Pathology & Oncology Research (2020) 26:1935–1945
https://doi.org/10.1007/s12253-019-00780-4

ORIGINAL ARTICLE

Comprehensive Analysis of Aberrantly Expressed Profiles of lncRNAs and miRNAs with Associated ceRNA Network in Lung Adenocarcinoma and Lung Squamous Cell Carcinoma

Ruolan Dong[1] · Jiawei Liu[2] · Wei Sun[2] · Wei Ping[2]

Received: 13 June 2019 / Accepted: 19 November 2019 / Published online: 2 January 2020

Abstract

Lung cancer (LC) continues to be the leading cause of cancer-related deaths worldwide and the prognosis remains poor worldwide. At present, the long non-coding RNAs (lncRNAs) was considered as a part of competing endogenous RNA (ceRNA) network act as natural microRNA (miRNA) sponges to regulate protein-coding gene expression. However, functional roles of lncRNA-mediated ceRNAs in LC are insufficiently understood. To classify the specific mechanism of lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), we comprehensively compared the expression profiles of mRNAs, lncRNAs and miRNAs obtained from 509 LUAD, 473 LUSC tissues and 49 adjacent non-cancerous lung tissues, based on The Cancer Genome Atlas (TCGA). After screening for differently expressed (DE) mRNAs, DEmiRNAs, DElncRNAs and weighted gene co-expression network analysis (WGCNA) ($|\log2FC| > 2.0$ and an adjusted p value <0.05), a total of 4478 DEmRNAs, 526 DElncRNAs and 75 DEmiRNAs in LUAD, while 6237 DEmRNAs, 843 DElncRNAs and 117 DEmiRNAs in LUSC were discovered. Interaction (PPI) network analysis was performed to identify 656 nodes and 2987 edges (minimum required interaction score > 0.9), as well as 8 different protein-protein interactions. Gene ontology (GO) analysis mainly associated with cell proliferation. KEGG pathway enrichment analyses most partly associated with metabolism pathway and cytokine-cytokine receptor interaction. Finally, the dysregulated lncRNA-miRNA-ceRNA network was constructed based on correlation analyses and a total of 62 dysregulated lncRNAs, 28 DEmRNAs and 18 DEmiRNAs were involved. The most significant lncRNAs included DElncRNAs, LINC00641 and AC004947.2, miRNAs included miR-6860, miR-1285-3p, miR-767-3p and miR-7974, mRNAs included MAP3K3, FGD3 and ATP1B2. Then we analyzed and described the potential characteristics of biological function and pathological roles of the LUAD and LUSC ceRNA co-regulatory network. Our findings revealed ceRNA network will be beneficial for promoting the understanding of lncRNA-mediated ceRNA regulatory mechanisms in the pathogenesis of LUAD and LUSC.

Keywords Lung adenocarcinoma (LUAD) · Lung squamous cell carcinoma (LUSC) · lncRNA · ceRNA · miRNA

Introduction

Lung cancer (LC) remains the leading cause of cancer-related deaths worldwide. More than 85% are categorized as non-small cell lung cancer (NSCLC), of which lung adenocarcinoma (LUAD) accounts for about 50% and lung squamous cell carcinoma (LUSC) accounts for about 30% [1]. Despite developments in experimental and clinical oncology, the prognosis of NSCLC remains poor, with a 5-year overall survival (OS) rate of around 15% [2]. Emerging data reveals that the initiation and progression of NSCLC require a variety of changes in gene regulation [3, 4]. Hence, identification of potential biomarkers and therapeutic targets to improve NSCLC treatment effectiveness, especially LUAD and LUSC, is urgently needed.

Non-coding RNAs (ncRNAs) are a class of functional RNA molecules that could not be translated into proteins. MicroRNAs (miRNAs) are a well characterized class of

✉ Wei Ping
247046170@qq.com

[1] Institute of Integrated Traditional Chinese and Western Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, China

[2] Department of Thoracic Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

ncRNAs (~22 nt) that can regulate the expression of mRNAs, typically by binding 3′-untranslated regions (3'-UTRs), leading to decreased translation and/or degradation of the target mRNAs. It is well known that miRNAs play important roles in a variety of tumor biological processes, including cell proliferation, cell apoptosis and tumor metastasis [5]. Recently, another class of ncRNAs, the long non-coding RNAs (lncRNAs) is rapidly gaining attention [6]. Unlike the miRNAs, lncRNAs are non-protein coding transcripts longer than 200 nt in length. Cumulative evidence has revealed that lncRNAs may regulate gene expression at different levels, such as transcription, post-transcription and translation, carcinogenesis and cancer metastasis [7, 8]. In addition, accumulating evidence has suggested that lncRNAs play an important role in a wide variety of biological processes, including cell differentiation, chromatin modification, dosage compensation and so on [9]. Currently, with the ceRNA (competing endogenous RNA) hypothesis been proposed, the influence of lncRNAs upon miRNAs function is rapidly emerging. The ceRNA hypothesis that lncRNAs and mRNAs can act as natural miRNA sponges to suppress miRNA function by binding shared 3'-UTRs. The complex crosstalk of ceRNA network has been demonstrated in different kinds of diseases including cancers [10]. For example, lncRNA-KRTAP5-AS1 can regulate CLDN4 expression by competitively binding to miR-596 and miR-3620-3p and induced EMT in gastric cancer [11]. In hepatocellular carcinoma lnc-CASC2 could function as a ceRNA by sponging miR-367 [12].

The Cancer Genome Atlas (TCGA) database, which is a publicly available platform, contained sequencing data of lncRNAs, miRNAs and mRNAs from different cancer including LC. Based on ceRNA theory and TCGA database, ceRNA network has been constructed in hepatocellular cancer, gastric cancer, muscle-invasive bladder cancer, pancreatic cancer and breast cancer. In order to provide new insight into the molecular mechanisms based on lncRNA-miRNA-mRNA ceRNA for LUAD and LUSC, we conducted a comprehensive bioinformatics analysis based on TCGA database in this study.

Materials and Methods

TCGA Dataset of LUAD and LUSC

RNA sequencing (RNA-Seq) data from patients with LUAD and LUSC, which were calculated on Illumina HiSeq RNA-Seq platform, were retrieved from TCGA, containing 509 LUAD, 473 LUSC tissues and 49 adjacent non-cancerous lung tissues up to November 9, 2016. Detailed clinical data consisted of demographic data (age, sex and race), smoking history, tumor status and pathologic stages. All patients had a certain follow-up time. Since the data was obtained from TCGA, additional approval by the ethics committee of our hospital was not required. This study meets the publication guidelines provided by TCGA (http://cancergenome.nih.gov/publications/publicationguidelines).

RNA Sequence Data Procession

The LUAD and LUSC miRNASeq and RNAseq data (level 3) of the corresponding patients were downloaded from TCGA, including 509 LUAD, 473 LUSC tissues and 49 adjacent non-cancerous lung tissues. To detect the differential expression of miRNA, mRNA and lncRNA, samples were divided into LUAD tumor tissues vs. adjacent non-tumor lung tissues and LUSC tumor tissues vs. adjacent non-tumor lung tissues. Tumor sample and normal sample data were combined and data approximate to zero were deleted. The sequenced data generated from Illumina HiSeq_miRNASeq and Illumina HiSeq_RNASeq sequencing platforms were all publicly available data. Each sample contained the corresponding miRNAseq and RNA-seq data.

Identification of Differentially Expressed mRNAs (DEmRNAs), lncRNAs (DElncRNAs) and miRNAs (DEmiRNAs) in LUAD and LUSC

The differently expressed mRNA and lncRNA were analyzed by Ballgown software, a Bioconductor package based on R language, subsequently for the calculation of DERNAs and DElncRNAs (Padj <0.05, and the absolute log2 fold change >2), respectively. Screening of DEmiRNAs was used by DESeq2, based on R language for differential gene expression analysis using the nbinomTest command by instructions from the DESeq reference manual and analyzed by EdgeR [13] (Padj <0.001 and the absolute log2 fold change >3). For all the *P*-values, a false discovery rate (FDR) < 0.05 was applied to correct the statistical significance of multiple testing.

Bioinformatics Analysis

For functional enrichment analysis, we used the Gene Ontology database (http://www.geneontol-ogy.org) to perform gene ontology (GO) analysis on the DERNAs. After the analyses for significance and false discovery rate (FDR), GO terms were selected from the significantly enriched gene sets ($P < 0.05$ and FDR < 0.05). GO analysis covered enrichment of biological processes(BP), cellular components(CC) and specific molecular functions(MF) of divergently expressed genes. Pathway analysis was used to identify significant pathways for the DERNAs according to the Kyoto Encyclopedia of Genes and Genomes (KEGG). The significant KEGG pathways were selected by Fisher's exact and chi-square tests ($P < 0.05$ and FDR < 0.05).

Table 1 Clinicopathological characteristics of patients with LUAD or LUSC

Parameter	Subtype	LUAD n, (%)	LUSC n, (%)
Age (years)	Not available	19, (3.73)	8, (1.69)
	>68	209, (41.06)	225, (47.57)
	≤68	281, (55.21)	240, (50.74)
Gender	Female	272, (53.44)	123, (26.00)
	Male	237, (46.56)	350, (74.00)
Race	Not available	66, (13.60)	105, (22.91)
	Asian	7, (1.38)	9, (1.90)
	Black or African American	53, (10.41)	29, (6.13)
	White	383, (75.25)	330, (69.77)
Smoking history (years)	Not available	298, (58.55)	266, (56.24)
	>15	114, (22.39)	8, (1.69)
	≤15	97, (19.06)	199, (42.07)
Tumor status	Not available	54, (10.61)	70, (14.80)
	Tumor free	289, (56.78)	299, (63.21)
	With tumor	166, (32.61)	104, (21.99)
Pathologic stage	Not Available	7, (1.38)	4, (0.85)
	Stage I	5, (0.98)	83, (17.55)
	Stage IA	131, (25.74)	143, (30.23)
	Stage IB	139, (27.31)	3, (0.63)
	Stage II	1, (0.20)	62, (13.11)
	Stage IIA	50, (9.82)	90, (19.03)
	Stage IIB	68, (13.36)	3, (0.63)
	Stage IIIA	73, (14.34)	61, (12.90)
	Stage IIIB	11, (2.16)	15, (3.17)
	Stage IV	24, (4.72)	6, (1.27)

Pathological stage was based on NCCN Guidelines Version 3.2011 NSCLC

Protein-Protein Interaction (PPI) Network Construction

Protein interaction analysis is an important tool to discover LC-related DEGs of different modules based on protein network. PPI network was constructed to identify the key coding genes and important gene modules which were common different expression in LUAD and LUSC compared with tumor free tissue. PPIs information of different expression genes was initially extracted from Search Tool for the Retrieval of Interacting Genes (STRING) database, an online database integrating known and predicted protein interactions [14] (http://www.string-db.org/).

Construction of lncRNA-miRNA-mRNA ceRNA Network

In the present study, we constructed the lncRNA-miRNA-mRNA ceRNA network using Cytoscape software (Version3.5.1) to identify gene interactions, which was based on the theory that lncRNAs can have a direct interaction by invoking the miRNA sponge to regulate the activity of mRNAs. Predicted targets of DEmiRNAs in this study were retrieved using miRBase targets (http://www.targetscan.org/ and http://www.microrna.org/microrna/). LncRNA-miRNA interactions were predicted by miRcode (http://www.mircode.org/). The gene co-expression network was built according to the normalized signal intensity of specific expressed genes. We calculated the Pearson correlation coefficient between two genes.

Statistical Analysis

The significance level was set as 0.05 as default to control the false discovery rate (FDR). *P* value less than 0.05 were considered statistically significant unless specifically indicated. The statistical analyses were performed using a statistical software package (version 17.0. SPSS, Chicago, IL).

Results

Patient Characteristics

The baseline and clinicopathological characteristics of 509 patients that pathologically diagnosed as LUAD and 473

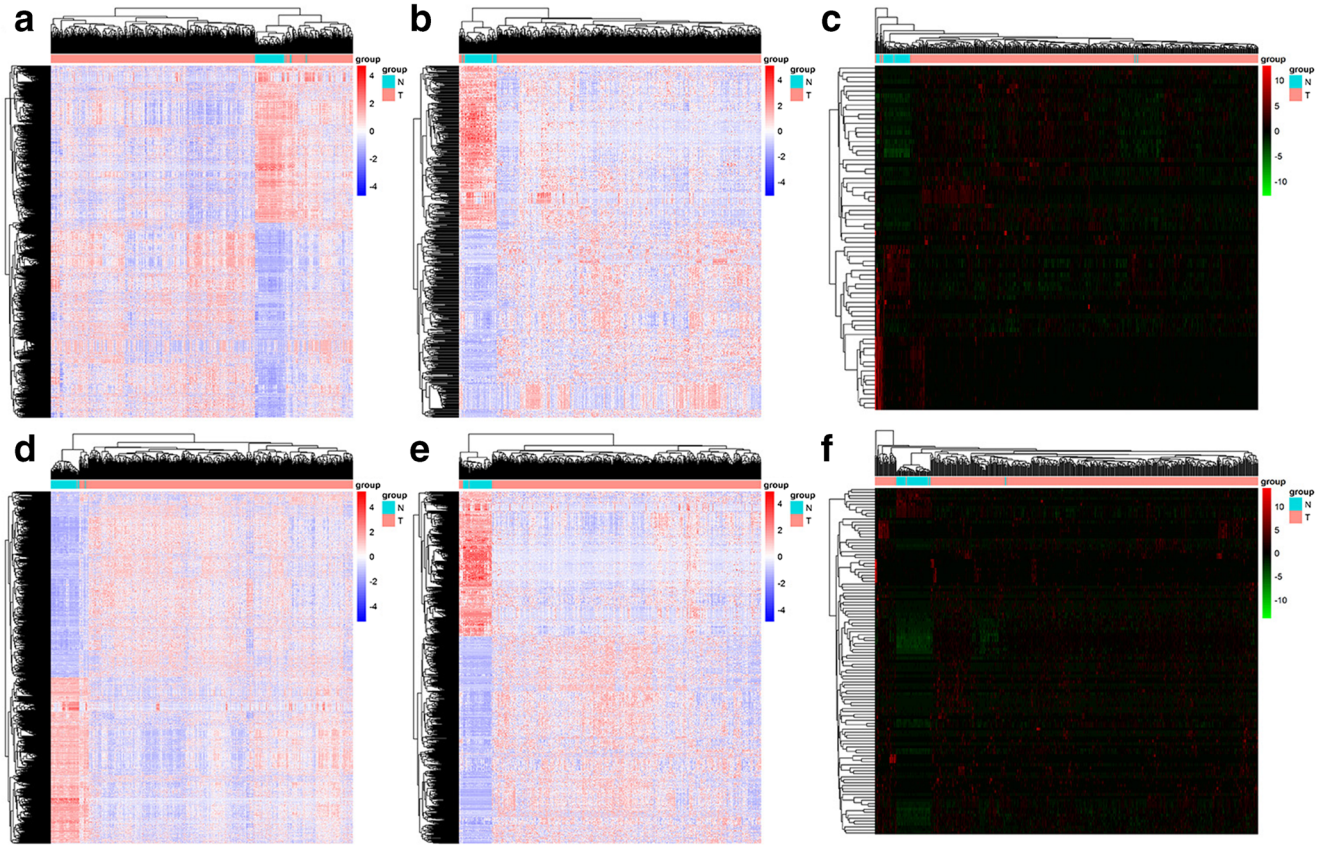

Fig. 1 Heat map showing expression profiles of mRNAs, lncRNAs and miRNAs. Differentially expressed mRNAs (**a, d**), lncRNAs (**b, e**) and miRNAs (**c, f**) in LUAD **(a, b, c)** and LUSC (**d, e, f**) were analyzed using hierarchical clustering. Compared with tumor free tissues. Red indicates high relative expression and blue or green indicates low relative expression. FC > 1.5 or FC < 0.7, $p < 0.05$. N, tumor free tissues; T, tumor tissues

patients pathologically diagnosed as LUSC, the parameters including demographic data (age, sex and race), smoking history, tumor status and pathologic stages were shown in Table 1.

Expression Profiles of lncRNAs, mRNAs and miRNAs in LUAD and LUSC

Hierarchical clustering (heat map) was created to reveal the expression clusters of these DERNAs, the results showed that a total of 4478 differently expressed mRNAs (Fig. 1a, FC >2 or FC < 0.5, $p < 0.05$), 526 differently expressed lncRNAs (Fig. 1b, FC >2 or FC < 0.5, $p < 0.05$) and 75 significant differently expressed miRNAs (Fig. 1c, FC > 3 or FC < −3, p < 0.05) in LUAD. Among these differently expressed RNAs, 2479 mRNAs, 282 lncRNAs and 39 miRNAs were up-regulated, 1999 mRNAs, 244 lncRNAs and 36 miRNAs were down-regulated (Table 2). Meanwhile, 6237 differently expressed mRNAs (Fig. 1d, FC >2 or FC < 0.5, p < 0.05), 843 differently expressed lncRNAs (Fig. 1e, FC >2 or FC < 0.5, p < 0.05) and 117 significant differently expressed miRNAs (Fig. 1f, FC > 3 or FC < −3, $p < 0.05$) in LUSC were revealed by heat map, and in this DERNAs, 3297 mRNAs, 495 lncRNAs and 107 miRNAs were up-regulated, 2940 mRNAs, 348 lncRNAs and 10 miRNAs were down-regulated (Table 2). What's more, the Venn diagram suggested that 1711 mRNAs, 127 lncRNAs and 25 miRNAs were common up-regulated, 1768 mRNAs, 213 lncRNAs and 5 miRNAs were common down-regulated in both LUAD and LUSC (Fig. 2a-c, Table 2). The top 10 DERNAs in LUAD and LUSC were shown in Tables 3 and 4).

Table 2 Common differentially expressed mRNA, lncRNA and miRNA in LUAD and LUSC

	Up-regulated			Down-regulated		
	mRNA	lncRNA	miRNA	mRNA	lncRNA	miRNA
LUAD	2479	282	39	1999	244	36
LUSC	3297	495	107	2940	348	10
Common	1711	127	25	1768	213	5

Compare with tumor free, respectively. FC>1.5 or FC<0.7

Protein-Protein Interaction (PPI) Analysis

In order to depict the complex relationships of the 1410 significant common differently expressed protein coding

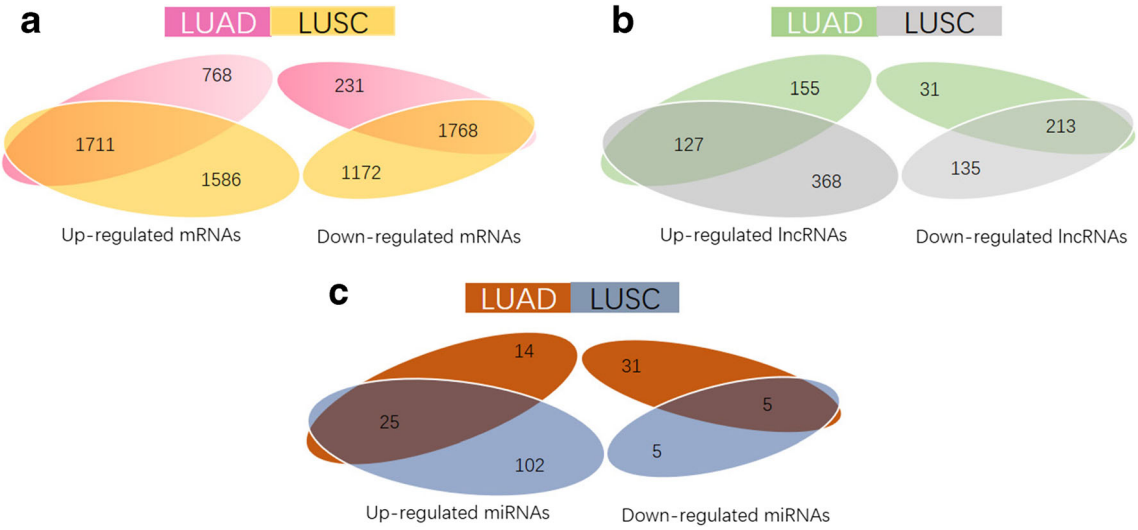

Fig. 2 Venn diagram showed the common differentially expressed RNAs in LUAD and LUSC compared with tumor free tissue, respectively. (**a**) 1711 mRNAs were common up-regulated and 1768 mRNAs were common down-regulated; (**b**) 127 lncRNAs were common induced, 213 were reduced in LUAD and LUSC, simultaneously; (**c**) 25 and 5 miRNAs were common up-regulated and down-regulated, respectively. $p < 0.05$, $FC > 1.5$ or $FC < 0.7$

genes in LUAD and LUAC, $p < 0.05$, $FC > 2.0$ or $FC < 0.5$, the PPI network was created. In this network, 656 nodes and 2987 edges were involved (minimum required interaction score > 0.9), and 8 different interactions of protein-protein were marked with different colors, respectively (Fig. 3).

Gene Ontology and KEGG Analysis

To identify the biological roles of the differently expression genes in LUAD and LUSC, the GO enrichment analyses and KEGG pathway analyses were performed. The GO analyses result showed the top 30 GO enrichment of

Table 3 Top 10 DERNAs (up-regulated or down-regulated) in LUAD

mRNA	FC (T/N)	Status	lncRNA	FC (T/N)	Status	miRNA	FC (T/N)	Status
SFTPC	0.0046	down	NAPSB	0.6049	down	hsa-miR-4732-3p	0.0453	down
AGER	0.0130	down	C10orf25	0.6651	down	hsa-miR-486-5p	0.0476	down
CLDN18	0.0198	down	SIGLEC17P	0.5817	down	hsa-miR-4634	0.0510	down
SCGB1A1	0.0275	down	RP11-218 M22.1	0.5479	down	hsa-miR-520a-3p	0.0528	down
FABP4	0.0322	down	RP4-539 M6.14	0.6382	down	hsa-miR-1323	0.0537	down
SFTPA1	0.0356	down	HLA-V	0.5974	down	hsa-miR-4732-5p	0.0633	down
SFTPA2	0.0409	down	MIR22HG	0.3798	down	hsa-miR-139-3p	0.0672	down
HBB	0.0427	down	TMSB4XP8	0.6452	down	hsa-miR-202-3p	0.0680	down
UPK3B	0.0476	down	COLCA1	0.5520	down	hsa-miR-517b-3p	0.0717	down
MCEMP1	0.0573	down	EEF1A1P5	0.3123	down	hsa-miR-517a-3p	0.0718	down
UBE2C	8.5087	up	CTA-384D8.34	2.2224	up	hsa-miR-301b-5p	28.7861	up
B3GNT3	8.6869	up	RP11-465 N4.5	1.6736	up	hsa-miR-577	29.6713	up
IGHG4	9.2968	up	RP5-963E22.6	1.5904	up	hsa-miR-1269b	31.6219	up
CTHRC1	9.3869	up	RP11-12G12.7	2.3048	up	hsa-miR-4652-5p	34.7719	up
CEACAM5	9.6498	up	RP11-631 N16.4	1.6986	up	hsa-miR-767-3p	36.4526	up
PYCR1	9.9004	up	PRSS3P2	2.4829	up	hsa-miR-767-5p	36.6590	up
SPINK1	10.8809	up	RP11-785D18.3	1.6998	up	hsa-miR-9-5p	38.9545	up
FAM83A	11.3036	up	IGHV3–41	1.8030	up	hsa-miR-1269a	46.1291	up
CRABP2	18.2451	up	SNHG4	1.6843	up	hsa-miR-196a-5p	58.5485	up
SPP1	20.3542	up	BLACAT1	2.0189	up	hsa-miR-105-5p	68.3640	up

DERNAs, differentially expressed RNAs; FC, fold change; T, tumor tissues; N, tumor free tissues; $p<0.05$

Table 4 Top 10 DERNAs (up-regulated or down-regulated) in LUSC

mRNA	FC (T/N)	Status	lncRNA	FC (T/N)	Status	miRNA	FC (T/N)	Status
SFTPC	0.0014	down	C10orf95	0.4895	down	hsa-miR-486-5p	0.0816	down
SFTPA1	0.0062	down	NAPSB	0.3188	down	hsa-miR-4678	0.0895	down
AGER	0.0073	down	KRT87P	0.6106	down	hsa-miR-144-5p	0.0913	down
SFTPA2	0.0077	down	CYP4Z2P	0.6431	down	hsa-miR-206	0.0932	down
CLDN18	0.0109	down	SIGLEC16	0.5803	down	hsa-miR-490-3p	0.0957	down
PGC	0.0180	down	C10orf25	0.6102	down	hsa-miR-4732-3p	0.0965	down
SFTPB	0.0208	down	CMAHP	0.3183	down	hsa-miR-451a	0.1020	down
SFTPD	0.0214	down	SIGLEC17P	0.4760	down	hsa-miR-3195	0.1104	down
NAPSA	0.0225	down	LINC00982	0.6637	down	hsa-miR-133b	0.1137	down
SCGB1A1	0.0226	down	RP11-218 M22.1	0.5283	down	hsa-miR-338-5p	0.1215	down
SLC2A1	39.3476	up	RP11-9E17.1	2.0700	up	hsa-miR-6499-5p	65.2579	up
KRT14	41.3925	up	RP5-1056H1.2	1.6455	up	hsa-miR-944	79.1490	up
CALML3	43.3539	up	RP11-245D16.4	2.0204	up	hsa-miR-4778-3p	82.2281	up
SPRR1B	44.3222	up	AC068831.16	1.7423	up	hsa-miR-1269a	83.8966	up
KRT6B	46.4879	up	LINC01451	1.8416	up	hsa-miR-767-5p	95.4303	up
S100A2	54.0973	up	SH3PXD2A-AS1	2.1173	up	hsa-miR-105-5p	102.7657	up
KRT16	67.6713	up	SNHG4	2.1894	up	hsa-miR-767-3p	118.3120	up
KRT17	76.3447	up	LINC01176	1.7261	up	hsa-miR-4652-5p	178.2873	up
KRT5	122.2849	up	BLACAT1	1.7855	up	hsa-miR-6499-3p	321.0574	up
KRT6A	241.9939	up	SAMD12-AS1	1.5618	up	hsa-miR-1269b	436.8335	up

DERNAs, differentially expressed RNAs; FC, fold change; T, tumor tissues; N, tumor free tissues; $p<0.05$

biological processes(BP), cellular components(CC) and specific molecular functions(MF) of up-regulated and down-regulated expressed genes, many of these terms were associated with cell proliferation, such as "mitotic cell cycle, GO:0000278", "M phase of mitotic cell cycle, GO:0000087", "cell division, GO:0051301", "mitotic anaphase, GO:0000090", "S phase of mitotic cell cycle, GO:0000084", "G1/S transition of mitotic cell cycle, GO:0000082", "DNA replication, GO:0006260", "cell cycle checkpoint, GO:0000075", "translation, GO:0006412", "positive regulation of cell proliferation, GO:0008284" and "cell proliferation, GO:0008283".(Fig. 4a and b). The KEGG pathway analysis showed the top 30 pathway enrichment of up-regulated and down-regulated expressed genes, and these differently expressed mRNAs mainly enriched in "metabolism pathways, pathway_ID: has01100", "Cell cycle, pathway_ID: has4110", "RNA transport, pathway_ID: has3013", "Cytokine-cytokine receptor interaction, pathway_ID: has4060" and "Pathways in cancer, pathway_ID: has5200" (Fig. 4c and d). This suggests that these GO terms and KEGG pathways may contribute to the pathogenesis and biochemical characteristics of LUAD and (or) LUSC in humans. Among these GO enrichments and pathway enrichments, 1226 up-regulated mRNAs and 1115 down-regulated genes were overlapped (Fig. 4e), and those overlapped genes might be the key point of LUAD and (or) LUSC development.

Co-Expression Analysis between Differentially Expressed miRNAs, lncRNAs and Coding Genes in LUAD and LUSC

To predict and describe the function of differentially expressed noncoding RNAs (miRNAs, lncRNAs) and coding genes in LUAD and LUSC, the dysregulated lncRNA-miRNA-mRNA ceRNA network was constructed based on correlation analysis. First, all differentially expressed miRNAs, lncRNAs and mRNAs from the present study were analyze by R language, LncRNAs and mRNAs with Pearson correlation coefficient (PCC) ≥ 0.99 were selected to draw the network, respectively. Pearson Correlation: $r > 0.7$ or < -0.7, $p < 0.05$. Finally, in the ceRNA network, a total of 62 dysregulated lncRNAs, included 7 up-regulated and 55 down-regulated lncRNAs, 28 differentially expressed mRNAs, included 5 up-regulated and 23 down-regulated mRNAs, and 18 dysregulated miRNAs, included 14 up-regulated and 4 down-regulated miRNAs, were involved and 208 connections were established (degree ≥ 2). The degree parameter represents the relative pivotal role of a gene and is important to evaluating the centrality of a gene in the network analysis. (Fig. 5). Meanwhile, the degree of miR-6860, miR-1285-3p, miR-767-3p, miR-7974, lncRNA-LINC00641, lncRNA-AC004947.2, lncRNA-109,642.1, lncRNA-MAGI2-AS3, lncRNA-SHANK3, SHROOM4, MAP3K3, FGD3, ATP1B2 and SAP30L were significantly higher, more importantly, some of them have been reported

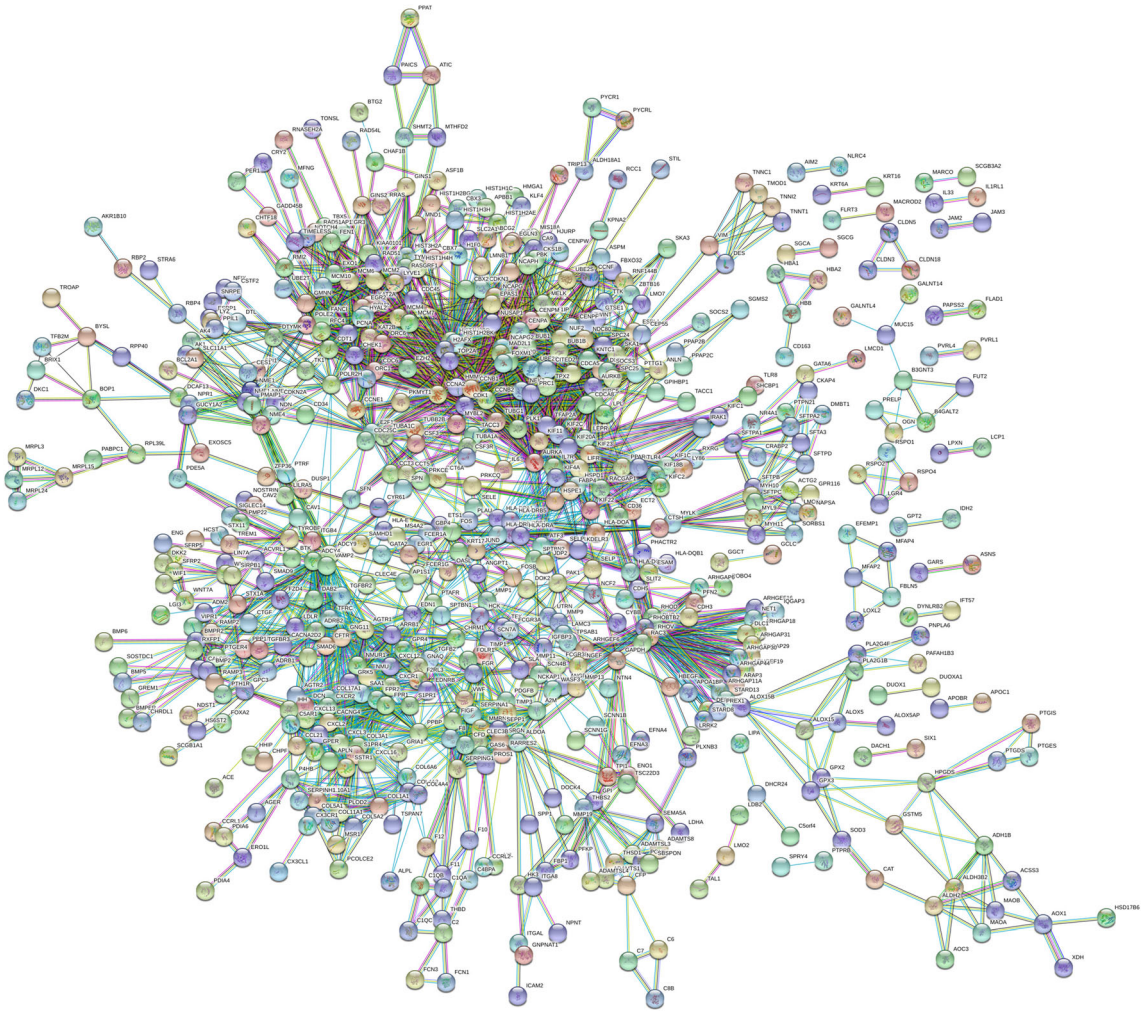

Fig. 3 Protein-protein interaction (PPI) network of overlapping genes both in LUAD and LUSC. PPI network was analyzed using an online tool, STRING. The high confidence (0.9) of minimum required interaction score had been performed and the disconnected nodes had been hidden in the network; The different colors of network edges represent different interactions of protein-protein

to be cancer-associated genes such as MAP3K3 [15–17], FGD3 [18], ATP1B2 [19].

Discussion

NSCLC is the most common type of lung cancer which is the leading cause of cancer-related mortality and its prognosis still remains poor. The therapeutic options for NSCLC are mainly dependent on the stage of the cancer. Although the five-year survival rate of NSCLC for local LC is about 50%, majority of patients miss the diagnosed opportunity in this stage [20]. Therefore, it is of great importance to further exploring the molecular mechanism leading to pathogenesis of NSCLC and ultimately assisting in the development of personalized diagnosis and treatment against the diseases.

Recently growing evidence indicated that ncRNAs take critical effects in regulating key pathogenic mechanisms in carcinogenesis and cancer metastasis. LncRNAs are considered to show the superior potential as diagnostic and prognostic biomarkers and clinical therapeutic applications. The significant upregulation of LncPVT1 in NSCLC, thought to be intensively associated poorer overall survival, facilitated the up-expression of HIF-1α by functioning as ceRNA of miR-199a-5p [21, 22]. LncRNA MEG3 down-regulation in NSCLC tissues partly mediated by DNA methylation and p53 activation may represent a new biomarker for poor diagnosis [23]. LncRNA UCA1 exerts oncogenic functions in promoting NSCLC progression by miR-193a-3p sponge and indicated a novel UCA1–miR-193a-3p–ERBB4 signaling pathway regulatory network [24]. Consequently, accumulating evidence suggested the possibility of elucidating the function how lncRNAs interact with miRNA as ceRNAs contributes to pathogenic process of NSCLC. Therefore, we comprehensively analyze the interaction of lncRNAs with miRNA to act as RNA pools to identify the lncRNA-miRNA-mRNA regulatory network mediated in the ceRNA network in NSCLC.

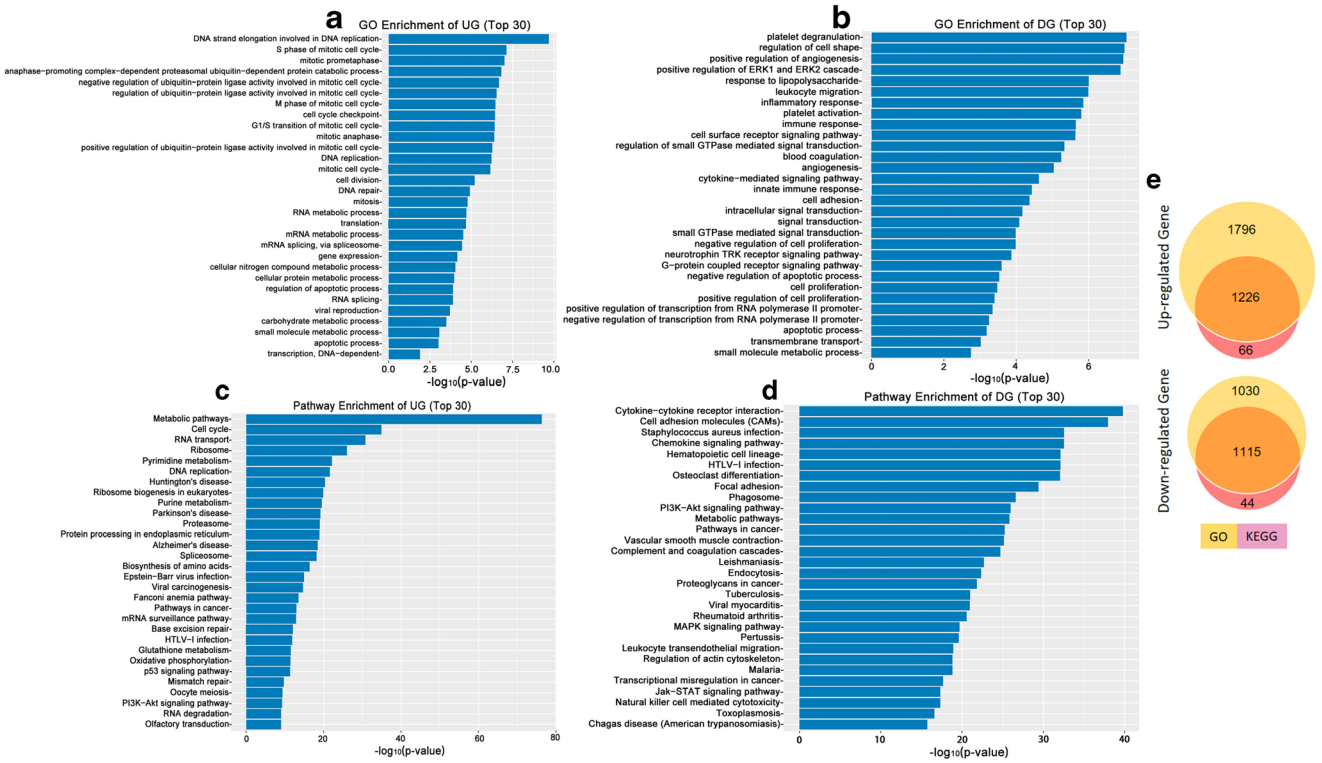

Fig. 4 Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment of different expressed genes. (**a, b**) The top 30 GO of up-regulated and down-regulated genes that related to biological process (BP), cellular component(CC) and molecular function(MF); (**c, d**) The top 30 KEGG pathway of up-regulated and down-regulated genes that correlated with cellular process, metabolism, environmental information processing and organismal systems. (**e**) Venn diagram showing the overlapped genes (1226 up-regulated genes; 1115 down-regulated genes) according to GO and KEGG pathway enrichment. FC > 1.5 or FC < 0.7; UG, up-regulated Gene; DG, Down-regulated Gene. The *P* value denotes the significance of Go terms or KEGG pathway enrichment in the UG and DG, $P < 0.05$ was recommended as the cut-off value

We performed lncRNA, miRNA and mRNA expression profiling in LUAD, LUSC tissues and adjacent normal lung tissues. The microarray expression profiles showed a total of 4478 DEmRNAs, 526 DElncRNAs and 75 DEmiRNAs in LUAD, whereas 6236 DEmRNAs, 843 DElncRNAs and 117 DEmiRNAs in LUSC. Venn diagram analysis showed 1711 mRNAs, 127 lncRNAs and 25 miRNAs were common up-regulated, 1768 mRNAs, 213 lncRNAs and 5 miRNAs were shared down-regulated in both LUAD and LUSC. Subsequently GO analysis and KEGG analysis were implemented to illustrate both the unique and shared mRNA, lncRNAs and miRNAs in LUSC and LUAD. The terms of up-regulated genes in GO analysis results mainly associated with cell proliferation such as DNA strand elongation, mitotic cell cycle, M phase of mitotic cell cycle and cell division while the down-regulation mRNA were enriched in platelet degranulation, regulation of cell shape and positive regulation of angiogenesis. Subsequent KEGG analysis showed upregulated mRNA mainly involved in metabolism pathway, Cell cycle, RNA transport and ribosomes while downregulated mRNA mainly associated with cytokine-cytokine receptor interaction, cell adhesion molecules and chemokine signaling pathway. Based on the GO and KEGG analysis, 1226 up-regulated mRNAs and 1115 down-regulated genes were overlapped, which could be the potential prognosis biomarkers in both LUSC and LUAD.

After co-expression analysis the dysregulated lncRNA-miRNA-mRNA ceRNA network was constructed. Totally 62 dysregulated lncRNAs (7 up-regulated and 55 down-regulated lncRNAs), 28 differentially expressed mRNAs (5 up-regulated and 23 down-regulated mRNAs), and 18 dysregulated miRNAs (14 up-regulated and 4 down-regulated miRNAs), were distinctly enriched in ceRNA network and 208 connections were established. Furthermore, we observed two DElncRNAs, LINC00641 and AC004947.2, the most remarkable lncRNAs in the network, strongly suggesting that the two lncRNAs may serve as key oncogenes as well as prognostic markers in lung cancer progression. After detailed analysis of the regulatory network, we identified LncRNA LINC00641 was the center of the ceRNA network that regulate the expressions of target genes by interacting with miR-6860 and miR-1285-3p. LINC00641 is a novel lncRNA whose function remains unclear. In recent studies down-regulation of LINC00641 in bladder cancer was related to poor diagnosis, and it could compete for miR-197-3p to promote KLF10 expression, thus accelerated proliferation,

migration and invasion of bladder cancer [25]. LINC00641 could act as a competitive endogenous RNA of miR-153-3p under nutrition deprivation stress in intervertebral disc degeneration [26]. In our study, it can be noticed that LINC00641 with high-expression can compete with miR-1285-3p to modulate the target genes PIGA, AHCYL1, ATP1B2, while interact with miR-6860 to regulate the expression of MAP3K3, SHRK004M and FGD3 involved in the network.

Besides, AC004947.2 has the second largest interactions in the network. AC004947.2 was a new lncRNA whose roles were totally undefined. Recent integrated analysis of AC004947.2 in lncRNA expression profiles with endometrial cancer, lung cardinocarcinoma and acute T cell–mediated rejection of renal allografts suggested it might be a novel biomarker to predict poorer prognosis in these diseases [27–29]. Also, AC004947.2 regulate the expression of MAP3K3, SHRK004M and FGD3 by competing for miR-6860. In the ceRNA network, another three lncRNAs, lncRNA-109,642.1, lncRNA-MAGI2-AS3 and lncRNA-SHANK3 can also serve as remarkable biomarkers for lung cancer and these RNAs were the first time to determine as biomarkers for diagnosis and prognosis, providing reasonable possiblity in future clinical treatment of lung cancer.

Unlike miRNAs which have been well investigated in various human diseases including lung cancer, merely a small amount of lncRNAs have been clearly defined and characterized. In our research, the ceRNA network was established to integrated illustrate interrelation of lncRNAs and miRNAs. Our study identified 18 miRNAs, among which the most significant ones were miR-6860, miR-1285-3p, miR-767-3p and miR-7974. Circulating miR-1285-3p with downregulation was predictive of poor therapeutic effects in hepatocellular carcinoma while upregulation in chronic heart failure correlated to disease severity [30, 31]. miR-767-3p was reported to be a potential serum biomarker in colorectal cancer, malignant pleural mesothelioma, non-invasive lung cancer and glioblastoma [32–35]. Our study showed high-expression of miR-1285-3p and miR-767-3p in lung cancer and further researches are required to annotate its impact on tumorigenesis, development and progression.

Many mRNAs involved in the new ceRNA network were recently considered as pivotal oncogenes in lung cancer development and progression. MAP3K3 was evidenced to be a

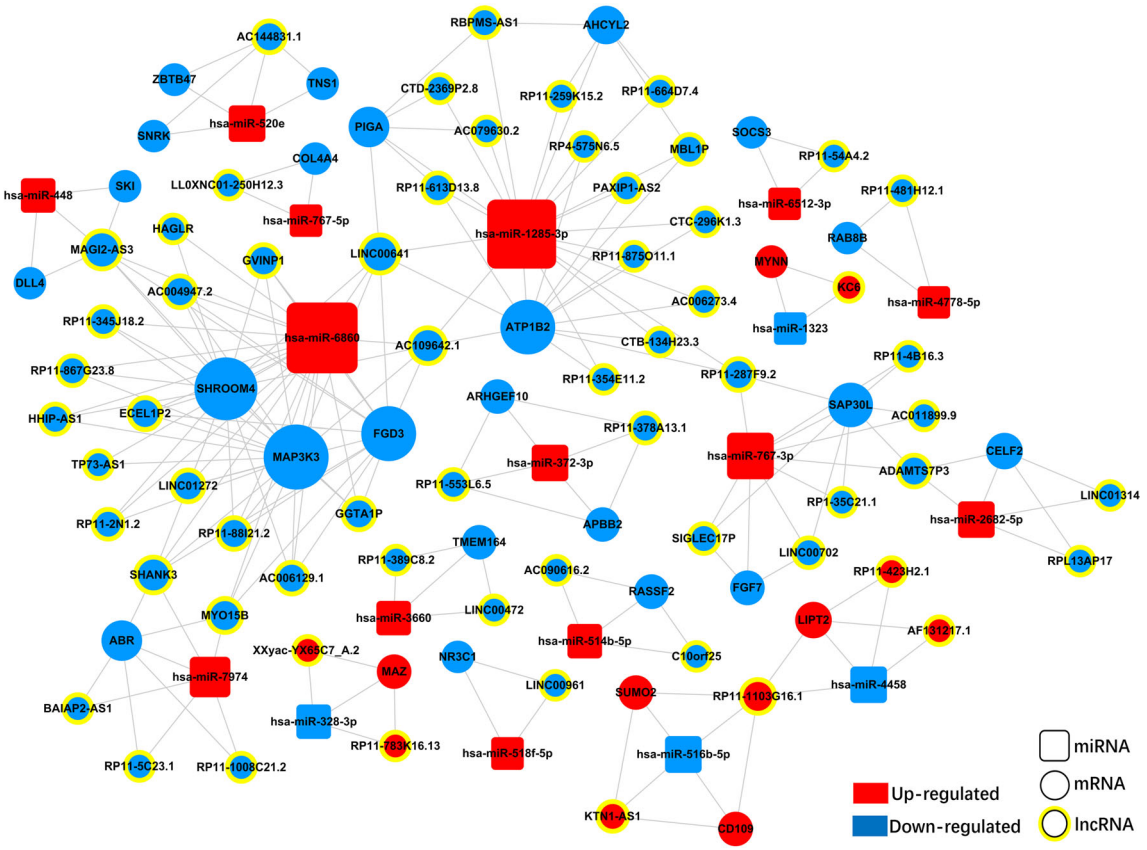

Fig. 5 Co-expression network analyses. Co-expression network analyses were performed using Cytoscape 3.5.1 software to characterize the 62 differentially expressed lncRNAs (7 up-regulated and 55 down-regulated), 28 differentially expressed mRNAs (5 up-regulated and 23 down-regulated) and 18 miRNAs (14 up-regulated and 4 down-regulated) based on the different expressed data. Red nodes indicate upregulated, and blue nodes indicated down-regulated. Square nodes denote miRNAs, circular nodes designate mRNAs, and circular nodes with yellow edge indicate lncRNAs. Degree ≥2, Pearson Correlation: $r > 0.7$ or < -0.7, $p < 0.05$, Pearson correlation coefficient ≥ 0.99

tumor suppressor in progression of NSCLC model [15]. FGD3 was reported to be predictive of better prognosis in breast cancer [18, 36]. Genetic variation of ATP1B2 was considered to be associated with ER negative breast cancer [19]. KEGG pathway revealed several crucial tumor related pathways involved in the ceRNA network of LC, including metabolic pathway, cell cycle, cytokine-cytokine receptor interaction pathway and cell adhesion molecules (CAMs). Recent studies have identified cytokine-cytokine receptor interaction modulate metabolic events to promote tumorigenesis [37]. LncRNAs in the network, LINC00641, AC004947.2, may take specific effects in LC through modulation of the four pathways.

Currently, lncRNA-miRNA-mRNA network is rapidly gaining attentions and much similar work has been done in lung cancer. Li X et al. and Sun J et al. constructed ceRNA network in lung adenocarcinoma and Ning P et al. established the network in lung squamous cell carcinoma [38–40]. Our study combined the clinical data of LUSC and LUAD to comprehensively analyze RNA biomarkers for diagnosis and treatment in LC.

In summary, our investigation analyzed non-coding and coding RNA expression profiling between normal adjacent tissues and cancer tissues from TCGA database. Further, we performed weighted gene co-expression network analysis and constructed lncRNA-miRNA-mRNA ceRNA crosstalk network by DERNAs to reveal a new regulatory mechanism for further studies of LC.

Acknowledgements This study is partly supported by Special Fund for Basic Research of Chen Xiaoping Foundation for Scientific and Technological Development (CXPJJH1180008).

Authors' Contributions Ruolan Dong conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper. Jiawei Liu performed the experiments, prepared Figures and/or tables. Wei Sun and Wei Ping analyzed the data, reviewed drafts of the paper.

Data Availability The datasets used and/or analyzed during the current study are available from TCGA database.

Compliance with Ethical Standards

Ethics Approval and Consent to Participate According to the TCGA ethic committee, additional approval was not required.

Consent for Publication Not applicable.

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Siegel RL, Miller KD, Jemal A (2017) Cancer statistics, 2017. CA Cancer J Clin 67(1):7–30. https://doi.org/10.3322/caac.21387
2. Goldstraw P, Ball D, Jett JR, Le Chevalier T, Lim E, Nicholson AG, Shepherd FA (2011) Non-small-cell lung cancer. Lancet 378(9804): 1727–1740. https://doi.org/10.1016/s0140-6736(10)62101-0
3. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, Sougnez C, Greulich H, Muzny DM, Morgan MB, Fulton L, Fulton RS, Zhang Q, Wendl MC, Lawrence MS, Larson DE, Chen K, Dooling DJ, Sabo A, Hawes AC, Shen H, Jhangiani SN, Lewis LR, Hall O, Zhu Y, Mathew T, Ren Y, Yao J, Scherer SE, Clerc K, Metcalf GA, Ng B, Milosavljevic A, Gonzalez-Garay ML, Osborne JR, Meyer R, Shi X, Tang Y, Koboldt DC, Lin L, Abbott R, Miner TL, Pohl C, Fewell G, Haipek C, Schmidt H, Dunford-Shore BH, Kraja A, Crosby SD, Sawyer CS, Vickery T, Sander S, Robinson J, Winckler W, Baldwin J, Chirieac LR, Dutt A, Fennell T, Hanna M, Johnson BE, Onofrio RC, Thomas RK, Tonon G, Weir BA, Zhao X, Ziaugra L, Zody MC, Giordano T, Orringer MB, Roth JA, Spitz MR, Wistuba II, Ozenberger B, Good PJ, Chang AC, Beer DG, Watson MA, Ladanyi M, Broderick S, Yoshizawa A, Travis WD, Pao W, Province MA, Weinstock GM, Varmus HE, Gabriel SB, Lander ES, Gibbs RA, Meyerson M, Wilson RK (2008) Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455:1069–1075. https://doi.org/10.1038/nature07423. https://www.nature.com/articles/nature07423#supplementary-information
4. Wistuba II, Lam S, Behrens C, Virmani AK, Fong KM, LeRiche J, Samet JM, Srivastava S, Minna JD, Gazdar AF (1997) Molecular damage in the bronchial epithelium of current and former smokers. J Natl Cancer Inst 89(18):1366–1373
5. Felekkis K, Touvana E, Stefanou C, Deltas C (2010) microRNAs: a newly described class of encoded molecules that play a role in health and disease. Hippokratia 14(4):236–240
6. Ponting CP, Oliver PL, Reik W (2009) Evolution and functions of long noncoding RNAs. Cell 136(4):629–641. https://doi.org/10.1016/j.cell.2009.02.006
7. Dykes IM, Emanueli C (2017) Transcriptional and post-transcriptional gene regulation by Long non-coding RNA. Genomics, proteomics & bioinformatics 15(3):177–186. https://doi.org/10.1016/j.gpb.2016.12.005
8. Prensner JR, Chinnaiyan AM (2011) The emergence of lncRNAs in cancer biology. Cancer discovery 1(5):391–407. https://doi.org/10.1158/2159-8290.Cd-11-0209
9. Rinn JL, Chang HY (2012) Genome regulation by long noncoding RNAs. Annu Rev Biochem 81:145–166. https://doi.org/10.1146/annurev-biochem-051410-092902
10. Kristensen LS, Hansen TB, Veno MT, Kjems J (2018) Circular RNAs in cancer: opportunities and challenges in the field. Oncogene 37(5):555–565. https://doi.org/10.1038/onc.2017.361
11. Song YX, Sun JX, Zhao JH, Yang YC, Shi JX, Wu ZH, Chen XW, Gao P, Miao ZF, Wang ZN (2017) Non-coding RNAs participate in the regulatory network of CLDN4 via ceRNA mediated miRNA evasion. Nat Commun 8(1):289. https://doi.org/10.1038/s41467-017-00304-1
12. Wang Y, Liu Z, Yao B, Li Q, Wang L, Wang C, Dou C, Xu M, Liu Q, Tu KJMC (2017) Long non-coding RNA CASC2 suppresses epithelial-mesenchymal transition of hepatocellular carcinoma cells through CASC2/miR-367/FBXW7 axis. Mol Cancer 16(1):123
13. Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11(10):R106. https://doi.org/10.1186/gb-2010-11-10-r106
14. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 41(Database issue): D808–D815. https://doi.org/10.1093/nar/gks1094
15. Zhao L, Ni X, Zhao L, Zhang Y, Jin D, Yin W, Wang D, Zhang W (2018) MiroRNA-188 acts as tumor suppressor in non-small-cell lung Cancer by targeting MAP3K3. Mol Pharm 15(4):1682–1689. https://doi.org/10.1021/acs.molpharmaceut.8b00071

16. Wei J, Dong Y, Lin T, Pang L, Yan R, Liang W, Jiang J, Gang C, Wen JZ, Yuan XJHP (2016) MAP3K3 overexpression is associated with poor survival in ovarian carcinoma. Hum Pathol 50:162–169
17. He Y, Wang L, Liu W, Zhong J, Bai S, Wang Z, Thomas DG, Lin J, Reddy RM, Ramnath N, Carrott PW, Lynch WR, Orringer MB, Chang AC, Beer DG, Chen G (2015) MAP3K3 expression in tumor cells and tumor-infiltrating lymphocytes is correlated with favorable patient survival in lung cancer. Sci Rep 5:11471–11413. https://doi.org/10.1038/srep11471
18. Ou YT, Cheng WY, Zheng T, Maurer MA, Anastassiou DJCEBP (2014) Breast cancer prognostic biomarker using attractor metagenes and the FGD3-SUSD3 metagene. Cancer Epidemiol Biomark Prev 23(12):2850–2856
19. Garcia-Closas MJIJC (2010) Common genetic variation in TP53 and its flanking genes, WDR79 and ATP1B2, and susceptibility to breast cancer. Int J Cancer 121(11):2532–2538
20. Vijayvergia N, Shah PC, Denlinger CS (2015) Survivorship in non-small cell lung Cancer: challenges faced and steps forward. J Natl Compr Cancer Netw 13(9):1151–1161
21. Yang Y-R, Zang S-Z, Zhong C-L, Li Y-X, Zhao S-S, X-JJIjoc F (2014) Increased expression of the lncRNA PVT1 promotes tumorigenesis in non-small cell lung cancer. Int J Clin Exp Pathol 7(10):6929–6935
22. Wang C, Han C, Zhang Y, Liu F (2018) LncRNA PVT1 regulate expression of HIF1alpha via functioning as ceRNA for miR199a5p in nonsmall cell lung cancer under hypoxia. Mol Med Rep 17(1):1105–1110. https://doi.org/10.3892/mmr.2017.7962
23. Lu K-h, Li W, X-h L, Sun M, M-l Z, Wu W-q, Xie W-p, Y-yJBc H (2013) Long non-coding RNA MEG3 inhibits NSCLC cells proliferation and induces apoptosis by affecting p53 expression. BMC Cancer 13(1):461
24. Nie W, Ge HJ, Yang XQ, Sun X, Huang H, Tao X, Chen WS, Li BJ (2016) LncRNA-UCA1 exerts oncogenic functions in non-small cell lung cancer by targeting miR-193a-3p. Cancer Lett 371(1):99–106
25. Li Z, Hong S, Liu Z (2018) LncRNA LINC00641 predicts prognosis and inhibits bladder cancer progression through miR-197-3p/KLF10/PTEN/PI3K/AKT cascade. Biochem Biophys Res Commun 503(3):1825–1829. https://doi.org/10.1016/j.bbrc.2018.07.120
26. Wang XB, Wang H, Long HQ, Li DY, Zheng X (2018) LINC00641 regulates autophagy and intervertebral disc degeneration by acting as a competitive endogenous RNA of miR-153-3p under nutrition deprivation stress. J Cell Physiol. https://doi.org/10.1002/jcp.27466
27. Lorenzen JM, Schauerte C, Kölling M, Hübner A, Knapp M, Haller H, Thum T (2015) Long Noncoding RNAs in Urine Are Detectable and May Enable Early Detection of Acute T Cell–Mediated Rejection of Renal Allografts. Clin Chem. https://doi.org/10.1373/clinchem.2015.243600
28. Shi X, Tan H, Le X, Xian H, Li X, Huang K, Luo VY, Liu Y, Wu Z, Mo H, Chen AM, Liang Y, Zhang J (2018) An expression signature model to predict lung adenocarcinoma-specific survival. Cancer Manag Res 10:3717
29. Zhou M, Zhang Z, Zhao H, Bao S, Sun J (2018) A novel lncRNA-focus expression signature for survival prediction in endometrial carcinoma. BMC Cancer 18(1):39
30. Liu J, Yan J, Zhou C, Ma Q, Jin Q, Yang Z (2015) miR-1285-3p acts as a potential tumor suppressor miRNA via downregulating JUN expression in hepatocellular carcinoma. Tumour Biol 36(1):219–225. https://doi.org/10.1007/s13277-014-2622-5
31. Li H, Fan J, Yin Z, Wang F, Chen C, Wang DWJO (2016) Identification of cardiac-related circulating microRNA profile in human chronic heart failure. Oncotarget 7(1):33–45
32. Keller A, Leidinger P, Gislefoss R, Haugen A, Langseth H, Staehler P, Lenhof H-P, Meese E (2011) Stable serum miRNA profiles as potential tool for non-invasive lung cancer diagnosis. RNA Biol 8(3):506–516
33. Kirschner MB, Cheng YY, Badrian B, Kao SC, Creaney J, Edelman JJ, Armstrong NJ, Vallely MP, Musk AW, Robinson BW, McCaughan BC, Klebe S, Mutsaers SE, van Zandwijk N, Reid G (2012) Increased circulating miR-625-3p: a potential biomarker for patients with malignant pleural mesothelioma. J Thorac Oncol 7(7):1184–1191
34. Kreth S, Limbeck E, Hinske LC, Schütz SV, Thon N, Hoefig K, Egensperger R, Kreth FW (2013) In human glioblastomas transcript elongation by alternative polyadenylation and miRNA targeting is a potent mechanism of MGMT silencing. Acta Neuropathol 125(5):671–681
35. H-s LIU, Y-y MA, H-sJT XIAO (2012) The diagnostic value of serum microRNAs including miR-129-3p, miR-767-3p and miR-877* for colorectal cancer. Tumor 32(1):42–48
36. Cheng WY, Ou Yang TH, Anastassiou D (2013) Development of a prognostic model for breast cancer survival in an open challenge environment. Sci Transl Med 5(181):181ra150. https://doi.org/10.1126/scitranslmed.3005974
37. Gao D, Cazares LH, Fish EN (2017) CCL5-CCR5 interactions modulate metabolic events during tumor onset to promote tumorigenesis. BMC Cancer 17(1):834. https://doi.org/10.1186/s12885-017-3817-0
38. 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. Oncol Lett 16(5):5697–5708. https://doi.org/10.3892/ol.2018.9336
39. Ning P, Wu Z, Hu A, Li X, He J, Gong X, Xia Y, Shang Y, Bian H (2018) Integrated genomic analyses of lung squamous cell carcinoma for identification of a possible competitive endogenous RNA network by means of TCGA datasets. PeerJ 6:e4254. https://doi.org/10.7717/peerj.4254
40. 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. Int J Oncol 49(5):2023–2036. https://doi.org/10.3892/ijo.2016.3716