

RESEARCH ARTICLE

Open Access



# Identification of LINC02310 as an enhancer in lung adenocarcinoma and investigation of its regulatory network via comprehensive analyses

Wenyuan Zhao<sup>1,2</sup>, Jun Wang<sup>1</sup>, Qingxi Luo<sup>1</sup>, Wei Peng<sup>1</sup>, Bin Li<sup>2</sup>, Lei Wang<sup>3</sup>, Chunfang Zhang<sup>1</sup> and Chaojun Duan<sup>1,2,3,4\*</sup>

## Abstract

**Background:** Lung adenocarcinoma (LADC) is a major subtype of non-small cell lung cancer and has one of the highest mortality rates. An increasing number of long non-coding RNAs (LncRNAs) were reported to be associated with the occurrence and progression of LADC. Thus, it is necessary and reasonable to find new prognostic biomarkers for LADC among LncRNAs.

**Methods:** Differential expression analysis, survival analysis, PCR experiments and clinical feature analysis were performed to screen out the LncRNA which was significantly related to LADC. Its role in LADC was verified by CCK-8 assay and colony. Furthermore, competing endogenous RNA (ceRNA) regulatory network construction, enrichment analysis and protein–protein interaction (PPI) network construction were performed to investigate the downstream regulatory network of the selected LncRNA.

**Results:** A total of 2431 differentially expressed LncRNAs (DELncRNAs) and 2227 differentially expressed mRNAs (DEmRNAs) were from The Cancer Genome Atlas database. Survival analysis results indicated that Inc-YARS2-5, Inc-NPR3-2 and LINC02310 were significantly related to overall survival. Their overexpression indicated poor prognostic. PCR experiments and clinical feature analysis suggested that LINC02310 was significantly correlated with TNM-stage and T-stage. CCK-8 assay and colony formation assay demonstrated that LINC02310 acted as an enhancer in LADC. In addition, 3 targeted miRNAs of LINC02310 and 414 downstream DEmRNAs were predicted. The downstream DEmRNAs were then enriched in 405 Gene Ontology terms and 11 Kyoto Encyclopedia of Genes and Genomes pathways, which revealed their potential functions and mechanisms. The PPI network showed the interactions among the downstream DEmRNAs.

**Conclusions:** This study verified LINC02310 as an enhancer in LADC and performed comprehensive analyses on its downstream regulatory network, which might benefit LADC prognoses and therapies.

**Keywords:** LADC, LncRNA, TCGA, ceRNA regulatory network, PPI

## Background

NSCLC is a major cause of cancer-derived death worldwide, accounting for 85% of all lung cancers [1]. It can be classified into four subtypes: Lung adenocarcinoma (LADC), lung large cell carcinoma (LCLC), lung squamous cell carcinoma (LSCC) and lung neuroendocrine tumor (Lung NET) [2]. LADC, as a major type of NSCLC, has one of the highest mortality rates. Although

\*Correspondence: duancjxy@126.com

<sup>1</sup> Department of Thoracic Surgery, Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China

Full list of author information is available at the end of the article



LADC treatments, including surgery, chemotherapy, radiotherapy and other molecular target therapies have been improved, the long-term survival remains dim. Therefore, it is quite necessary to find new biomarkers for precise LADC treatments and predicting prognostic.

LncRNAs, which were previously considered as transcriptional noise, have been found to participate in regulatory biological processes such as transcriptional regulation, RNA processing and chromatin interaction. LncRNAs also participate in the stability and translation of mRNA and influence cell signaling [3, 4]. In addition, LncRNAs have been reported to promote LADC cell proliferation, invasion, metastasis and tumor progression, such as LncRNA NEAT1 [5], DGCR5 [6], H19 [7]. Moreover, previous studies showed that LncRNAs play an important role in cancer drug resistance [8–11]. Therefore, it is reasonable to find certain new prognostic biomarkers for LADC among LncRNAs.

In addition, LncRNAs were widely considered to function as ceRNAs, which regulate mRNA expression by sponging miRNAs in cancers [12–14]. These ceRNA regulatory networks constructed by coding and non-coding transcripts can result in several diseases when dysregulated [15]. Moreover, researches have demonstrated the significance of ceRNA regulatory network during the oncogenesis and progression of various cancers [16–19]. Some researchers have further identified potential prognostic mRNAs for LADC based on ceRNA regulatory network [20–29].

This study identified LINC02310 as an enhancer in LADC and investigated its regulatory network through bioinformatic analyses and experimental validation. Firstly, differential expression analyses based on the expression data from TCGA were performed to screen DELncRNAs and DEMRNAs. Next, the correlation p-values of overall survival and DELncRNA expression level in LADC patients were estimated by survival analysis. The top-3 significant DELncRNAs associated with overall survival were selected for further analysis. Then, PCR experiments and the relevant clinical feature analyses were conducted to evaluate the clinical correlation of the three DELncRNAs in LADC, among which LINC02310 was the most meaningful. Moreover, CCK-8 assays and colony formation assays were performed to demonstrate that LINC02310 acted as an enhancer in LADC. Furthermore, ceRNA regulatory network was constructed to find the targeted miRNAs and downstream DEMRNAs. In addition, GO functional annotation and KEGG pathway analyses of the downstream DEMRNAs provided a comprehension of their functions and mechanisms. Finally, PPI network was constructed to visualize the interactions among these downstream DEMRNAs.

## Methods

### LADC data sets and preprocessing

The LADC data sets, including RNA expression quantification data and the relevant clinical data, were downloaded from TCGA database. The expression quantification data of mRNAs and LncRNAs were obtained from HTSeq-counts files on Illumina platform. A total of 533 LADC samples and 59 adjacent normal samples were collected. The IDs of mRNAs and LncRNAs were converted using GENCODE Human Release 31.

### Differential expression analysis

DELncRNAs and DEMRNAs between LADC and adjacent normal samples were analyzed using the R package DESeq2 (version 1.22.2) [30]. Firstly, the samples were divided into two groups: LADC and adjacent normal samples. Next, the genes whose sum of expression is less than 20 were removed since too small gene expression can lead to large calculation errors. The DESeq function in the R package DESeq2 was then used to analyze the differential expression of the remaining genes and obtain the result dataset. Subsequently, the LncRNAs in the result data set were screened to obtain DELncRNAs according to the criteria of  $|\log_2$  fold change (LFC)| > 1.5 and false discovery rate (FDR, or adjusted p-value) < 0.05. As for DEMRNAs, the criteria were  $|\text{LFC}| > 2$  and  $\text{FDR} < 0.01$ . Finally, the R packages ggplot2 (version 3.1.1) and pheatmap (version 1.0.12) were used to plot the volcano plots and heatmaps of DELncRNAs and DEMRNAs.

### Survival analysis

To begin with, the samples with missing survival status information were removed from the expression data of DELncRNAs. Next, the R package DESeq2 was used to normalize the expression data of DELncRNAs. Next, the LADC samples were divided into two groups: high expression (normalized expression > median expression) and low expression (normalized expression < median expression) according to the expression level of the DELncRNAs in these samples. The R package survival (version 2.44-1.1) was then used to build the survival information dataset and to perform the log-rank test to calculate the correlation p-value of overall survival and DELncRNA expression level. The top-3 significant DELncRNAs related to overall survival were selected for further investigation. The survival curves of the three DELncRNAs were calculated with Kaplan–Meier (KM) method and visualized using the R package survminer (version 0.4.4). Lastly, the significance was calculated again by performing Fisher's accurate test and Pearson's chi-squared test to validate the correlation between the overall survival and gene expression level.

**PCR experiments**

LADC and adjacent normal tissues were obtained from 22 patients who suffered from LADC. These patients were treated in Xiangya Hospital, Central South University in 2015. All subjects had written consents, and the protocol was approved by the ethics committee of Xiangya Hospital (No. 201503303). All procedures conducted in the study were in accordance with the Ethics Standards Institutions Research Committee and the 1964 Helsinki Declaration and its subsequent amendments or similar ethical standards. All patients had no history of other malignancies and never received radiotherapy or chemotherapy. All tissue samples were immediately frozen in liquid nitrogen and stored at - 80 °C. All tumors and matched normal tissues were confirmed by two pathologists. Total RNA was extracted using the Trizol Reagent kit (Invitrogen) and was reversely transcribed into cDNAs by using HiScript II Q Select RT SuperMix for qPCR (Vazyme). Real-time PCR was performed on

edition, which is the latest revision proposed by International Association for the Study of Lung Cancer (IASLC) [31, 32]. T-stage represented the size and location of tumor, N-stage reflected the spread of lung cancer in lymph nodes, and M-stage indicated whether the cancer cell had metastasized to a distant site. They were explained detailly in Table 1. TNM-stage gave an overall evaluation of the tumor progression via integrating the information given by T-stage, N-stage and M-stage. Its stage I, II/III and IV represented the early, middle and advanced stages of tumor, respectively.

The clinical feature analyses of the top-3 significant DELncRNAs were conducted based on the expression levels identified by PCR experiments and the clinical features of the 22 LADC patients. Firstly, the CT values of the experimental genes (DELncRNAs) and reference genes in LADC and adjacent normal samples were obtained from the PCR experimental results. Then, the ΔCT values were calculated as follows.

$$\Delta CT_{LADC \text{ sample}} = CT_{LADC \text{ sample, experimental gene}} - CT_{LADC \text{ sample, reference gene}}$$

$$\Delta CT_{\text{adjacent normal sample}} = CT_{\text{adjacent normal sample, experimental gene}} - CT_{\text{adjacent normal sample, reference gene}}$$

the cDNA templates using specific primers (TSINGKE) and the GeneCopoeia BlazeTaq™ SYBR® Green qPCR mix (GeneCopoeia). The LncRNAs relative expression levels were calculated as a ratio normalized to U6 expression. Comparative quantification was calculated with the 2<sup>-ΔΔCt</sup> method. The primers sequences used in this study are listed. lnc-YARS2-5: F, 5'- ggtaccagaagcagcacct-3'; and R, 5'- aaaagaactcggccaagctc-3'. lnc-NPR3-2: F, 5'- aagcaagcatactcgtccct-3'; and R, 5'- gagccaagacgtagaggagg-3'. LINC02310: F, 5'-gaggaggtgcttgcttctc-3'; and R, 5'-atgaaccgagtcctggagtc-3'.

**Clinical feature analysis**

The TNM-staging system has become the standard method for clinicians and medical scientists to stage malignant tumors. In this paper, we adopted its 8th

The ΔΔCT values were calculated by

$$\Delta\Delta CT = \Delta CT_{LADC \text{ sample}} - \Delta CT_{\text{adjacent normal sample}}$$

Let LFC = -ΔCT, and the LADC patients with |LFC| ≤ 0.5 were removed. The remaining patients were separated into two groups: high expression (LFC > 0.5) and low expression (LFC < -0.5) according to the expression level of DELncRNAs in their tissues. Finally, the significance was calculated by performing Fisher's accurate test and Pearson's chi-squared test to evaluate the correlation between the expression level of each of the three DELncRNAs and each clinical feature.

From the above analyses, LINC02310 can be screened out. Therefore, further validation of the clinical relevance of LINC02310 were performed based on the expression data and clinical data from TCGA. Firstly, LINC02310 expression

**Table 1 Descriptions of T-stage, N-stage and M-stage**

| T-stage |                          | N-stage |  | M-stage |                                    |
|---------|--------------------------|---------|--|---------|------------------------------------|
| Stage   | Description              | Stage   | Description  | Stage   | Description                        |
| T1      | tumor size ≤ 3 cm        | N0      | No cancer cells near the lymph nodes   | M0      | Cancer cells have not metastasized |
| T2      | 3 cm < tumor size ≤ 5 cm | N1      | Cancer cells have spread to the parabranchial nodes or the lymph nodes in the lungs of the primary tumor | M1      | Cancer cells have metastasized     |
| T3      | 5 cm < tumor size ≤ 7 cm | N2      | Cancer cells have spread to the mediastinal lymph nodes  | MX      | Cancer metastasis is unknown       |
| T4      | tumor size > 7 cm        | N3      | Cancer cells have spread to lymph nodes without the primary tumor  |         |                                    |
| TX      | undetectable             | NX      | Lymph nodes near the tumor were not tested   |         |                                    |

data were extracted from TCGA. Then, the LADC samples were divided into two groups: high expression (LINC02310 expression > median expression) and low expression (LINC02310 expression < median expression). Lastly, the correlation significance between LINC02310 expression level and each clinical feature was evaluated by performing Fisher's accurate test and Pearson's chi-squared test.

#### CCK-8 assay and colony formation assay

Two human lung cancer cell lines (H1299, PC-9) were used in the CCK-8 assay and colony formation assay since they can be efficiently infected according to our previous studies. They were purchased from Chinese Academy of Sciences Cell Bank (Shanghai, China).

OE: the primers G0157014-1F (AGGATCGCTAGCGCTACCGGACTCAGATCTCGAGAGTCTCCTCTTG CAGATCAGATAACCACC) and G0157014-1R (TACCCGGTGAATTATCTAGAGTCGCGGGATCCAGATAC CTA AAAGGCACAAATCCTTCTG) were synthesized. They were then cloned into pLVX-Puro at XhoI-BamHI site to obtain the LINC02310 overexpression vector, pLVX-Puro-LINC02310, which was transfected into the cells.

NC: a random sequence was cloned into pLVX-Puro at XhoI-BamHI site to obtain the control vector, which was then transfected into the cells.

Si-02310: three small interfering RNAs (siRNAs) of LINC02310 were synthesized and transfected into the cells to down-regulate LINC02310 expression.

Si-NC: A control sequence was transfected into the cells.

CCK-8 assay: 100  $\mu$ L of OE, NC, si-02310 and si-NC of H1299 and PC-9 ( $1 \times 10^3$ /mL) were transplanted in 96-well plate. Next, 10  $\mu$ L of CCK-8 solution (Dojindo Laboratories, Kumamoto, Japan) was added to each well. Then, 40 min of incubation at 37 °C was involved. Finally, the absorbance was measured at 450 nm (OD450) by automatic microplate reader (Tecan, NANOQUANT, Switzerland) on days 0, 1, 2, 3 and 4, respectively. The experiment was repeated three times.

Colony formation assay: OE, NC, si-02310 and si-NC of H1299 ( $1 \times 10^3$ ) were transplanted into 6-well plate and then incubated at 37 °C for 10 days. Colonies were dyed with dyeing solution containing 0.1% crystal violet and 20% methanol. Cell colonies were then counted and analyzed.

#### CeRNA regulatory network construction

The lncRNASNP2 website [33, 34] was used to predict the targeted miRNAs of LINC02310. The targeted miRNAs whose reads of exon model per million mapped reads (RPM)  $\geq 1$  were screened out. The starBase database (version 3.0) [35] was used to predict the downstream mRNAs of the selected targeted miRNAs. The prediction programs TargetScan, microT, miRanda, miRmap, PITA, RNA22 and PicTar were integrated into starBase. Then,

these downstream mRNAs were intersected with the DEmRNAs to get the downstream DEmRNAs. Finally, the ceRNA (LINC02310-targeted miRNAs-downstream DEmRNAs) regulatory network was constructed and visualized by the software Cytoscape (version 3.7.1).

#### Enrichment analysis

The enrichment analysis consisted of Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. It was performed using the R packages clusterProfiler (version 3.10.1) [36] and org.Hs.eg.db (version 3.7.0), which provided analysis functions and human gene annotation information, respectively. They were used to analyze and visualize the functional profiles of the predicted downstream DEmRNAs.

The GO functional annotation was performed, including molecular function (MF), biological process (BP) and cellular component (CC) [37]. Firstly, the gene IDs of downstream DEmRNAs were listed. The enrichGO function was then used to generate the enrichment result data set of MF, BP and CC, respectively. Finally, the barplot function was used to visualize the result.

In addition, the KEGG pathway analysis was conducted. Firstly, the gene IDs of downstream DEmRNAs were converted into a format suitable for KEGG pathway analysis according to HGNC database. Secondly, the enrichKEGG function was used to generate KEGG pathway analysis result data set. Lastly, the result was visualized via barplot function.

#### PPI network

The PPI network was constructed using the STRING database (version 11.0) and Cytoscape according to the following steps. First, the gene names of downstream DEmRNAs were uploaded to the STRING database to build a PPI dataset. In the dataset, the relevance between every two proteins (nodes) were evaluated by a combined score. The score ranges from 0 to 1, and a higher score represents a closer interaction. Second, the interactions whose combined score > 0.7 were screened out and downloaded. Finally, the downloaded interactions were divided into different modules using MCODE (version 1.5, a plugin for Cytoscape). The module with the highest average combined score was selected to construct the PPI network, which was visualized by Cytoscape.

## Results

### DELncRNAs and DEmRNAs in LADC

The expression data of 12,459 lncRNAs and 19,647 mRNAs in LADC and adjacent normal samples were extracted from TCGA database. The differential expression analysis results of these lncRNAs and mRNAs were

visualized in the volcano plots (Fig. 1), respectively. The red dots and blue dots represent the significantly high-expressed RNAs and low-expressed RNAs, respectively, and the gray dots are the remaining RNAs. As a result, 2431 DELncRNAs (Additional file 1: Table S1, 1931 high-expressed and 500 low-expressed) and 2227 DEMRNAs (Additional file 2: Table S2, high-expressed and 570 low-expressed) were identified. The expression of the top-50 DELncRNAs and DEMRNAs in LADC and adjacent normal samples were visualized in the heatmaps (Fig. 2), respectively. Significant difference can be seen in the expression of each RNA in the two groups.

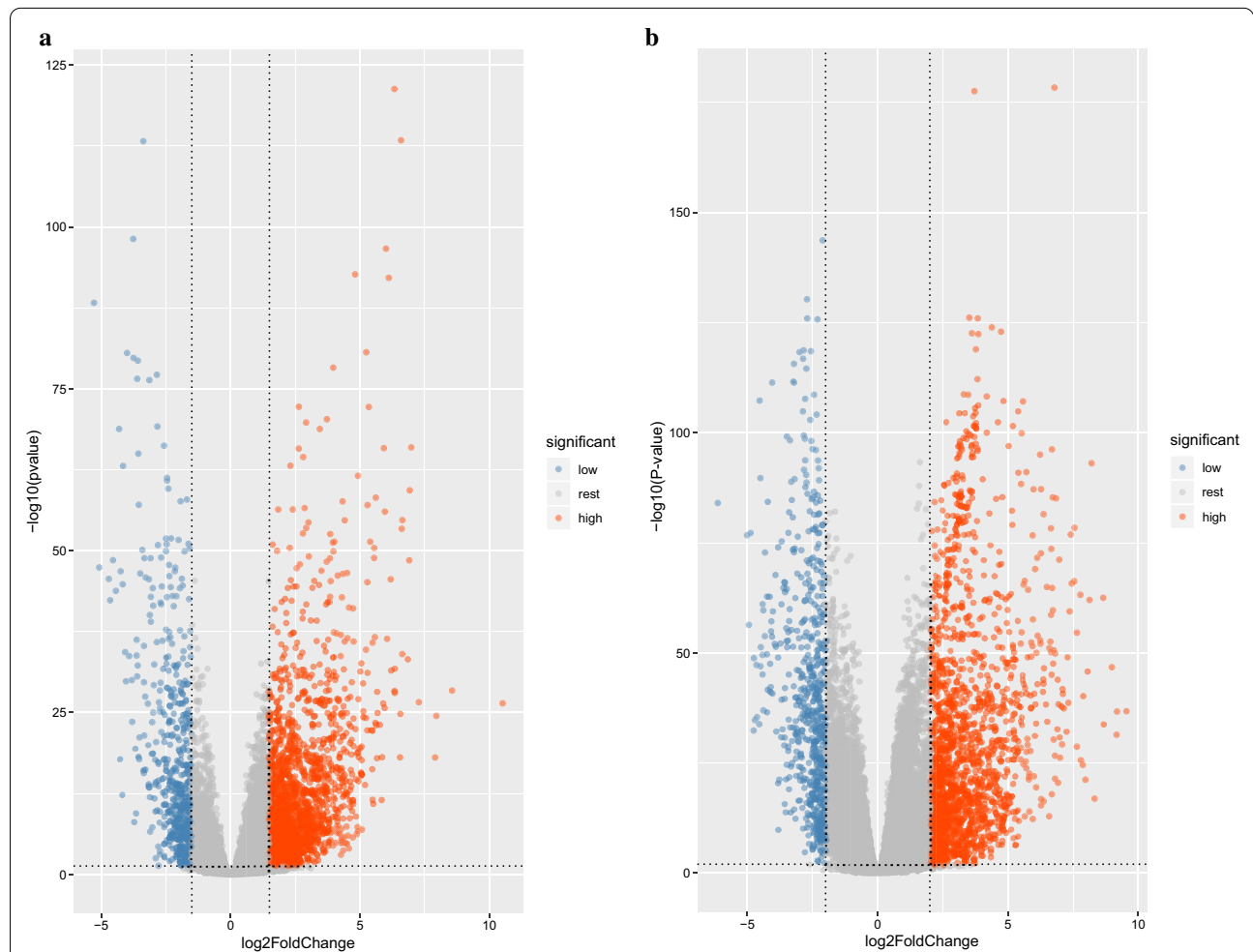
**Survival analysis of the DELncRNAs**

Setting  $p < 0.05$  as a cutoff, 389 DELncRNAs were identified as significantly related to overall survival (Additional file 3: Table S3). Lnc-YARS2-5, Lnc-NPR3-2 and LINC02310 were the top-3 significant DELncRNAs

related to overall survival. The KM survival plots of the three DELncRNAs were shown in Fig. 3. It can be observed that high expression of the three DELncRNAs lead to poor survival probabilities. In Table 2, p1 and p2 were the correlation p-values of Fisher’s accurate test and Pearson’s chi-squared test, respectively. The results further confirmed that these three DELncRNAs were significantly related to overall survival.

**PCR experiments and clinical feature analyses**

Lnc-YARS2-5, Lnc-NPR3-2 and LINC02310 were further investigated by PCR experiments based on the expression level in LADC tissues and its related clinical features. The expression quantification (Fig. 4) were obtained with the  $2^{-\Delta\Delta Ct}$  method and showed that the three DELncRNAs were mainly up-regulated in the LADC tissues, which was consistent with the survival analysis results.



**Fig. 1** Volcano plots of LncRNAs and mRNAs. **a** LncRNAs: the red dots represent the LncRNAs with LFC > 1.5 and  $p < 0.05$ ; the blue dots represent the LncRNAs with LFC < -1.5 and  $p < 0.05$ ; the gray dots are the remaining LncRNAs; **b** mRNAs: the red dots represent the mRNAs with LFC > 2 and  $p < 0.01$ ; the blue dots represent the mRNAs with LFC < -2 and  $p < 0.01$ ; the gray dots are the remaining mRNAs

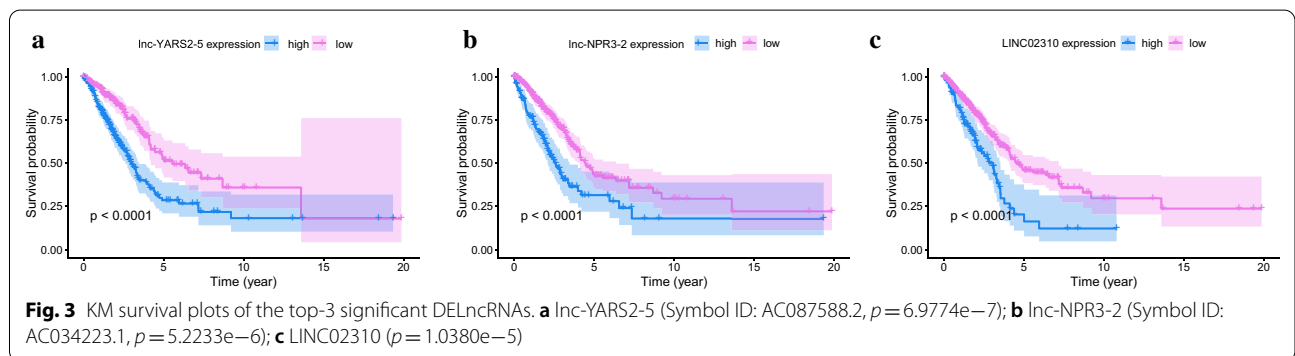
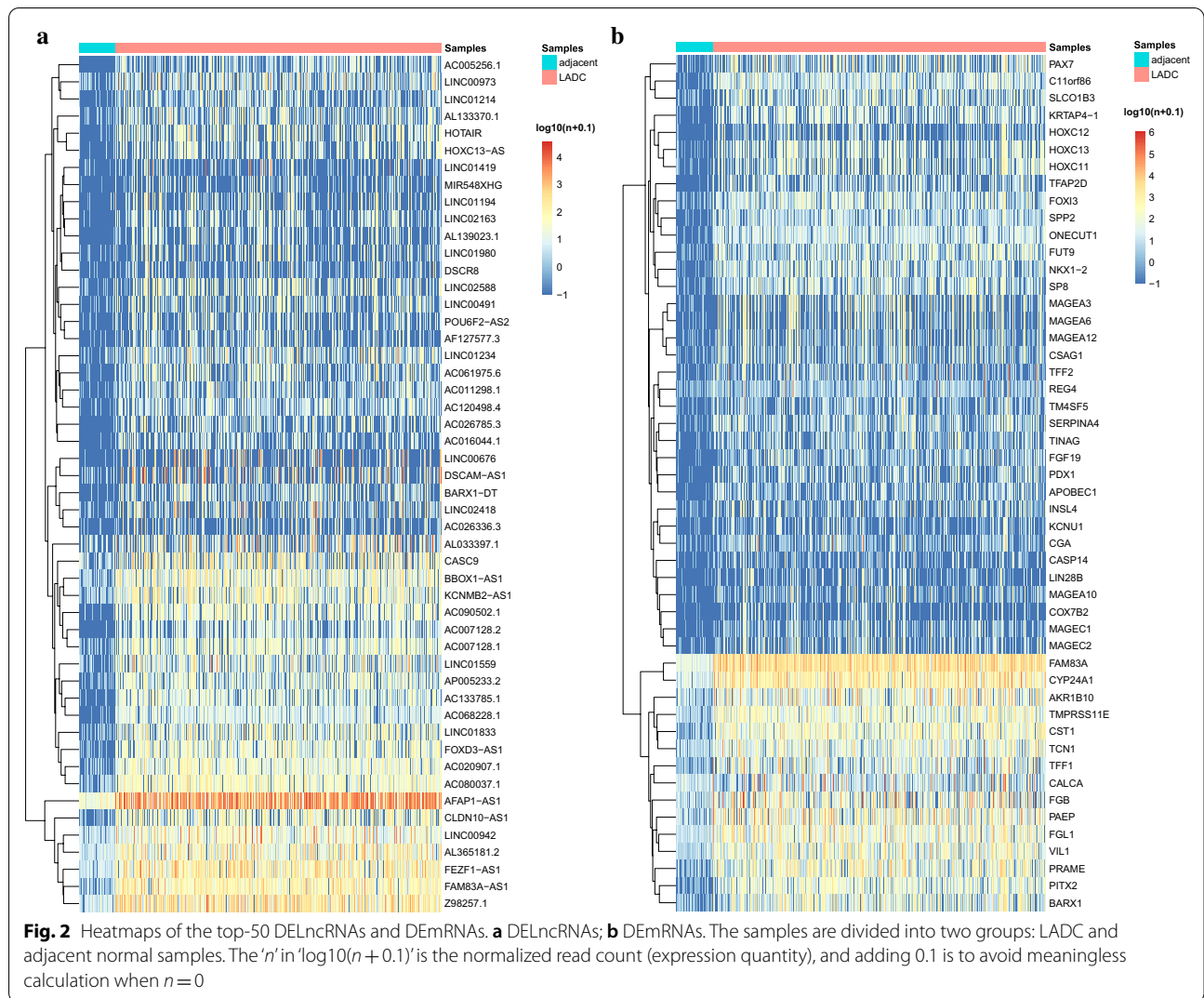


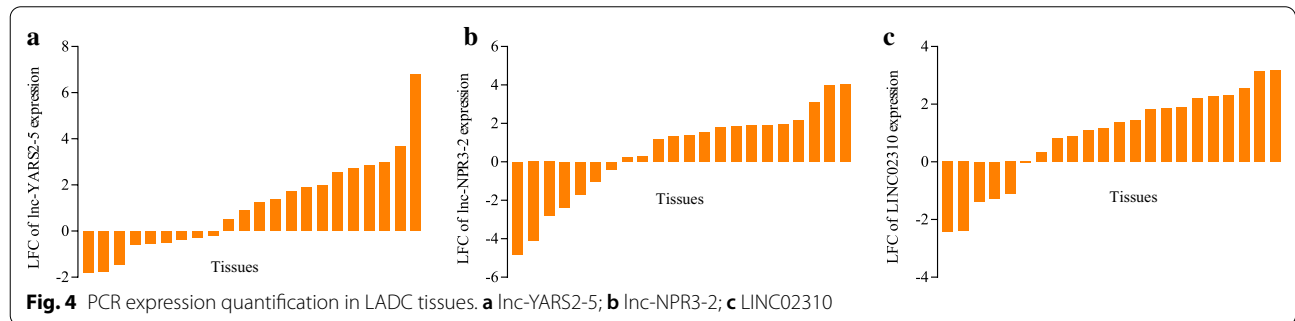
Table 3 showed the clinical feature analyses of the three DELncRNAs by performing Fisher's accurate test ( $p_1$ ) and Pearson's chi-squared test ( $p_2$ ), respectively. These DELncRNAs were correlated with clinical T-stage and

TNM-stage in LADC tissues from Xiangya Hospital. It can be observed that LINC02310 had the most significant effect on tumor proliferation. This result was further validated through the clinical analysis of LINC02310

**Table 2 Validation of the correlation significance between the overall survival and gene expression level**

| Status | Inc-YARS2-5 |      |         |         | Inc-NPR3-2 |      |         |         | LINC02310 |      |         |         |
|--------|-------------|------|---------|---------|------------|------|---------|---------|-----------|------|---------|---------|
|        | Low         | High | $p_1$   | $p_2$   | Low        | High | $p_1$   | $p_2$   | Low       | High | $p_1$   | $p_2$   |
| Dead   | 66          | 117  | 3.28e-6 | 2.31e-6 | 117        | 66   | 7.77e-5 | 4.76e-5 | 126       | 57   | 3.54e-4 | 2.74e-4 |
| Alive  | 186         | 135  |         |         | 258        | 63   |         |         | 266       | 55   |         |         |

'Low' and 'high' represent the LADC samples from TCGA in which the DELncRNAs were low expressed and high expressed, respectively.  $p_1$  and  $p_2$  are the correlation p-values of Fisher's accurate test and Pearson's chi-squared test, respectively



**Table 3 LncRNA expression level and clinical feature analysis of 22 cases of LADC from Xiangya hospital**

| Clinical features | Inc-YARS2-5 |      |       |       | Inc-NPR3-2 |      |       |       | LINC02310 |      |       |       |
|-------------------|-------------|------|-------|-------|------------|------|-------|-------|-----------|------|-------|-------|
|                   | Low         | High | $p_1$ | $p_2$ | Low        | High | $p_1$ | $p_2$ | Low       | High | $p_1$ | $p_2$ |
| Age               |             |      |       |       |            |      |       |       |           |      |       |       |
| < 55              | 1           | 5    | 0.615 | 0.457 | 3          | 5    | >0.99 | 0.636 | 2         | 6    | >0.99 | >0.99 |
| ≥ 55              | 4           | 8    |       |       | 3          | 8    |       |       | 3         | 9    |       |       |
| Gender            |             |      |       |       |            |      |       |       |           |      |       |       |
| Male              | 4           | 8    | 0.615 | 0.457 | 6          | 7    | 0.109 | 0.044 | 5         | 9    | 0.260 | 0.091 |
| Female            | 1           | 5    |       |       | 0          | 6    |       |       | 0         | 6    |       |       |
| T-stage           |             |      |       |       |            |      |       |       |           |      |       |       |
| T1                | 5           | 4    | 0.029 | 0.009 | 1          | 8    | 0.141 | 0.069 | 0         | 10   | 0.033 | 0.010 |
| T2/T3/T4          | 0           | 9    |       |       | 5          | 5    |       |       | 5         | 5    |       |       |
| N-stage           |             |      |       |       |            |      |       |       |           |      |       |       |
| N0                | 5           | 7    | 0.114 | 0.063 | 1          | 10   | 0.041 | 0.013 | 2         | 11   | 0.290 | 0.176 |
| N1/N2/N3          | 0           | 6    |       |       | 5          | 3    |       |       | 3         | 4    |       |       |
| TNM-stage         |             |      |       |       |            |      |       |       |           |      |       |       |
| I                 | 4           | 4    | 0.118 | 0.060 | 1          | 7    | 0.177 | 0.127 | 0         | 9    | 0.038 | 0.020 |
| II/III/IV         | 1           | 9    |       |       | 5          | 6    |       |       | 5         | 6    |       |       |
| Differentiation   |             |      |       |       |            |      |       |       |           |      |       |       |
| Low               | 1           | 5    | 0.542 | 0.395 | 2          | 4    | >0.99 | 0.937 | 2         | 4    | 0.346 | 0.310 |
| Mid-low           | 1           | 2    |       |       | 1          | 2    |       |       | 1         | 2    |       |       |
| Mid               | 1           | 3    |       |       | 2          | 2    |       |       | 0         | 4    |       |       |
| Mid-high          | 1           | 0    |       |       | 1          | 1    |       |       | 2         | 1    |       |       |

'Low' and 'high' denote the cases where the DELncRNAs were low expressed and high expressed, respectively.  $p_1$  and  $p_2$  are the p-values obtained via Fisher's accurate test and Pearson's chi-squared test, respectively

based on TCGA database, as shown in Table 4. Both Fisher's accurate test ( $p_1$ ) and Pearson's chi-squared test ( $p_2$ ) demonstrated that LINC02310 expression was

significantly related to TNM-stage, T-stage and N-stage. Based on the above results, LINC02310 was screened out and reasonably assumed to act as an enhancer in LADC.

**Table 4 Clinical feature analysis of LINC02310 based on TCGA clinical data**

| Clinical features                | Low | High | $p_1$ | $p_2$ |
|----------------------------------|-----|------|-------|-------|
| Gender                           |     |      |       |       |
| Female                           | 216 | 60   | 0.831 | 0.776 |
| Male                             | 183 | 54   |       |       |
| Race                             |     |      |       |       |
| American Indian or Alaska native | 1   | 0    | 0.762 | 0.745 |
| Asian                            | 6   | 1    |       |       |
| Black or African American        | 43  | 9    |       |       |
| White                            | 300 | 87   |       |       |
| Age                              |     |      |       |       |
| < 55                             | 56  | 15   | 0.878 | 0.803 |
| ≥ 55                             | 328 | 95   |       |       |
| TNM-stage                        |     |      |       |       |
| I                                | 222 | 52   | 0.033 | 0.027 |
| II                               | 95  | 26   |       |       |
| III                              | 55  | 29   |       |       |
| IV                               | 21  | 5    |       |       |
| T-stage                          |     |      |       |       |
| T1/T2                            | 352 | 92   | 0.043 | 0.038 |
| T3/T4/TX                         | 47  | 22   |       |       |
| N-stage                          |     |      |       |       |
| N0                               | 268 | 62   | 0.035 | 0.035 |
| N1                               | 71  | 24   |       |       |
| N2/N3/NX                         | 60  | 27   |       |       |
| M-stage                          |     |      |       |       |
| M0                               | 265 | 79   | 0.726 | 0.655 |
| M1                               | 21  | 4    |       |       |
| MX                               | 111 | 29   |       |       |
| Pack-years smoked                |     |      |       |       |
| < 20                             | 49  | 10   | 0.389 | 0.310 |
| ≥ 20                             | 225 | 67   |       |       |

'Low' and 'high' represent the number of LADC samples where LINC02310 were low expressed and high expressed, respectively.  $p_1$  and  $p_2$  are the p-values of Fisher's accurate test and Pearson's chi-squared test, respectively

### CCK-8 assay and colony formation assay

CCK-8 assay and colony formation assay were performed to verify the assumption that LINC02310 promotes cell proliferation in LADC. Firstly, the linear transcript expression vector was constructed. Next, OE, NC, SI and SI-NC were transfected into H1299 and PC-9 cell lines, respectively. Subsequently, the effect of LINC02310 on cell proliferation was examined. CCK-8 assays showed that forced expression of LINC02310 enhanced proliferation of H1299 and PC-9 (Fig. 5). Similarly, colony formation assay indicated that H1299 cell proliferation was significantly increased by forced expression of LINC02310 (Fig. 6). It can be concluded that LINC02310 overexpression significantly promoted cell proliferation

and depressing LINC02310 expression conspicuously inhibited cell proliferation.

### CeRNA regulatory network construction

MiRNA target sites on LINC02310 were illustrated in Fig. 7 according to LncRNASNP2 database. The targeted miRNAs with  $RPM \geq 1$  were selected to interact with 2227 DEmRNAs via starBase. As a result, has-miR-1270, has-miR-506-3p and has-miR-330-3p were involved in 561 interactions with 414 downstream DEmRNAs (Additional file 4: Table S4). The downstream DEmRNAs with  $|LFC| > 3.5$  and  $FDR < 0.01$  were further selected from these interactions to construct the ceRNA regulatory network (Fig. 8) using Cytoscape. The ceRNA regulatory network contained 117 nodes and 149 edges with LINC02310, 3 targeted miRNAs and 113 downstream DEmRNAs involved.

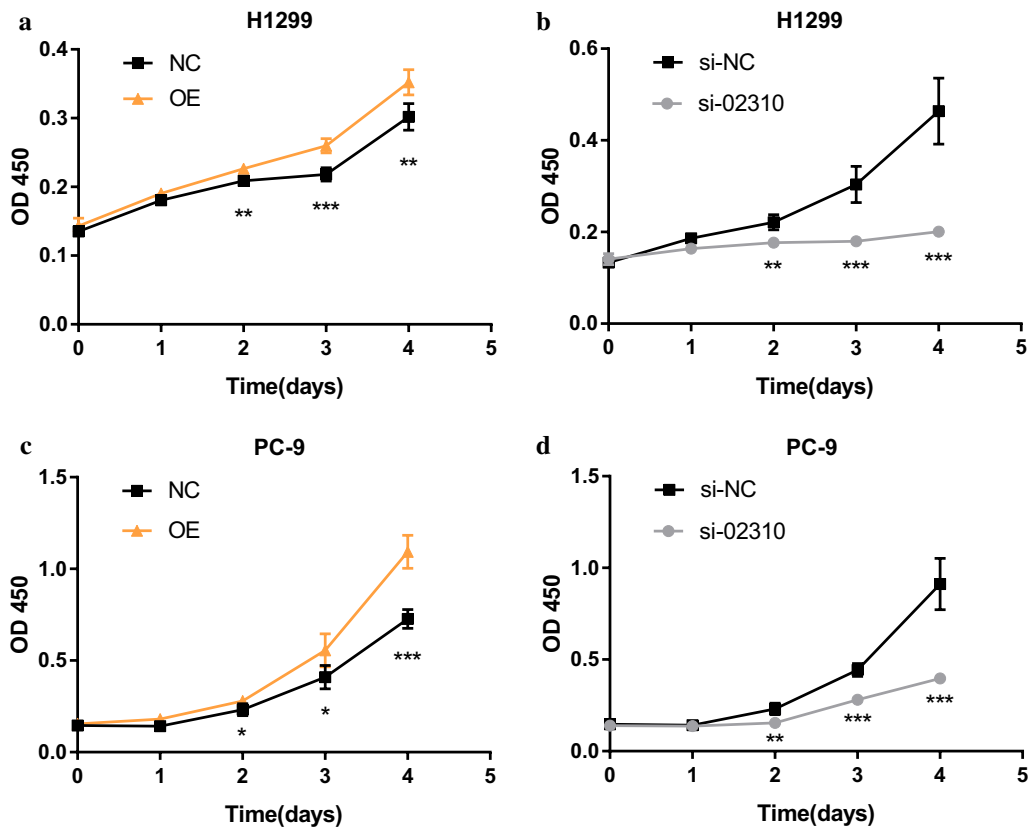
### GO functional annotation and KEGG pathway analysis

GO analysis result (Additional files 5–7: Tables S5–S7) showed that the 414 downstream DEmRNAs were significantly related to 405 GO terms (361 BPs, 17 MFs and 27 CCs) with p-value  $< 0.05$ . The top-10 significantly enriched GO terms were shown in Fig. 9. Most of the downstream DEmRNAs enriched in BPs (Fig. 9a) took part in nuclear division, chromosome segregation, morphogenesis, organelle fission, forebrain development and positive regulation of cell cycle. MF result (Fig. 9b) indicated that the downstream DEmRNAs were significantly associated with DNA binding activities. In addition, the most enriched CC was spindle (Fig. 9c). KEGG pathway analysis (Fig. 9d and Additional file 8: Table S8) showed that 11 pathways were significantly enriched with p-value  $< 0.05$ , which indicated that the downstream DEmRNAs were conspicuously associated with metabolism activities.

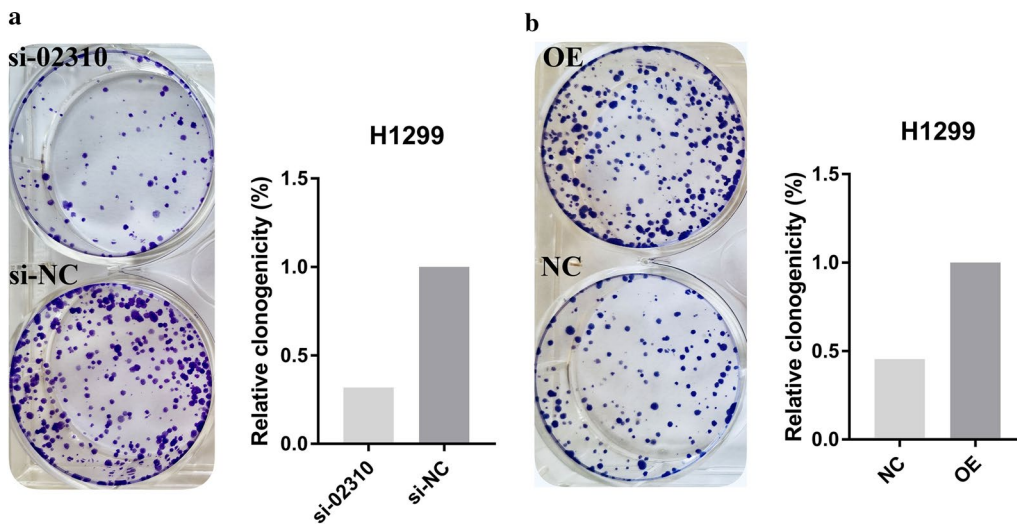
### Protein–protein interaction network analysis

A total of 1081 interactions (Additional file 9: Table S9) involving 226 nodes (downstream DEmRNAs) were identified using the STRING database with combined score  $> 0.7$ . These interactions were analyzed and divided into different modules using MCODE. The module with the highest combined score were screened out (Additional file 10: Table S10, 35 nodes and 534 interactions). For better visualization, these interactions were further screened with combined score  $> 0.95$  to construct the PPI network (Fig. 10, 34 nodes and 156 interactions). The three downstream DEmRNAs with the highest degree are CDK1 (degree = 57), KIF11 (degree = 51) and TOP2A (degree = 50).





**Fig. 5** CCK-8 assays. LINC02310 promoted the cell proliferation of H1299 and PC-9. **a** H1299 OE/NC; **b** H1299 si-02310/si-NC; **c** PC-9 OE/NC; **d** PC-9 si-02310/si-NC. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

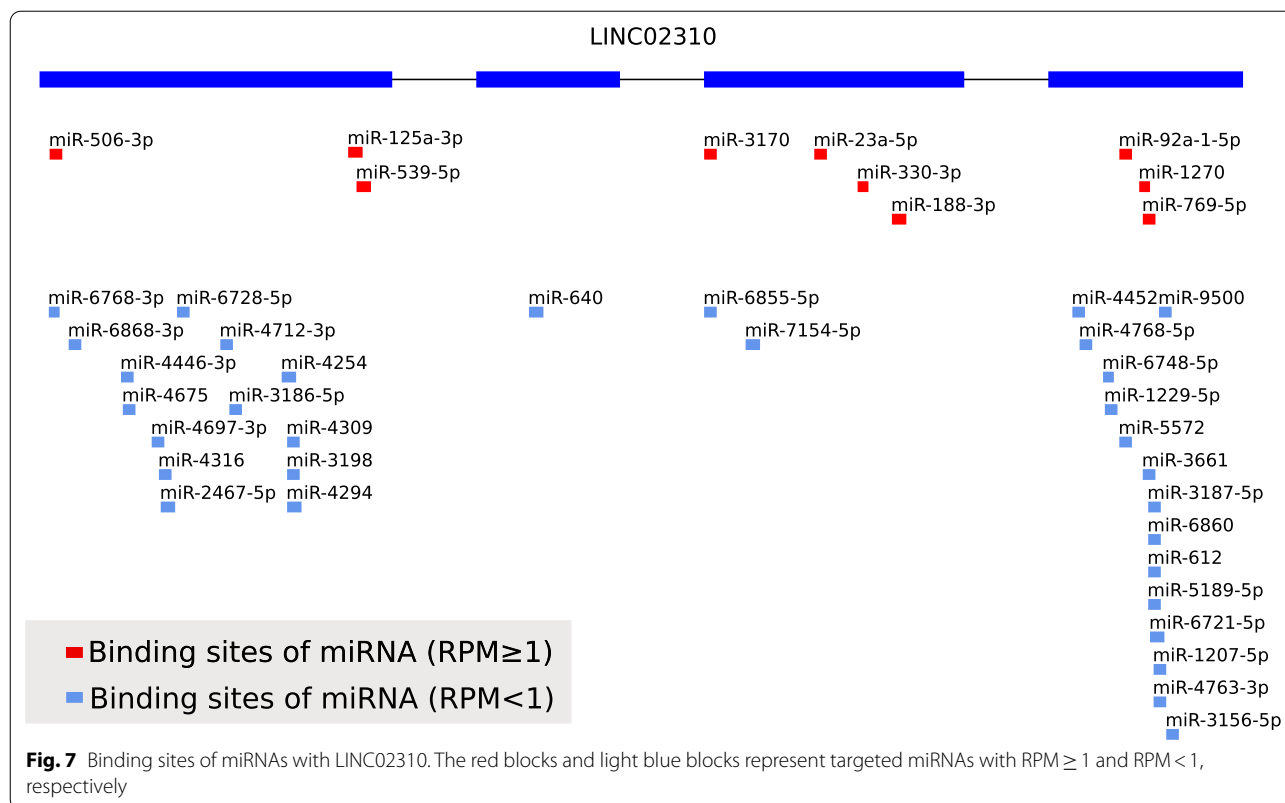


**Fig. 6** Colony formation assays. LINC02310 promoted the clone formation of H1299 cells. **a** H1299 si-02310/si-NC; **b** H1299 OE/NC

**Discussion**

LADC is a major subtype of NSCLC and has a very high mortality rate [38]. The best opportunity for surgery

is often lost because the early symptoms of LADC are insipid and the prognosis is still unsatisfactory. Thus, it is quite necessary to provide prognostic biomarkers for



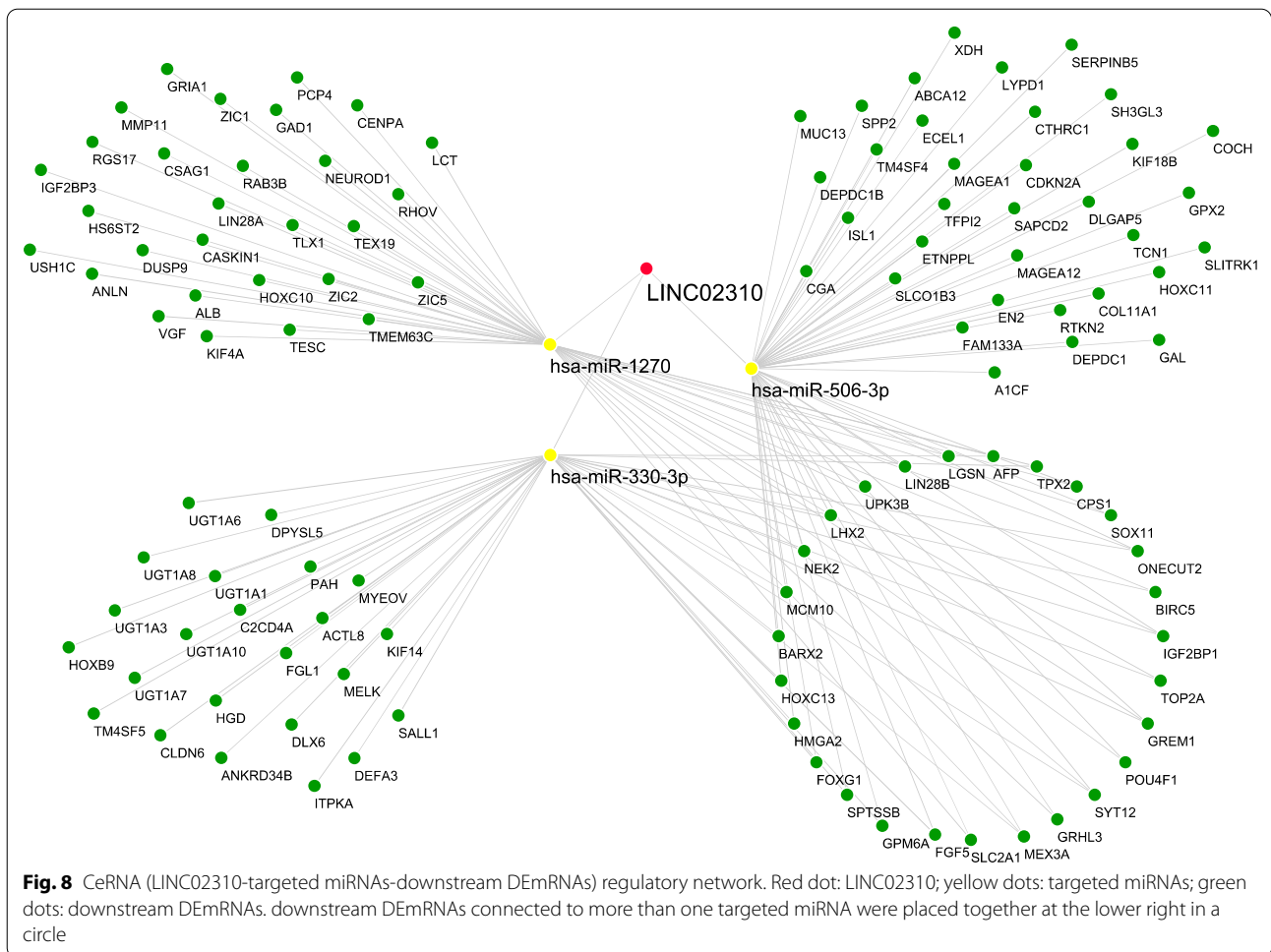
LADC treatments. With the development of molecular biology, lncRNA has been found to be related to the regulation of gene expression. Its ectopic expression can interact with DNA, proteins and RNA, leading to malignant transformation of cells, which can lead to a variety of cancers.

This study identified LINC02310 as an enhancer and promising biomarker in LADC via comprehensive analyses of TCGA data and experimental validation. LINC02310 has almost not been reported except by Yan-Yan Li et al. [39] up until now. However, they simply identified LINC02310 as one of the ten prognostic markers in LADC through a Cox regression model. This study provided comprehensive bioinformatic analyses and further experiments to identify the significance of LINC02310. Further investigations about its targeted miRNAs and downstream DEMRNAs were performed.

To begin with, a total of 2431 DELncRNAs and 2227 DEMRNAs were identified. Next, lnc-YARS2-5, lnc-NPR3-2 and LINC02310 were identified as the top-3 significant DELncRNAs associated with overall survival, and their high expression means low survival probability. The PCR experimental results also showed that these three DELncRNAs were up-regulated in most of the 22 pairs of LADC tissues from Xiangya Hospital, which was in accordance with the survival analysis. The clinical

feature analysis of the 22 pairs of LADC tissues showed that LINC02310 was significantly associated with T-stage and TNM-stage in LADC, which was further verified by analysis of TCGA clinical data. These results indicated that LINC02310 acted as an enhancer in LADC. The results of CCK-8 assay and colony formation assay suggested that up-regulating LINC02310 significantly promoted the growth and proliferation of the LADC cells PC-9 and H1299. On the contrary, after down-regulating the expression of LINC02310, the growth and proliferation of the two cell lines were inhibited. In conclusion, LINC02310 was identified as an enhancer in LADC in this study.

In addition, it is necessary to investigate the downstream RNAs of LINC02310 to reveal the potential functions and mechanisms involved in the regulating pathways. 3 targeted miRNAs (has-miR-1270, has-miR-506-3p and has-miR-330-3p) and 414 downstream DEMRNAs were predicted via the lncRNASNP2 website and starBase database. Has-miR-1270 was reported to interacted with circRNAs to regulate the tumorigenesis and development of LADC [40] and the cell invasion and migration in NSCLC [41]. Has-miR-506-3p was identified as a tumor suppressor in several cancers, including NSCLC [42] and colorectal cancer [43]. It can also reverse gefitinib resistance in NSCLC by targeting Yes-associated

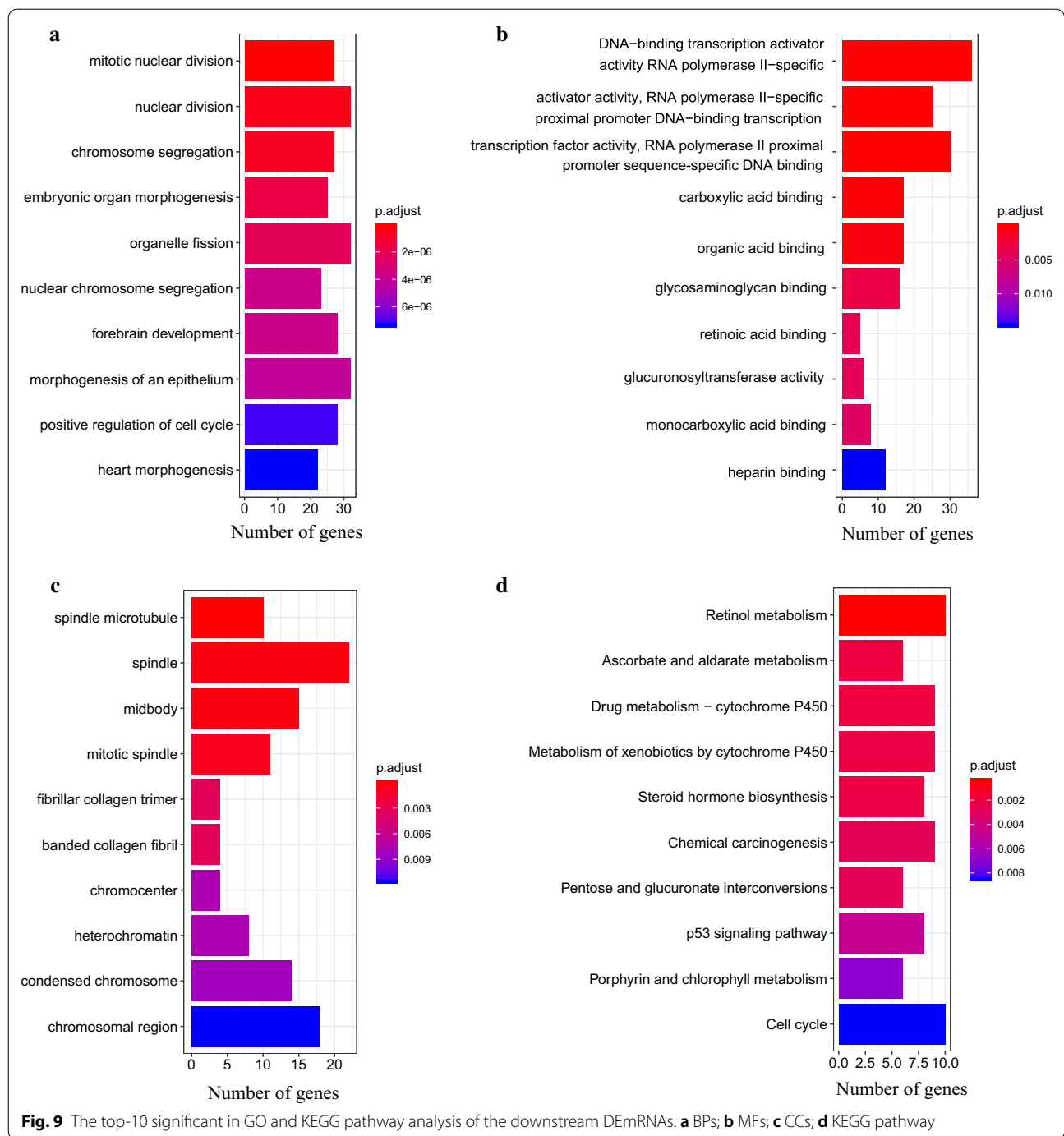


protein 1 [44]. Has-miR-330-3p promoted cell invasion and metastasis [45] and controlled cell proliferation [46] in NSCLC. In conclusion, these three targeted miRNAs played an important role in NSCLC. However, they were rarely reported to interact with LncRNAs in lung cancer. 113 downstream DE mRNAs with  $|LFC| > 3.5$  and  $FDR < 0.01$  were further screened out to construct the ceRNA regulatory network, which revealed the interactions of LINC02310 with its 3 targeted miRNAs and downstream DE mRNAs.

GO and KEGG pathway analysis provided a further understanding of the functions and mechanisms associated with the downstream DE mRNAs. The top-10 GO terms and KEGG pathways were visualized, most of which were reported to associate with cancers. For example, nuclear division, which was the most significant BP term, is associated with promoting malignant progression in colorectal cancer [47]. Transcription factors play important roles in various cancers [48–50]. More potential biomarkers for LADC may be found by investigating these downstream DE mRNAs.

Furthermore, the PPI network of the downstream DE mRNAs was constructed and the three downstream DE mRNAs of the highest degrees were obtained, including CDK1, TOP2A and KIF11. CDK1 was reported to regulate the development of lung cancer by interacting with LncRNA [51], miRNA [52] and circRNA [53]. It also served as a prognostic biomarker for lung cancer [54]. The prognostic significance of TOP2A in NSCLC was revealed by bioinformatic analysis [55]. In addition, the mutations of TOP2A are associated with chemotherapy resistance in lung cancer [56]. KIF11 has been identified as potential biomarkers for NSCLC by several studies [57, 58]. These findings validated the significance of the downstream regulatory network analysis of LINC02310 in this study.

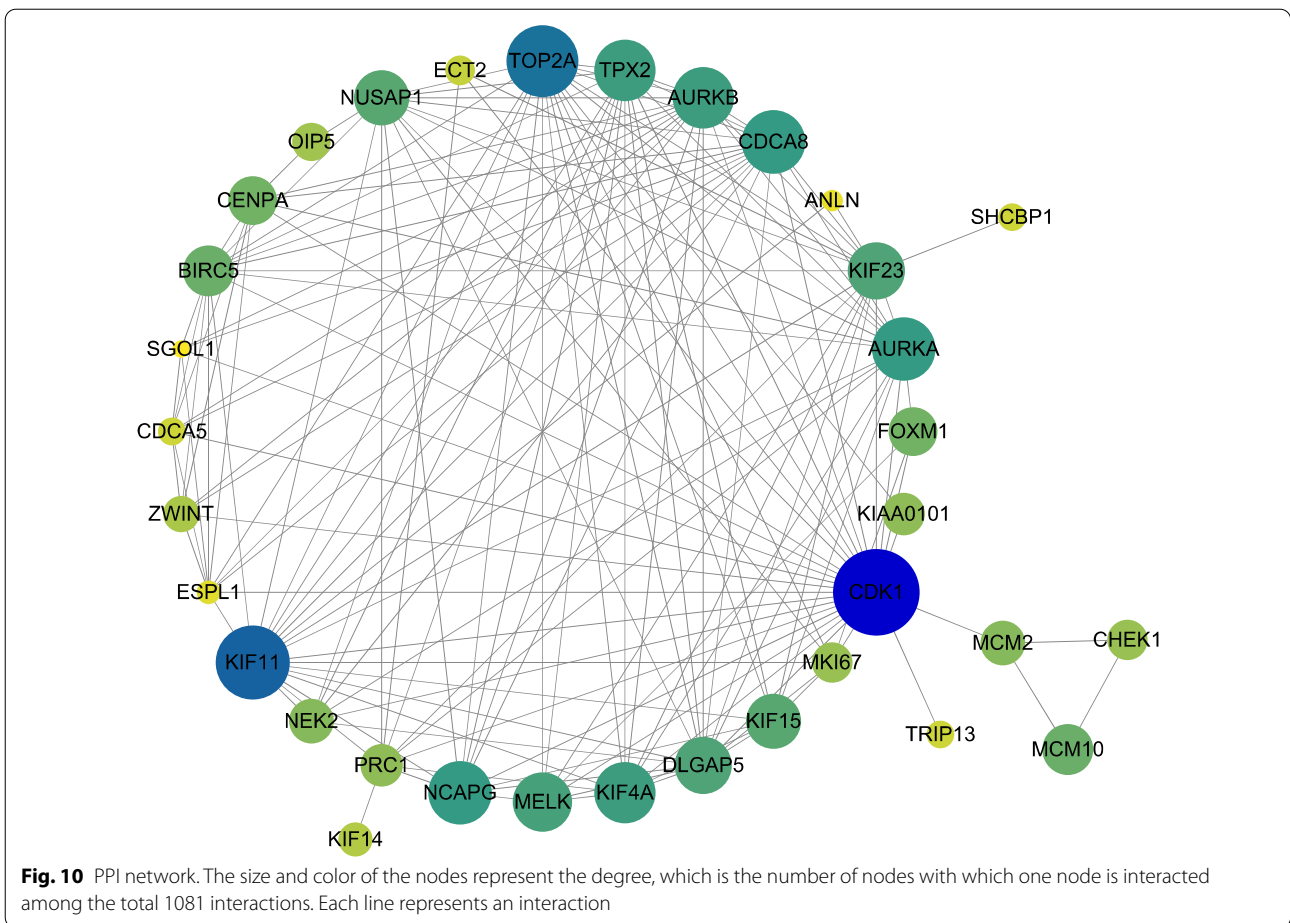
In summary, this study verified that LINC02310 significantly promoted the growth and proliferation of LADC cells, and was closely related to the clinical features of LADC patients. The regulatory network of LINC02310 in the carcinogenesis of LADC was further investigated, which may benefit future studies. However, to reveal the



specific mechanisms of LINC02310 still requires a lot of experiments. The tumorigenesis and progression of LADC is a complex process regulated by multiple factors. In order to realize the early diagnosis of LADC and improve the prognosis of patients, we still need to investigate its potential regulatory mechanisms, which can improve the current diagnostic and treatment methods.

### Conclusions

In this study, LINC02310, which has rarely been investigated previously, was identified as an enhancer in LADC through bioinformatic analyses and experiments, including PCR experiment, CCK-8 assays and colony formation assays. In addition, the ceRNA regulatory network was constructed to identify the targeted miRNAs of LINC02310 and downstream DEmRNAs. GO functional



annotation and KEGG pathway analyses showed the functions and mechanisms of the downstream DEmRNAs. The PPI network revealed the interactions among the downstream DEmRNAs. In conclusion, this research identified LINC02310 as a promising prognostic biomarker for LADC therapies and provided comprehensive analyses on its regulatory network.

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12920-020-00834-6>.

- Additional file 1. Table S1:** List of DELncRNAs.
- Additional file 2. Table S2:** List of DEmRNAs.
- Additional file 3. Table S3:** 389 DELncRNAs significantly related to overall survival.
- Additional file 4. Table S4:** targeted miRNAs interacted with DEmRNAs.
- Additional file 5. Table S5:** GO analysis result (BP terms).
- Additional file 6. Table S6:** GO analysis result (MF terms).
- Additional file 7. Table S7:** GO analysis result (CC terms).
- Additional file 8. Table S8:** KEGG analysis result.

**Additional file 9. Table S9:** downstream DEmRNAs interactions.

**Additional file 10. Table S10:** PPI network clustered by MCODE.

### Abbreviations

LADC: Lung adenocarcinoma; NSCLC: Non-small cell lung cancer; LSCC: Lung squamous cell carcinoma; LCLC: Lung large cell carcinoma; TCGA: The Cancer Genome Atlas; DEmRNA: Differentially expressed mRNA; DELncRNA: Differentially expressed LncRNA; KM: Kaplan–Meier; ceRNA: Competing endogenous RNA; GO: Gene ontology; MF: Molecular function; BP: Biological process; CC: Cellular component; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: Protein–protein interaction.

### Acknowledgments

The authors thank Dr. Suiyu Chen for the helpful discussion. In addition, WZ thanks the inimitable care and support of SC over the past 5 years, and she wants to tell him, 'grow old along with me, the best is yet to be.'

### Authors' contributions

WZ conceived the research and designed the experiments. WZ, JW and QL carried out the experiments. WZ, QL, WP, BL and LW performed data analysis and visualized the results together. WZ, JW, CZ and CD drafted and revised the manuscript. All authors have read and approved the final manuscript.

### Funding

This work was supported by the National Natural Science Foundation of China (Nos. 81572281, 81974367, 81372515), Natural Science Foundation of Hunan Province (2018JJ6131). The funding bodies played no role in the design of the

study and collection, analysis, and interpretation of data and in writing the manuscript.

#### Availability of data and materials

The LADC data sets, the expression quantification data of RNAs, and the relevant clinical datasets are available in TCGA database, (<https://portal.gdc.cancer.gov/>). GENCODE Human Release 31 (GRCh38.p12) was obtained from GENCODE, (<https://www.genencodegenes.org/human/>). The R packages used in this study were downloaded from Bioconductor, (<http://www.bioconductor.org/packages/release/bioc/>). The clinical data of 22 LADC patients and the raw data of PCR experiments are available in figshare website, (<https://doi.org/10.6084/m9.figshare.13266500>). Other data and materials are available as follows: LncRNASNP2 website, (<http://bioinfo.life.hust.edu.cn/lncRNASNP#!/>); StarBase database, (<http://starbase.sysu.edu.cn/>); the software Cytoscape, (<https://cytoscape.org/download.html>); HGNC database, (<https://www.genenames.org/>); STRING database, (<https://string-db.org/>); MCODE, (<http://apps.cytoscape.org/apps/mcode>).

#### Ethics approval and consent to participate

This research was approved by the ethics committee of Xiangya Hospital of Central South University (No. 201503303). All procedures conducted in the study were in accordance with the Ethics Standards Institutions Research Committee and the 1964 Helsinki Declaration and its subsequent amendments or similar ethical standards. All participants signed the informed consent voluntarily.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup> Department of Thoracic Surgery, Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China. <sup>2</sup> Department of Oncology, Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China. <sup>3</sup> Institute of Medical Sciences, Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China. <sup>4</sup> National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China.

Received: 20 February 2020 Accepted: 24 November 2020

Published online: 11 December 2020

#### References

- Liu G, Pei F, Yang F, Li L, Amin AD, Liu S, Buchan JR, Cho WC. Role of autophagy and apoptosis in non-small-cell lung cancer. *Int J Mol Sci*. 2017;18(2):367.
- Wang Q, Liu S, Zhao X, Wang Y, Tian D, Jiang W. MiR-372-3p promotes cell growth and metastasis by targeting FGF9 in lung squamous cell carcinoma. *Cancer Med*. 2017;6(6):1323–30.
- Chujo T, Hirose T. Nuclear bodies built on architectural long noncoding RNAs: unifying principles of their construction and function. *Mol Cells*. 2017;40(12):889–96.
- Noh JH, Kim KM, McClusky WG, Abdelmohsen K, Gorospe M. Cytoplasmic functions of long noncoding RNAs. *Wiley Interdiscip Rev RNA*. 2018;9(3):e1471.
- Zhou W, Chen X, Hu Q, Chen X, Chen Y, Huang L. Galectin-3 activates TLR4/NF-kappaB signaling to promote lung adenocarcinoma cell proliferation through activating lncRNA-NEAT1 expression. *BMC Cancer*. 2018;18(1):580.
- Dong HX, Wang R, Jin XY, Zeng J, Pan J. LncRNA DGCR5 promotes lung adenocarcinoma (LUAD) progression via inhibiting hsa-mir-22-3p. *J Cell Physiol*. 2018;233(5):4126–36.
- Liu L, Liu L, Lu S. LncRNA H19 promotes viability and epithelial-mesenchymal transition of lung adenocarcinoma cells by targeting miR-29b-3p and modifying STAT3. *Int J Oncol*. 2019;54(3):929–41.
- Zhao W, Shan B, He D, Cheng Y, Li B, Zhang C, Duan C. Recent progress in characterizing long noncoding RNAs in cancer drug resistance. *J Cancer*. 2019;10(26):6693–702.
- Bhan A, Soleimani M, Mandal SS. Long noncoding RNA and cancer: a new paradigm. *Cancer Res*. 2017;77(15):3965–81.
- Renganathan A, Felley-Bosco E. Long noncoding RNAs in cancer and therapeutic potential. *Adv Exp Med Biol*. 2017;1008:199–222.
- Chandra GS, Nandan TY. Potential of long non-coding RNAs in cancer patients: From biomarkers to therapeutic targets. *Int J Cancer*. 2017;140(9):1955–67.
- Chen X, Chen Z, Yu S, Nie F, Yan S, Ma P, Chen Q, Wei C, Fu H, Xu T, et al. Long noncoding RNA LINC01234 functions as a competing endogenous RNA to regulate CBFβ expression by sponging miR-204-5p in gastric cancer. *Clin Cancer Res*. 2018;24(8):2002–14.
- Zhao X, Liu Y, Li Z, Zheng S, Wang Z, Li W, Bi Z, Li L, Jiang Y, Luo Y, et al. Linc00511 acts as a competing endogenous RNA to regulate VEGFA expression through sponging hsa-miR-29b-3p in pancreatic ductal adenocarcinoma. *J Cell Mol Med*. 2018;22(1):655–67.
- Qiu YY, Wu Y, Lin MJ, Bian T, Xiao YL, Qin C. LncRNA-MEG3 functions as a competing endogenous RNA to regulate Treg/Th17 balance in patients with asthma by targeting microRNA-17/RORγmat. *Biomed Pharmacother*. 2019;111:386–94.
- Afify AY, Ibrahim SA, Aldamsisi MH, Zaghloul MS, El-Ekiaby N, Abdelaziz AI. Competing endogenous RNAs in hepatocellular carcinoma—the pinnacle of rivalry. *Semin Liver Dis*. 2019;39:463–75.
- Zhou RS, Zhang EX, Sun QF, Ye ZJ, Liu JW, Zhou DH, Tang Y. Integrated analysis of lncRNA-miRNA-mRNA ceRNA network in squamous cell carcinoma of tongue. *BMC Cancer*. 2019;19(1):779.
- Jiang WD, Yuan PC. Molecular network-based identification of competing endogenous RNAs in bladder cancer. *PLoS ONE*. 2019;14(8):e220118.
- Guan YJ, Ma JY, Song W. Identification of circRNA-miRNA-mRNA regulatory network in gastric cancer by analysis of microarray data. *Cancer Cell Int*. 2019;19:183.
- Li Z, Jiang C, Yuan Y. TCGA based integrated genomic analyses of ceRNA network and novel subtypes revealing potential biomarkers for the prognosis and target therapy of tongue squamous cell carcinoma. *PLoS ONE*. 2019;14(5):e216834.
- Liu Y, Xie D, He Z, Zheng L. Integrated analysis reveals five potential ceRNA biomarkers in human lung adenocarcinoma. *PEERJ*. 2019;7:e6694.
- Li L, Peng M, Xue W, Fan Z, Wang T, Lian J, Zhai Y, Lian W, Qin D, Zhao J. Integrated analysis of dysregulated long non-coding RNAs/microRNAs/mRNAs in metastasis of lung adenocarcinoma. *J Transl Med*. 2018;16(1):372.
- Li X, Li B, Ran P, Wang L. Identification of ceRNA network based on a RNA-seq shows prognostic lncRNA biomarkers in human lung adenocarcinoma. *Oncol Lett*. 2018;16(5):5697–708.
- Wang X, Ding Y, Da B, Fei Y, Feng G. Identification of potential prognostic long noncoding RNA signatures based on a competing endogenous RNA network in lung adenocarcinoma. *Oncol Rep*. 2018;40(6):3199–212.
- Zhao J, Cheng W, He X, Liu Y, Li J, Sun J, Li J, Wang F, Gao Y. Construction of a specific SVM classifier and identification of molecular markers for lung adenocarcinoma based on lncRNA-miRNA-mRNA network. *Onco Targets Ther*. 2018;11:3129–40.
- Sui J, Li YH, Zhang YQ, Li CY, Shen X, Yao WZ, Peng H, Hong WW, Yin LH, Pu YP, et al. Integrated analysis of long non-coding RNA-associated ceRNA network reveals potential lncRNA biomarkers in human lung adenocarcinoma. *Int J Oncol*. 2016;49(5):2023–36.
- Li DS, Ainiwaer JL, Sheyhiding I, Zhang Z, Zhang LW. Identification of key long non-coding RNAs as competing endogenous RNAs for miRNA-mRNA in lung adenocarcinoma. *Eur Rev Med Pharmacol Sci*. 2016;20(11):2285–95.
- Qiu M, Feng D, Zhang H, Xia W, Xu Y, Wang J, Dong G, Zhang Y, Yin R, Xu L. Comprehensive analysis of lncRNA expression profiles and identification of functional lncRNAs in lung adenocarcinoma. *Oncotarget*. 2016;7(13):16012–22.
- Wang Y, Lu T, Wo Y, Sun X, Li S, Miao S, Dong Y, Leng X, Jiao W. Identification of a putative competitive endogenous RNA network for lung adenocarcinoma using TCGA datasets. *PEERJ*. 2019;7:e6809.
- Zhu TG, Xiao X, Wei Q, Yue M, Zhang LX. Revealing potential long non-coding RNA biomarkers in lung adenocarcinoma using long non-coding

- RNA-mediated competitive endogenous RNA network. *Braz J Med Biol Res.* 2017;50(9):e6297.
30. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
  31. Chansky K, Detterbeck FC, Nicholson AG, Rusch VW, Vallieres E, Groome P, Kennedy C, Krasnik M, Peake M, Shemanski L, et al. The IASLC Lung Cancer Staging Project: external validation of the revision of the TNM stage groupings in the eighth edition of the TNM classification of lung cancer. *J Thorac Oncol.* 2017;12(7):1109–21.
  32. Aokage K, Miyoshi T, Ishii G, Kusumoto M, Nomura S, Katsumata S, Sekihara K, Hishida T, Tsuboi M. Clinical and pathological staging validation in the eighth edition of the TNM classification for lung cancer: correlation between solid size on thin-section computed tomography and invasive size in pathological findings in the new T classification. *J Thorac Oncol.* 2017;12(9):1403–12.
  33. Gong J, Liu W, Zhang J, Miao X, Guo AY. IncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse. *Nucl Acids Res.* 2015;43(Database issue):D181–6.
  34. Miao YR, Liu W, Zhang Q, Guo AY. IncRNASNP2: an updated database of functional SNPs and mutations in human and mouse lncRNAs. *Nucl Acids Res.* 2018;46(D1):D276–80.
  35. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucl Acids Res.* 2014;42(Database issue):D92–7.
  36. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–7.
  37. Li X, Xiao X, Chang R, Zhang C. Comprehensive bioinformatics analysis identifies lncRNA HCG22 as a migration inhibitor in esophageal squamous cell carcinoma. *J Cell Biochem.* 2019;121:468–81.
  38. Niu R, Shao X, Shao X, Wang J, Jiang Z, Wang Y. Lung adenocarcinoma manifesting as ground-glass opacity nodules 3 cm or smaller: evaluation with combined high-resolution CT and PET/CT modality. *AJR Am J Roentgenol.* 2019;213:1–10.
  39. Li YY, Yang C, Zhou P, Zhang S, Yao Y, Li D. Genome-scale analysis to identify prognostic markers and predict the survival of lung adenocarcinoma. *J Cell Biochem.* 2018;119(11):8909–21.
  40. Gao N, Ye B. Circ-SOX4 drives the tumorigenesis and development of lung adenocarcinoma via sponging miR-1270 and modulating PLAGL2 to activate WNT signaling pathway. *Cancer Cell Int.* 2020;20:2.
  41. Xu Z, Xiang W, Chen W, Sun Y, Qin F, Wei J, Yuan L, Zheng L, Li S. Circ-IGF1R inhibits cell invasion and migration in non-small cell lung cancer. *Thorac Cancer.* 2020;11(4):875–87.
  42. Guo S, Yang P, Jiang X, Li X, Wang Y, Zhang X, Sun B, Zhang Y, Jia Y. Genetic and epigenetic silencing of miR-506-3p enhances COTL1 oncogene expression to foster non-small lung cancer progression. *Oncotarget.* 2017;8(1):644–57.
  43. Wei C, Yang C, Wang S, Shi D, Zhang C, Lin X, Liu Q, Dou R, Xiong B. Cross-talk between cancer cells and tumor associated macrophages is required for mesenchymal circulating tumor cell-mediated colorectal cancer metastasis. *Mol Cancer.* 2019;18(1):64.
  44. Zhu J, Tao L, Jin L. MicroRNA5063p reverses gefitinib resistance in non-small cell lung cancer by targeting Yes-associated protein 1. *Mol Med Rep.* 2019;19(2):1331–9.
  45. Wei CH, Wu G, Cai Q, Gao XC, Tong F, Zhou R, Zhang RG, Dong JH, Hu Y, Dong XR. MicroRNA-330-3p promotes cell invasion and metastasis in non-small cell lung cancer through GRIA3 by activating MAPK/ERK signaling pathway. *J Hematol Oncol.* 2017;10(1):125.
  46. Liu X, Shi H, Liu B, Li J, Liu Y, Yu B. miR-330-3p controls cell proliferation by targeting early growth response 2 in non-small-cell lung cancer. *Acta Biochim Biophys Sin (Shanghai).* 2015;47(6):431–40.
  47. Yan X, Huang L, Liu L, Qin H, Song Z. Nuclear division cycle 80 promotes malignant progression and predicts clinical outcome in colorectal cancer. *Cancer Med.* 2018;7(2):420–32.
  48. Dittmer J. The role of the transcription factor Ets1 in carcinoma. *SEMIN Cancer Biol.* 2015;35:20–38.
  49. Sun P, Song Y, Liu D, Liu G, Mao X, Dong B, Braicu EI, Sehoul J. Potential role of the HOXD8 transcription factor in cisplatin resistance and tumour metastasis in advanced epithelial ovarian cancer. *Sci Rep.* 2018;8(1):13483.
  50. Chen Y, Zhang C, Chen J, Zhang B, Zhang H, Yang X, Liu J, Wu Q. Expression of transcription factor 21 (TCF21) and upregulation its level inhibits invasion and metastasis in esophageal squamous cell carcinoma. *Med Sci Monit.* 2018;24:4128–36.
  51. Tong W, Han TC, Wang W, Zhao J. lncRNA CASC11 promotes the development of lung cancer through targeting microRNA-302/CDK1 axis. *Eur Rev Med Pharmacol Sci.* 2019;23(15):6539–47.
  52. Shi Q, Zhou Z, Ye N, Chen Q, Zheng X, Fang M. MiR-181a inhibits non-small cell lung cancer cell proliferation by targeting CDK1. *Cancer Biomark.* 2017;20(4):539–46.
  53. Wang B, Hua P, Zhao B, Li J, Zhang Y. Circular RNA circDLGAP4 is involved in lung cancer development through modulating microRNA-143/CDK1 axis. *Cell Cycle.* 2020;19(16):2007–17.
  54. Li M, He F, Zhang Z, Xiang Z, Hu D. CDK1 serves as a potential prognostic biomarker and target for lung cancer. *J Int Med Res.* 2020;48(2):1219697060.
  55. Ma W, Wang B, Zhang Y, Wang Z, Niu D, Chen S, Zhang Z, Shen N, Han W, Zhang X, et al. Prognostic significance of TOP2A in non-small cell lung cancer revealed by bioinformatic analysis. *Cancer Cell Int.* 2019;19:239.
  56. Wang TL, Ren YW, Wang HT, Yu H, Zhao YX. Association of topoisomerase II (TOP2A) and dual-specificity phosphatase 6 (DUSP6) single nucleotide polymorphisms with radiation treatment response and prognosis of lung cancer in Han Chinese. *Med Sci Monit.* 2017;23:984–93.
  57. Schneider MA, Christopoulos P, Muley T, Warth A, Klingmueller U, Thomas M, Herth FJ, Dienemann H, Mueller NS, Theis F, et al. 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.* 2017;50(2):365–72.
  58. Liu X, Liu X, Li J, Ren F. Identification and integrated analysis of key biomarkers for diagnosis and prognosis of non-small cell lung cancer. *Med Sci Monit.* 2019;25:9280–9.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

