the path to cancer. *In Vivo* 32, 1–5. doi: 10.21873/invivo.11197

Xu, C., Yang, M., Tian, J., Wang, X., and Li, Z. (2011). MALAT-1: a long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis. *Int. J. Oncol.* 39, 169–175. doi: 10.3892/ijo.2011.1007

Yan, X., Hu, Z., Feng, Y., Hu, X., Yuan, J., Zhao, S. D., et al. (2015). Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* 28, 529–540. doi: 10.1016/j.ccell.2015.09.006

Yu, F., Quan, F., Xu, J., Zhang, Y., Xie, Y., Zhang, J., et al. (2018). Breast cancer prognosis signature: linking risk stratification to disease subtypes. *Brief Bioinform.* 20, 2130–2140. doi: 10.1093/bib/bby073

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zeng, F., Wang, Q., Wang, S., Liang, S., Huang, W., Guo, Y., et al. (2019). Linc00173 promotes chemoresistance and progression of small cell lung cancer by sponging miR-218 to regulate Etk expression. *Oncogene* 39, 293–307. doi: 10.1038/s41388-019-0984-2

Zhang, C., Elkahloun, A. G., Robertson, M., Gills, J. J., Tsurutani, J., Shih, J. H., et al. (2011). Loss of cytoplasmic CDK1 predicts poor survival in human lung cancer and confers chemotherapeutic resistance. *PLoS One* 6:e23849. doi: 10.1371/journal.pone.0023849

Zhang, L. L., Feng, Z. L., Su, M. X., Jiang, X. M., Chen, X., Wang, Y., et al. (2018). Downregulation of Cyclin B1 mediates nagilactone E-induced G2 phase cell cycle arrest in non-small cell lung cancer cells. *Eur. J. Pharmacol* 830, 17–25. doi: 10.1016/j.ejphar.2018.04.020

Zhao, L., and Zheng, X. Y. (2016). MicroRNA-490 inhibits tumorigenesis and progression in breast cancer. *Oncol. Targets Ther.* 9, 4505–4516. doi: 10.2147/OTT.S100037

Zhou, M., Hu, L., Zhang, Z., Wu, N., Sun, J., and Su, J. (2018a). Recurrence-Associated Long Non-coding RNA signature for determining the risk of recurrence in patients with colon cancer. *Mol. Ther. Nucleic Acids* 12, 518–529. doi: 10.1016/j.omtn.2018.06.007

Zhou, M., Zhang, Z., Zhao, H., Bao, S., Cheng, L., and Sun, J. (2018b). An Immune-related Six-lncRNA signature to improve prognosis prediction of glioblastoma multiforme. *Mol. Neurobiol.* 55, 3684–3697. doi: 10.1007/s12035-017-0572-9

Zhou, M., Zhang, Z., Zhao, H., Bao, S., and Sun, J. (2018c). A novel lncRNA-focus expression signature for survival prediction in endometrial carcinoma. *BMC Cancer* 18:39. doi: 10.1186/s12885-017-3983-0

Zhou, M., Zhao, H., Xu, W., Bao, S., Cheng, L., and Sun, J. (2017). Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. *Mol. Cancer* 16:16. doi: 10.1186/s12943-017-0580-4