BMC Cancer

RESEARCH

Open Access

Establishment of potential lncRNA-related hub genes involved competitive endogenous RNA in lung adenocarcinoma

Yong Li[1], Danfei Shi[2], Yan Jiang[3], Yanqin Hu[1], Qiuxia Liu[1], Yanping Xie[3] and Xilin Zhang[4*]

Abstract

Long non-coding RNAs (lncRNAs) have a notable role in the diagnosis and prognosis of cancer. However, the associations between lncRNA-related hub genes (LRHGs) expression and the corresponding outcomes have not been fully understood in lung adenocarcinoma (LUAD). Here, a total of 71 patients diagnosed with LUAD and 60 healthy volunteers at The First Affiliated Hospital of Huzhou University from April, 2023 to December, 2023 were enrolled in the present study. A LRHGs model was established using least absolute shrinkage and selection operator analyses of The Cancer Genome Atlas-LUAD datasets. The underlying mechanisms of the LRHGs were investigated via Gene Set Enrichment Analysis and Gene Set Variation Analysis. Additionally, the diagnostic role of serum HOXD cluster antisense RNA 2 (HOXD-AS2) was assessed by receiver operating characteristic (ROC) curve analysis. Lastly, TCGA-LUAD samples were divided into high- and low-HOXD-AS2 expression groups based on the median expression. The associations between HOXD-AS2 expression and miR-4538 as well as Calmodulin-Dependent Protein Kinase Type II subunit Beta (CAMK2B) levels were conducted through Pearson correlation analysis. A comprehensive analysis identified 141 differentially expressed lncRNAs between 539 LUAD tissues and 59 normal samples. A prognostic marker for overall survival was established by constructing a predictive signature consisting of 9 LRHGs. Subsequently, 474 LUAD samples were categorized into a high or low-risk group based on the median of the risk score. An independent prognostic model was constructed to confirm the validity of this categorization. Further comparisons of the clinicopathological features and LRHG-related pathways were performed between the two groups. Examinations of LRHG expression in two LUAD clusters and of the association between LRHG expression and immune infiltration were also conducted. HOXD-AS2 expression was shown to be elevated in LUAD tissues compared with matched normal tissues, and the serum HOXD-AS2 level was also notably increased in LUAD samples compared with healthy controls. The results of the ROC analysis indicated that the sensitivity and specificity of HOXD-AS2 were higher than that of cytokeratin-19 fragment (CYFRA21-1), which is a serum marker for LUAD. Pearson analyses indicated that miR-4538 level was negatively associated with HOXD-AS2 expression, but CAMK2B level showed positive correlation in LUAD. The results of the present study therefore indicated that the constructed LRHG model, particularly HOXD-AS2, could independently diagnose and predict the prognosis of LUAD, which suggested the underlying mechanism of the HOXD-AS2/miR-4538/CAMK2B, and might offer efficient strategies for LUAD treatment.

Keywords Lung adenocarcinoma, LncRNA, Diagnosis, ceRNA

*Correspondence:
Xilin Zhang
zhangxilin16@126.com
Full list of author information is available at the end of the article

Introduction

Lung adenocarcinoma (LUAD) is the most common pathological type of non-small cell lung cancer (NSCLC), accounting for ~ 50% of primary NSCLC cases [1, 2]. With the widespread use of low-dose computed tomography for early screening and the rapid development of target drugs for the drivers of genetic changes, the treatment of LUAD has obtained great progress [1]. However, the 5-year survival rate of patients with LUAD has still not significantly improved. The LUAD serum tumor markers, carcinoembryonic antigen, cytokeratin-19 fragment (CYFRA21-1) and neuron-specific enolase (NSE), are used for clinical diagnosis, but the specificity and sensitivity of these protein markers are not sufficient [3]. Therefore, LUAD remains a serious public health concern and the discovery of novel effective diagnostic targets for these patients is vital.

Long non-coding RNAs (lncRNAs) are RNAs that are > 200 nucleotides in length which do not have protein-coding capacity. Accumulating evidence has demonstrated that lncRNAs have a pivotal role in the modulation of multiple pathological and physiological processes via interacting with sponging target micro (mi)RNAs, transcription factors and proteins [4–6]. It has been reported that lncRNAs could be regarded as diagnostic and prognostic markers in LUAD [7–9]. For instance, lncRNA metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) is one of the most notable lncRNA and is significantly promoted in LUAD, where it promotes cell proliferation, migration and metastasis, leading to LUAD progression [10]. The sensitivity and specificity of MALAT1 in NSCLC is 0.87 and 0.95, respectively [11]. In addition, Fan et al. [12] reported that the expression of lncRNA LITATS1 was reduced in LUAD tissues compared with matched adjacent normal tissues, and that patients with low LITATS1 expression had a poor prognosis. It was also observed that lncRNA HOXD-AS2 was highly expressed in cisplatin-resistance A549 cell lines than in cisplatin-sensitive cell lines, providing a new therapeutic target for patients with LUAD [13]. Thus, the identification of novel diagnostic and prognostic lncRNAs has contributed to the early detection, prognosis monitoring and treatment of patients with LUAD. However, a large number of lncRNAs have still not been studied in LUAD.

The present study was aimed to construct a lncRNA-related hub genes (LRHGs) model to diagnose and predict prognosis and immune infiltration of patients with LUAD using multiple databases, advanced bioinformatics tools and statistical methods. Firstly, a LRHGs model was constructed to investigate the role of LRHGs in the diagnosis and prognosis of LUAD, using the datasets from The Cancer Genome Atlas (TCGA) database. Subsequently, the LRHGs associated networks were enriched based on Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA) to certify their biological functions and downstream pathways. The associations between LRHGs expression and the tumor immune microenvironment were also investigated. Further experimental validations were conducted to confirm the expression levels of the LRHGs and the diagnostic value of HOXD cluster antisense RNA 2 (HOXD-AS2) in LUAD. Lastly, it was observed that the HOXD-AS2/miR-4538/CAMK2B axis plays an extremely important role in LUAD. Overall, the present study is one of the few studies to focus on the biological characteristics and mechanisms of lncRNAs in LUAD, which provides a new perspective for understanding the progression of LUAD.

Materials and methods

Patients

According to the inclusion and exclusion criteria, a total of 71 patients diagnosed with LUAD and 60 healthy controls at The First Affiliated Hospital of Huzhou University (Huzhou, China) from April, 2023 to December, 2023 were enrolled in the present study. The criteria for patient recruitment included: (a) diagnosed LUAD; (b) no neoadjuvant therapy; (c) complete basic information. We excluded patients who (a) previously combined with other malignant tumors or other severe diseases; (b) received radiotherapy, chemotherapy, targeted therapy, or/and other treatments before operation; (c) were unwilling to participate [14]. The LUAD group contained 37 men and 34 women (age range, 32–87 years; average age, 63 years). The healthy control group consisted of 28 men and 32 women (age range, 33–88 years; average age, 64 years). Serum samples were obtained from patients on preoperative day 1 and from healthy donors, and then stored at -80˚C until further use. In addition, total of 21 paired LUAD tissues and adjacent normal tissues were collected. Data on the clinicopathological features of patients were also collected. The present study was approved by The Medical Research and Clinical Trials Ethics Committee of The First Affiliated Hospital of Huzhou University (approval no. 2023KYLL014), in accordance with the Declaration of Helsinki. Written informed consent was obtained from all patients.

Data acquisition

The RNA-sequencing (RNA-seq) data of 598 specimens consisting of 539 LUAD and 59 normal lung tissues were downloaded from TCGA database (https://portal.gdc.cancer.gov/) [15]. After eliminating without complete clinicopathological information and survival data, 474 patients were included in the further analysis. The RNA-seq data were subsequently normalized into

the Fragments Per Kilobase per Million format using the "easyTCGA" package. Then, the batch effects were revised using the ComBat algorithm. The expression values for each gene were standardized as z-scores. The clinicopathological characteristics of the patients with LUAD were obtained from the UCSC Xena database (http://genome.ucsc.edu) [16].

Enrichment analyses of differentially expressed genes (DEGs)

To identify the potential mechanisms of the DEGs in LUAD, the RNA-seq data were firstly standardized using the "limma" package (https://bioinf.wehi.edu.au/limma/). To construct a reliable prognostic and diagnostic model, $|\log_2$ fold change (FC)$|>0$ and $P<0.05$ served as the screening criteria, which was consistent with previous reports [17, 18]. Then, a volcano plot of the DEGs was constructed using $|\log_2 FC|>0$ and false discovery rate (FDR) <0.05 as the criteria for significant differential expression. The genes with $\log_2 FC>0$ and FDR <0.05 were considered upregulated genes, and the genes with $\log_2 FC<0$ and FDR <0.05 were considered downregulated genes.

Construction and evaluation of a LRHGs model

According to the screening criteria ($|\log_2 FC|>0$ and $P<0.05$), 141 differentially expressed lncRNAs (DELRs) were obtained and a volcano plot was constructed using the "ggplot2" package. To investigate the relationship between the DELRs and overall survival (OS), univariate Cox regression analysis was conducted. It was observed that 10 lncRNAs were notably associated with the OS of patients with LUAD. Least absolute shrinkage and selection operator (LASSO) was subsequently used to obtain the LRHGs based on DELR expression through the "glmnet" package, and deviance and tenfold cross-validation were conducted [16, 19]. LASSO uses L1 regularization which not only contributes to overfitting reduction but also to selecting characteristics by shrinking certain coefficients to 0. Lastly, an efficient 9 LRHGs model was constructed on the basis of the following computational formula utilizing the LASSO regression algorithm: $\text{Risk } scores = \sum_i \text{coefficient}\left(\text{hub gene}_i\right) \text{ x lncRNA expression}\left(\text{hub gene}_i\right)$. 'Coefficient (hub gene$_i$)' and 'lncRNA expression (hub gene$_i$)' represent the LASSO coefficient and corresponding LRHGs expression levels, respectively. The risk score for each patients with LUAD was calculated using the computational formula, and the risk score threshold was measured using the "survminer" package. All patients with LUAD were divided into the high and low-risk groups according to the median risk score.

Prognostic value of the LRHGs model

The OS of the LRHGs high and low-risk groups were analyzed using Kaplan–Meier (KM) curve (https://CRAN.R-project.org/package=survival). In addition, the LRHGs model was evaluated by constructing a LRHGs expression diagram and a time-dependent ROC curve based on the risk scores [20]. The area under the curve (AUC) was also calculated to evaluate the diagnostic efficiency for predicting LUAD survival. The "pheatmap" R package was used to show the distribution of clinicopathological features, including tumor (T) stage, node (N) stage, metastasis (M) stage, age, gender, OS, disease-specific survival (DSS) and progression-free interval (PFI). Cox regression analysis was used to determine whether the risk scores were independent prognostic factors when combined with other clinical characteristics. A nomogram to predict the OS prognosis for LUAD was developed and validated (https://CRAN.R-project.org/package=rms) [21]. A calibration curve was also constructed to evaluate the performance of the LRHGs model in patients with LUAD [22]. The sensitivity and specificity of the risk scores model were evaluated using decision curve analysis (DCA) [23], which is a simple method to evaluate clinical predictive models, diagnostic tests and biomarkers.

GSEA

GSEA is a commonly used computational method for evaluating differentially expressed gene sets between two groups, such as tumor and normal gene sets [24]. To perform the GSEA, the DEG data from both the tumor and normal groups were firstly obtained through the "clusterProfiler" package. The analysis parameters were as follows: seed for 2022, number of calculations for 10,000, and at least 10 genes and a maximum of 500 genes for each gene set. The c2.cp.all.v2022.1.Hs.symbols.gmt gene cluster was subsequently downloaded from the Molecular signatures database 3.0, and KEGG and Hallmark gene sets served as a reference for pathway analysis [25]. Finally, the P-value was adjusted by the Benjamini-Hochberg (BH) method, and P.adj <0.05 and FDR <0.05 were considered as the screening criteria.

GSVA

GSVA was used to explore the biological processes and signaling pathways associated with the LRHGs between tumor and normal samples in TCGA-LUAD datasets [26]. The P-value was adjusted by the BH method, and P.adj <0.05 was considered as the screening criteria. The top 10 pathways were plotted and compared.

Construction of a lncRNA-RNA binding protein (RBP) network

The Encyclopedia of RNA Interactomes database (ENCORI; https://starbase.sysu.edu.cn/) is an open-source platform for revealing the miRNA-ncRNA, miRNA-mRNA, ncRNA-RNA, RNA-RNA, RBP-lncRNA and RBP-mRNA interactions from cross-linking and immunoprecipitation-seq and degradome-seq datasets [27]. The RBPs interacting with the LRHGs were predicted using the ENCORI database and the lncRNA-RBP network was constructed with the screening criteria, cluster number > 1.0 using Cytoscape software [28].

Identification of disease subtype

Consensus clustering is an algorithm based on resampling that identifies each sample and its subgroup, and verifies the rationality of the clustering [29]. To identify and screen the LRHGs in different LUAD subtypes, the consensus clustering analysis was performed using the "ConsensusClusterPlus" package [30]. The consensus clustering number was set between 2 and 8, and 80% of the total samples was clustered 50 times. The conditions, clusterAlg = "km" and distance = "euclidean" were also set. The difference was calculated and analyzed using the Wilcoxon rank sum test and $P < 0.05$ was considered to indicate a statistically significant difference.

Immune infiltration

The high specificity and sensitivity of the immune cell phenotype in tumor microenvironment of patients with LUAD were detected using the single sample (ss)GSEA algorithm [31]. The fraction scores for 28 immune cell subtypes were evaluated via CIBERSORT [32]. Immune infiltration was analyzed and compared by clustering and assessing the risk scores of the two groups. The associations between immune cells in tumor microenvironment and the LRHGs levels were obtained. A dot plot was then constructed using the R packages, "ggplot2" or "pheatmap".

RNA extraction and reverse transcription-quantitative PCR (RT-qPCR)

Total RNA from tissues and serum samples was isolated using TRIzol® reagent (Invitrogen; Thermo Fisher Scientific, Inc.). 500 ng RNA was subsequently reverse-transcribed using the PrimeScript™ RT Master Mix (Perfect Real Time; Takara Biotechnology Co., Ltd.). Briefly, the 10 µl reaction mix was as follows: 500 ng RNA, 2 µl 5X PrimeScript RT Master Mix (for Real Time) and RNase free dH_2O. The mixture was incubated for 15 min at 37˚C and for 5 s at 85˚C. qPCR analysis was conducted using an ABI 7500 instrument and UltraSYBR Green PCR Master mix (CWBio), according to the manufacturer's instructions. The reaction conditions were as follows: 95˚C for 10 min, followed by 40 cycles of 95˚C for 15 s and 60˚C for 1 min. Gene expression was normalized to 18S ribosomal RNA, and the relative gene expression levels were calculated using the $2^{-\Delta\Delta Cq}$ method [33]. The sequences of the primers (designed and synthesized by Sangon Biotech (Shanghai) Co., Ltd.) used for RT-qPCR are listed in Table SI.

ceRNA analysis

We downloaded LUAD gene expression data from the TCGA database [15] and divided the samples into high- and low-HOXD-AS2 expression groups based on the median expression. Differential expression analysis was performed to identify low expression miRNAs (LEmiRNAs) with $\log_2FC < 0$ and $P < 0.05$, and high expression mRNAs (HEmRNAs) with $\log_2FC > 0$ and $P < 0.05$. Pearson correlation analysis was used to investigate the association between miR-4538 and HOXD-AS2. Subsequently, we used the TargetScan [34] and miRWalk [35] databases to predict the target genes of miR-4538, intersecting these with the HEmRNAs and a list of necrosis-related genes [36], ultimately identifying CAMK2B as a necroptosis-related gene. The correlation analysis was also conducted and the Venn diagram were visualized through the "ggplot2" package.

Statistical analysis

All data analyses were conducted using R software (version 4.1.2; (https://mirrors.tuna.tsinghua.edu.cn/CRAN/). For the comparison of two sets of continuous variables, the Student's unpaired t-test was employed. The differences between non-normally distributed variables were compared using the Wilcoxon signed-rank test. Comparison of the survival between groups was investigated using KM analysis and the log-rank test. The AUC value was determined by receiver operating characteristic (ROC) curve analysis. The correlation between groups was calculated by Pearson correlation analysis. $P < 0.05$ was considered to indicate a statistically significant difference.

Results

Identification of the LRHGs

Figure 1 shows a schematic diagram of the methods conducted in the present study. From TCGA-LUAD RNA-seq datasets, 141 DELRs were selected for further analysis and were shown in a volcano plot in Fig. 2A. Among these DELRs, 105 were upregulated, while the remaining were downregulated in LUAD (Fig. 2A, Table SII). To identify the LRHGs, the relationship between the 141 DELRs levels and the OS of patients with LUAD was firstly determined by univariate Cox regression analysis. The results revealed that the 10 lncRNAs (AC004687.1, AC074011.1, AC090371.1, AC100781.1, AC110285.1, AL596087.2, AP000322.2, AP004609.1,

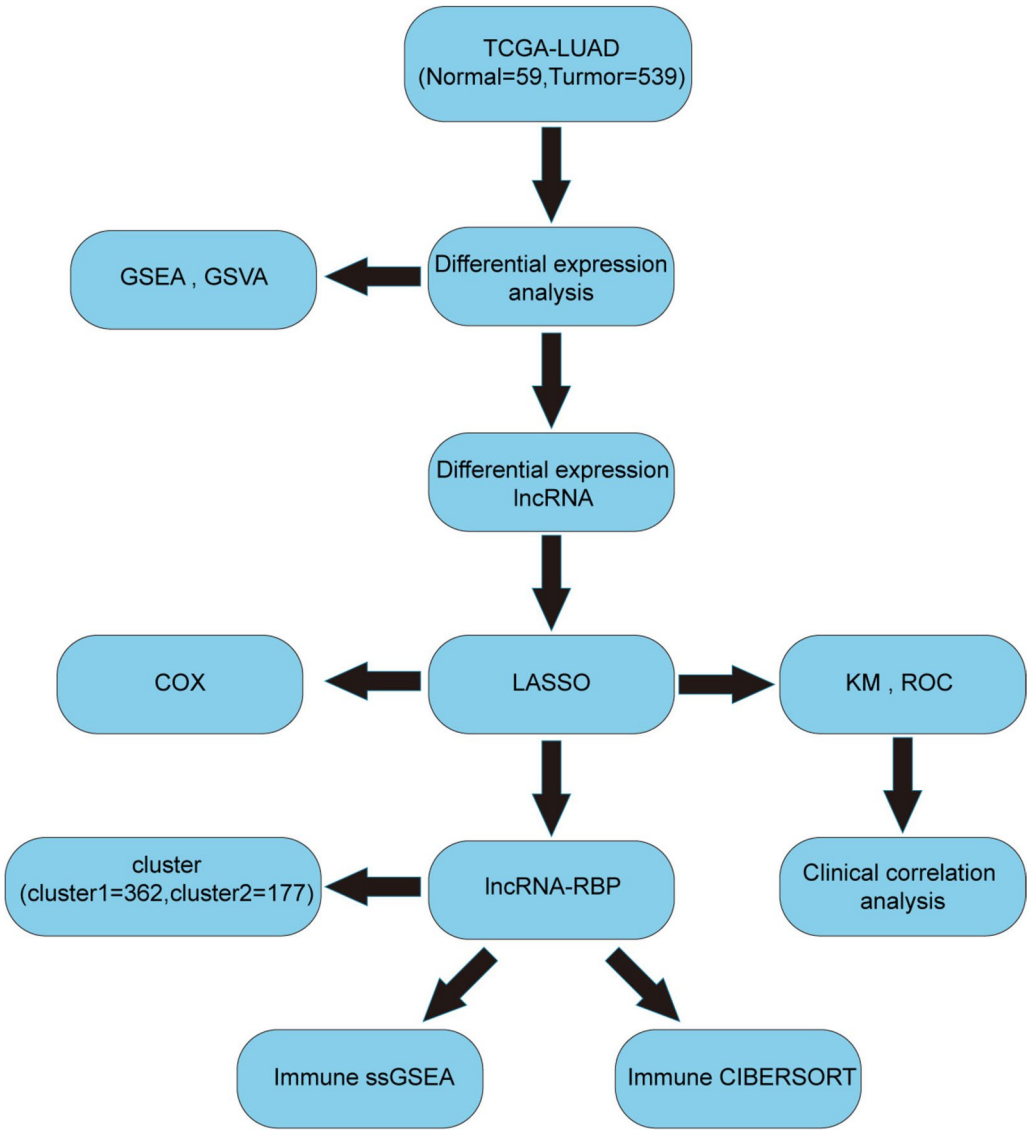

Fig. 1 Schematic diagram of the present study. TCGA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma; LASSO, least absolute shrinkage and selection operator; lncRNA, long non-coding RNA; LRHGs, lncRNA-related hub genes; ROC, receiver operating characteristic; KM, Kaplan–Meier; DEGs, differentially expressed genes; GSEA, Gene Set Enrichment Analysis; GSVA, Gene Set Variation Analysis

HOXD-AS2 and LINC02779) expressions were closely associated with the prognosis of patients with LUAD. LASSO was subsequently used to identify the LRHGs based on DELR expression through the "glmnet" package. It was observed that the key LRHGs model consisted of 9 optimal LRHGs: AC004687.1, AC074011.1, AC100781.1, AC110285.1, AL596087.2, AP000322.2, AP004609.1, HOXD-AS2 and LINC02779 (Fig. 2B). Moreover, a variable trajectory diagram of the 9 LRHGs was analyzed and visualized (Fig. 2C). To construct the risk score formula, the expression level of each of these 9 optimal LRHGs was weighted and represented as its corresponding regression coefficient. In addition, the overall risk score for each patient was calculated using the total weighted values, which has been widely used and validated in medicine [37]. Therefore, the designed risk score formula was as follows: Risk score = (AC004687.1 × 0.7750717) + (AC074011.1 × 0.03097888) + (AC100781.1 × 1.36442837) + (AC110285.1 x -1.24924136) + (AL596087.2 × 1.24701852) + (AP000322.2 × 21.393488) + (AP004609.1 × 0.84695884) + (HOXD-AS2 x -0.75041734) + (LINC02779 × 2.38601789). The median risk score was then obtained according to the scores calculated for each sample. Subsequently, the samples were divided into high and low-risk

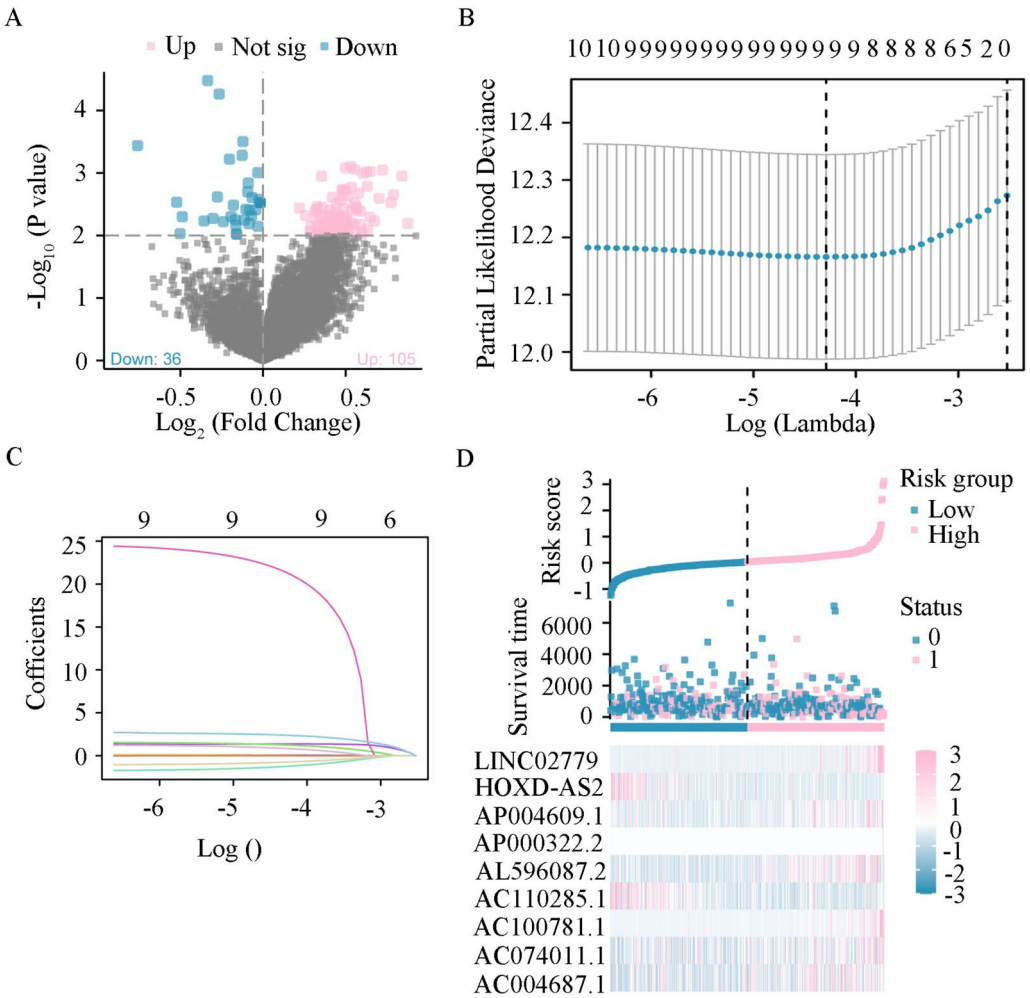

Fig. 2 Construction of a LRHGs model. **A** Volcano plot of the LRHGs in TCGA-LUAD datasets. **B** LRHGs model diagram following LASSO regression. **C** Variable trajectory diagram of the LASSO regression. **D** Risk factor map for the LRHGs model. TCGA, The Cancer Genome Atlas; LASSO, least absolute shrinkage and selection operator; LUAD, lung adenocarcinoma; LRHGs, long non-coding RNA-related hub genes

groups based on the median (Fig. 2D). Compared with the high-risk group, the OS of the low-risk group indicated a relatively improved prognosis for these patients (Fig. 2D).

Prognosis significance of the LRHGs in LUAD

To further investigate the expression of the 9 LRHGs, TCGA-LUAD datasets were analyzed. It was observed that AC100781.1, AC110285.1, AL596087.2 and HOXD-AS2 levels were abnormally expressed in LUAD ($P < 0.001$; Fig. 3A). Among these LRHGs, the expression levels of AC110285.1, AL596087.2 and HOXD-AS2 were upregulated, while AC100781.1 was downregulated in LUAD (Fig. 3A). To evaluate the role of AC100781.1, AC110285.1, AL596087.2 and HOXD-AS2 in the diagnosis of LUAD, the AUC values were determined, which were 0.626, 0.755, 0.774 and 0.683 for AC100781.1, AC110285.1, AL596087.2 and HOXD-AS2, respectively, demonstrating the signature had a comparatively good prediction value (Fig. 3B and C). In addition, according to the risk scores, the 1, 3 and 5-year OS AUCs were 0.683, 0.623 and 0.601, respectively (Fig. 3D), suggesting that the signature had a relatively good diagnostic value.

Moreover, the OS of patients in the high and low-risk groups were analyzed using KM plotter. As the expression levels of AC100781.1, AP000322.2 and LINC02779 were low, these data were excluded from the analysis. The risk score of the remaining LRHGs was analyzed. It was revealed that the high-risk group patients had a poor prognosis ($P < 0.001$; Fig. 3E). Subsequently, the relationships between the AC004687.1, AC074011.1, AC110285.1, AL596087.2, AP004609.1 and HOXD-AS2 expression levels and OS were investigated. The data showed that the patients with higher AC074011.1 and AL596087.2 expression levels had a

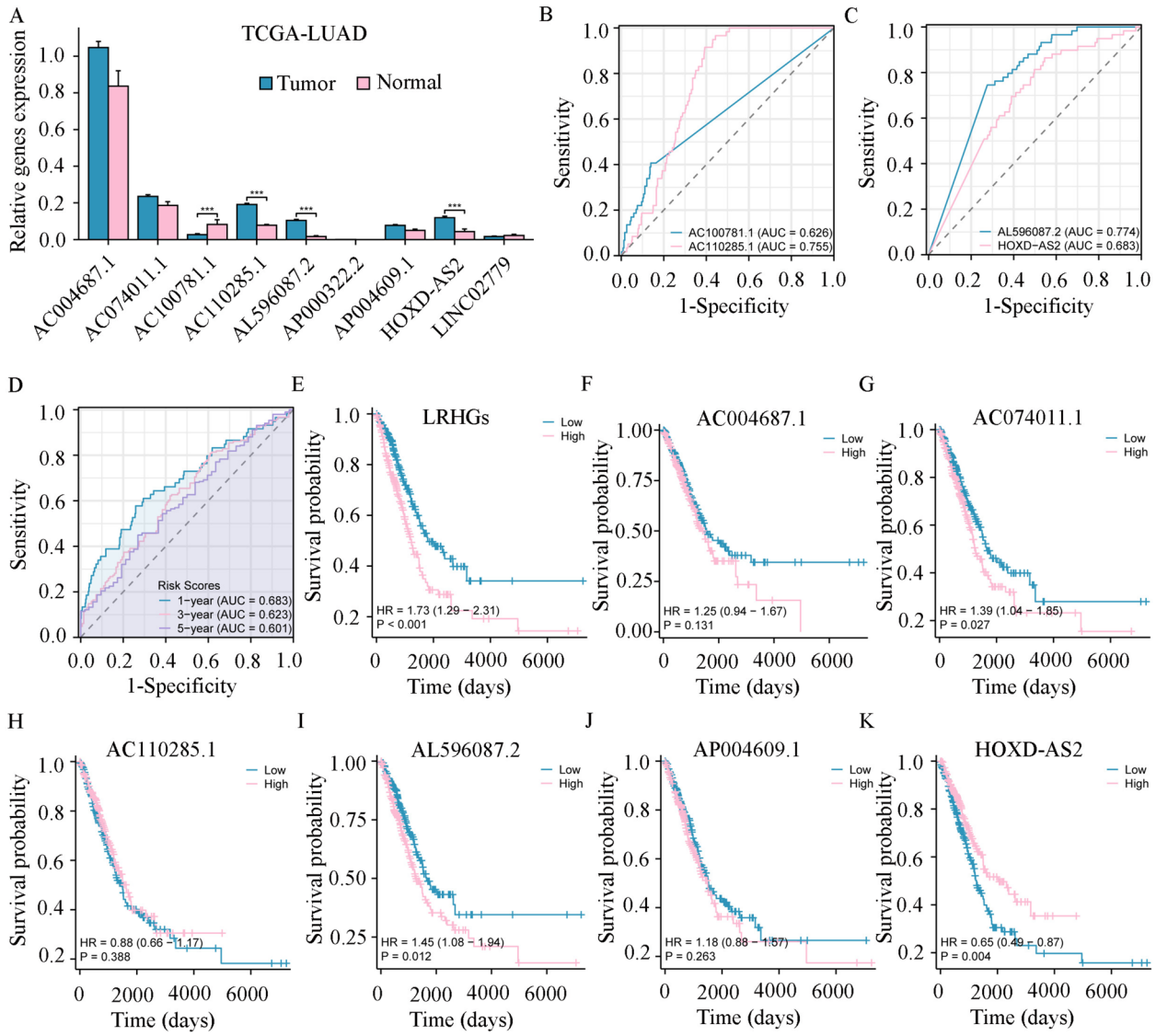

Fig. 3 Prognosis of the LRHGs. **A** Expression of the LRHGs in TCGA-LUAD datasets. **B** Diagnostic value of AC100781.1 and AC110285.1. **C** Diagnostic value of AL596087.2 and HOXD-AS2. **D** Receiver operating characteristics curves for evaluating the diagnostic value of the LRHGs model in the high and low-risk groups at 1, 3, and 5 years. **E** The survival curves of patients with LUAD in the high and low-risk groups. The association between (**F**) AC004687.1, **G** AC074011.1, **H** AC110285.1, **I** AL596087.2, **J** AP004609.1, or **K** HOXD-AS2 expression and OS of patients with LUAD. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ns, not significant. LUAD, lung adenocarcinoma; TCGA, The Cancer Genome Atlas; KM, Kaplan–Meier; OS, overall survival; LRHGs, long non-coding RNA-related hub genes; ROC, receiver operating characteristic; AUC, area under the curve; HOXD-AS2, HOXD cluster antisense RNA 2

poor prognosis ($P = 0.027$ and $P = 0.012$, respectively; Fig. 3G and I), while patients with higher HOXD-AS2 expression had an improved prognosis ($P = 0.004$; Fig. 3K). The remaining LRHGs were not significantly associated with the OS of patients (Fig. 3F-J). These results suggested that the LRHGs were notably distinguished the risk level in different groups, and therefore the signature may be an independent prognostic indicator in patients with LUAD.

Clinical significance of the LRHGs in TCGA-LUAD

To investigate the clinical significance of the LRHGs in the high and low-risk groups in TCGA-LUAD datasets, data on all clinicopathological factors, including T stage, N stage, M stage, stage, age, sex, OS, DSS and PFI, were collected and analyzed. It was revealed that T stage, N stage, M stage, stage, sex, OS and DSS were statistically varied in the two groups (Fig. 4A-I). However, age and PFI were not significantly different (Fig. 4E, I and Table

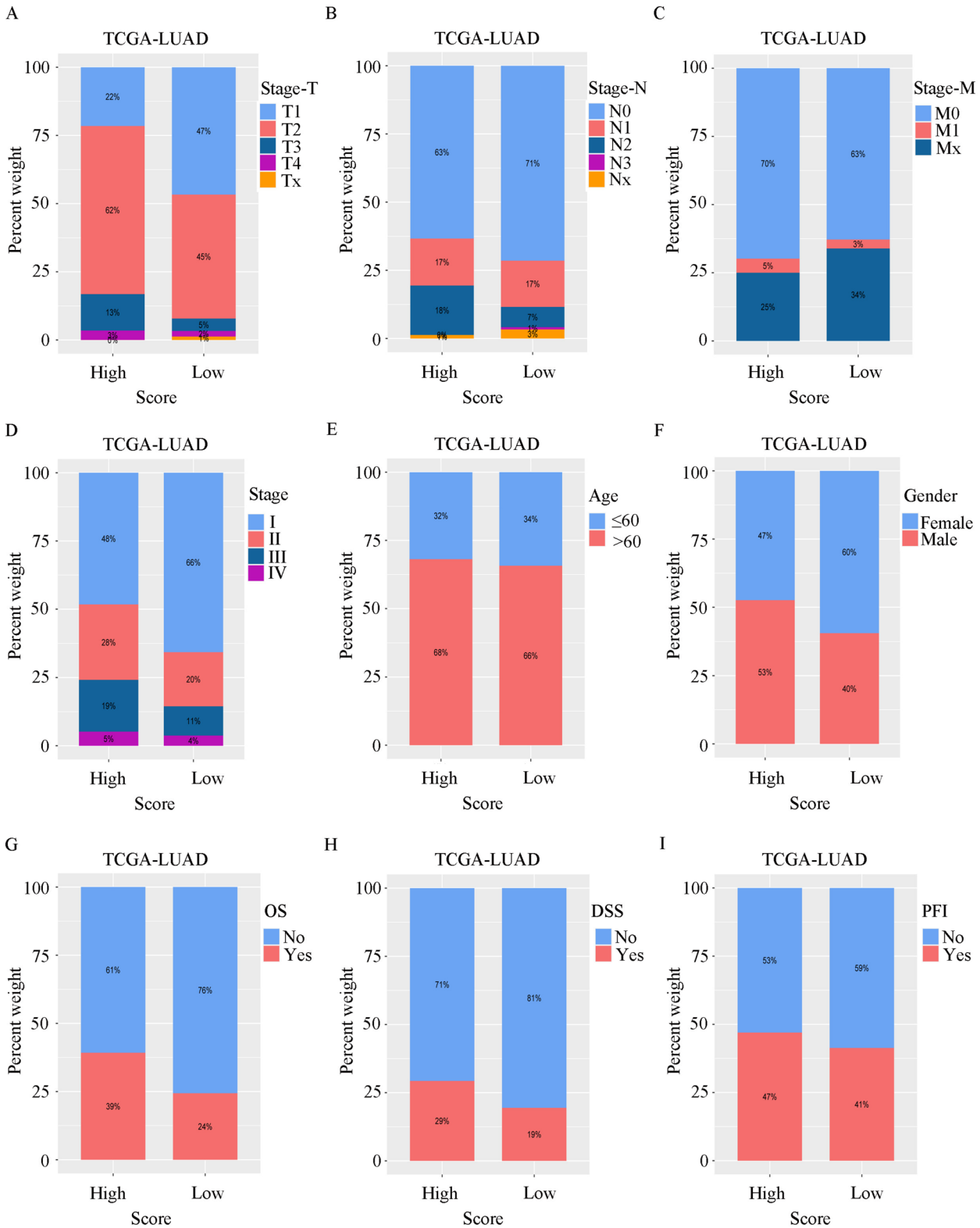

Fig. 4 Clinical significance of the LRHGs. Percentage weight of (**A**) T stage, (**B**) N stage, (**C**) M stage, (**D**) stage, (**E**) age, (**F**) sex, (**G**) OS, (**H**) DSS and (**I**) PFI between the high and low-risk groups. TCGA, The Cancer Genome Atlas; OS, overall survival; DSS, disease-specific survival; PFI, progression-free interval

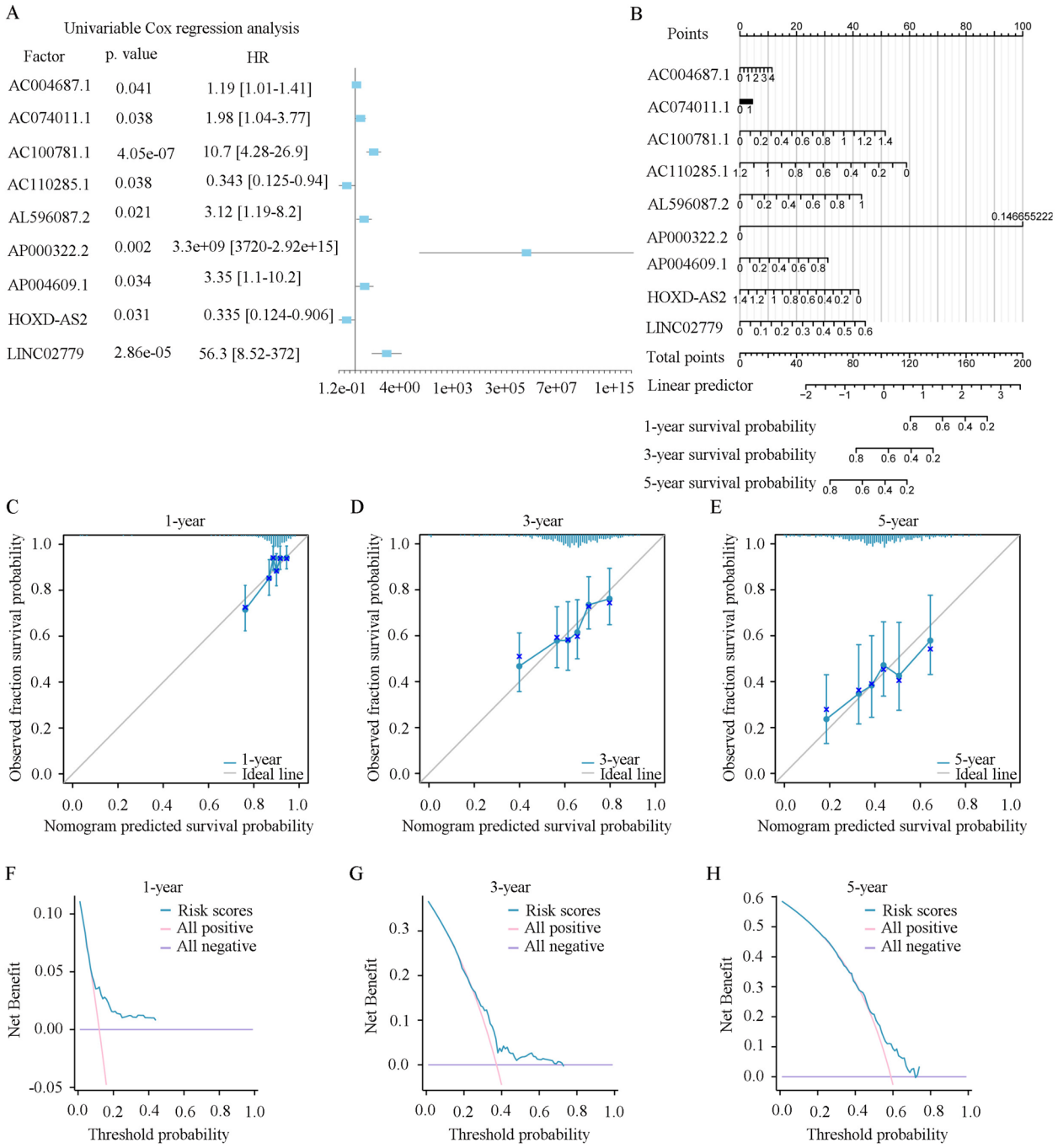

Fig. 5 Construction of a LRHGs prognostic model. **A** Univariate Cox regression analysis. **B** A nomogram of the LRHGs prognostic model. Calibration curves for **C** 1-year, **D** 3-year and **E** 5-year. DCA of (**F**) 1-year, (**G**) 3-year and (**H**) 5-year. TCGA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma; LRHGs, long non-coding RNA-related hub genes; DCA, decision curve analysis

SIII). In addition, the univariate Cox regression analysis indicated that the 9 LRHGs were closely associated with the OS of patients with LUAD ($P < 0.05$; Fig. 5A and Table 1), which served as an independent prognostic indicator in LUAD. A nomogram was further constructed to forecast the LUAD 1, 3, and 5-year survival. It was demonstrated that the value of AP000322.2 was notably varied compared with the rest of the 8 LRHGs (Fig. 5B). In addition, to evaluate the predictive efficacy of the LRHGs model, calibration curves were constructed. The findings

Table 1 COX analysis of dataset TCGA-LUAD

Factor	Univariate analysis		Multivariate analysis	
	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
AC004687.1	1.19 [1.01—1.41]	0.0409	1.11 [0.912—1.35]	0.301
AC074011.1	1.98 [1.04—3.77]	0.0375	1.13 [0.527—2.43]	0.752
AC100781.1	10.7 [4.28—26.9]	4.05E-07	3.79 [1.08—13.3]	0.038
AC110285.1	0.343 [0.125—0.94]	0.0375	0.169 [0.0524—0.542]	0.0028
AL596087.2	3.12 [1.19—8.2]	0.021	4.76 [1.81—12.5]	0.0016
AP000322.2	3.3e + 09 [3720—2.92e + 15]	0.00171	5.5e + 10 [51500—5.86e + 16]	0.000479
AP004609.1	3.35 [1.1—10.2]	0.0338	3.5 [0.869—14.1]	0.0781
HOXD.AS2	0.335 [0.124—0.906]	0.0312	0.337 [0.121—0.935]	0.0367
LINC02779	56.3 [8.52—372]	2.86E-05	14.6 [1.36—156]	0.0268

observed that the predicted survival rates were consistent with the observed 1, 3, and 5-year survival rates, and that the 5-year prediction was more consistent than the 3 and 1-year predictions (Fig. 5C-E). Finally, a DCA was conducted to evaluate the clinical effectiveness of the 1, 3, and 5-year predictive signatures, and the findings also showed that the 5-year prediction was more consistent than the 3 and 1-year predictions (Fig. 5F-H). The aforementioned results verified that the LRHGs were linked to the OS of patients with LUAD in the constructed predictive model.

Identification of LRHG-associated pathways

To identify the relationship between expression of the LRHGs and the underlying mechanism in LUAD, LRHG-associated pathways were identified by GSEA. It was observed that expression of the LRHGs was associated with tumor-associated pathways, such as the 'integrin 4' [normalized enrichment scores (NES) = -2.009 and nominal $P = 0.045$), 'advanced-glycosylation end product receptor signaling' (NES = -2.018 and nominal $P = 0.037$), 'mitochondrial-translation' (NES = 1.707 and nominal $P < 0.001$) and 'fatty-acid metabolism' (NES = 1.485 and nominal P = 0.006) pathways (Fig. 6A-E and Table SIV), leading to tumor metastasis and abnormal metabolism. Following GSVA enrichment analysis, the top 10 positively enriched pathways included 'breast cancer luminal', 'farmer breast cancer cluster', 'thyroid cancer poor survival', 'cell cycle', 'reactome phosphorylation', 'aging middle', 'DNA replication', 'unwinding', 'breast cancer kinome' and 'crosby-E2F4 targets' (Fig. 6F, 6G and Table SV). The top 10 negatively enriched pathways included 'prostate cancer', 'erythrocytes take up oxygen and release carbon dioxide', 'muscarinic acetylcholine receptors', 'alveolar epithelium', 'ficolins bind to repetitive carbohydrate structures on the target cell surface', 'hypoxia', 'ligand receptor interactions', 'LDL pathway', 'defective CSF2RB causes SMDP5', and 'activated NTRK2 signals through PI3K' (Fig. 6F, G and Table SV). These results revealed that the LRHGs probably promoted tumor progression via influencing cell cycle or DNA replication in LUAD.

Construction of a lncRNA-RBP interaction network

The ENCORI database was used to explore the interactions between the LRHGs and RNA binding proteins (RBPs). A lncRNA-RBP interaction network was subsequently visualized using Cytoscape software (Fig. 7 and Table SVI). The lncRNA-RBP interaction network was composed of 5 LRHGs and 40 RBPs (Fig. 7 and Table SVI).

Construction of a LUAD-associated cluster subtype

To investigate the expression of the LRHGs in LUAD-associated subtypes, the consistent clustering method was conducted using the "ConsensusClusterPlus" package. Subsequently, two LUAD clusters (clusters 1 and 2) were identified and analyzed (Fig. 8A). Among them, cluster 1 contained 362 LUAD cases and cluster 2 contained 177 cases. The cumulative distribution function (CDF) curve of consistent clustering and the Delta curve of the CDF were also visualized (Fig. 8B and C). As shown in Fig. 8B, the result of consistent clustering was the best when k = 2 recognized as the number of unsupervised clusters. Additionally, the differential expressions of the 9 LRHGs in the two LUAD clusters were determined by the "pheatmap" package, and it was revealed that the expression of AC004687.1 in cluster 2 was notably increased compared with in cluster 1, while the remaining LRHG levels were not different (Fig. 8D). Then, principal component analysis was performed and it was showed that the LRHGs levels in the two LUAD clusters were significantly varied (Fig. 8E). As shown in Fig. 8F, the expression levels of AC004687.1, AC074011.1,

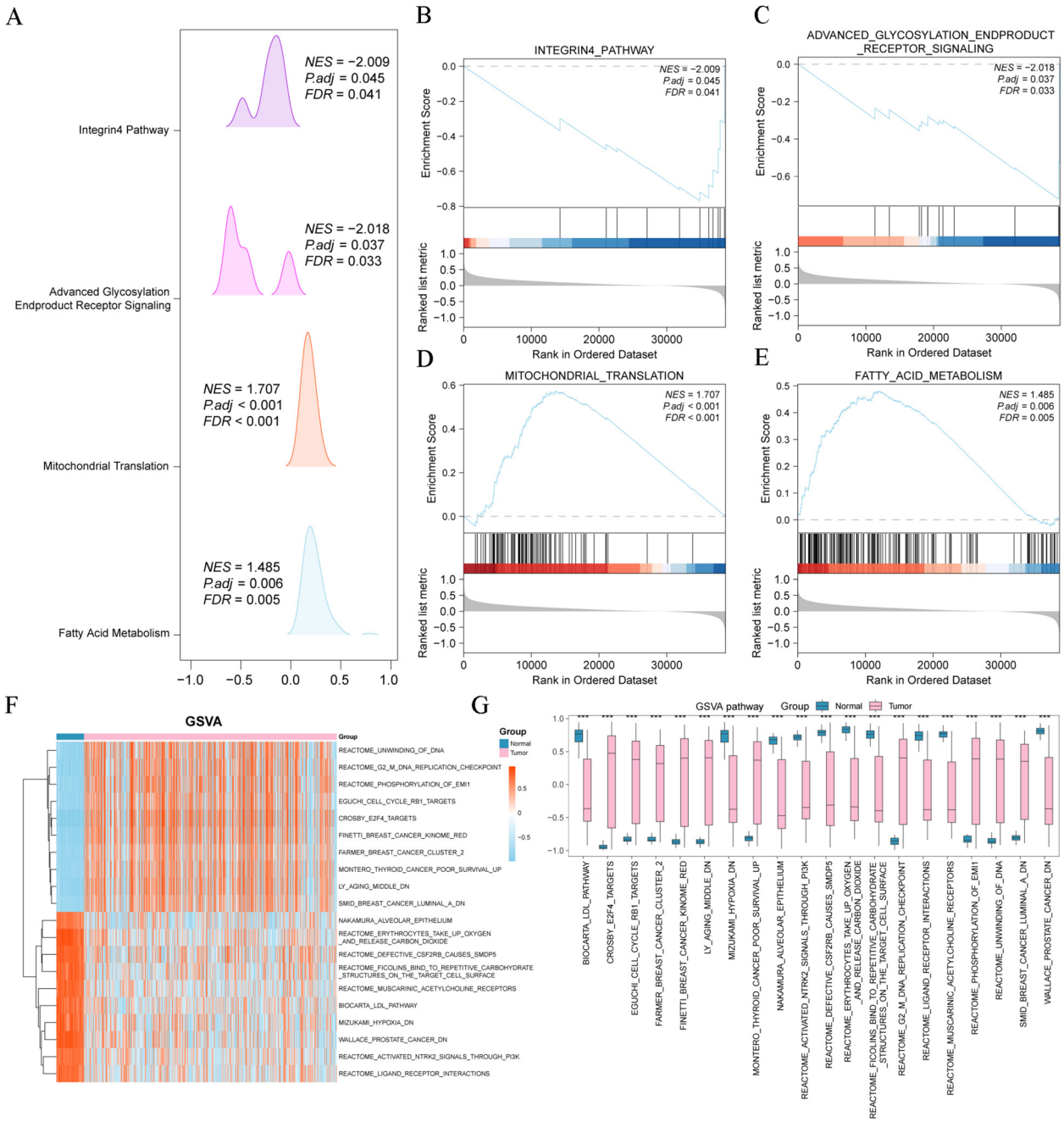

Fig. 6 GSEA and GSVA. **A** GSEA. **B** Intergin 4 pathway. **C** Advanced glycosylation end product receptor signaling. **D** Mitochondrial translation. **E** Fatty acid metabolism. **F** Heatmap of the top 20 pathways following GSVA. **G** Comparison curve of the GSVA. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ns, not significant. GSEA, Gene Set Enrichment Analysis; GSVA, Gene Set Variation Analysis

AC110285.1 and AP004609.1 in cluster 2 were higher than that in cluster 1. The stromal score, immune score, estimate score and tumor purity of the two LUAD clusters were further determined by the "estimate" package. It was observed that the three score types and tumor purity were markedly changed (Fig. 8G-J).

Immune infiltration profile

To investigate the association between the LRHGs and immune function, immune cell subgroup analysis was conducted by the ssGSEA algorithm. It was indicated that 23 immune cells (activated B cells, activated CD4 T cells, activated dendritic cells, CD56 dim natural killer cells, central memory CD4 T cells, central memory CD8 T

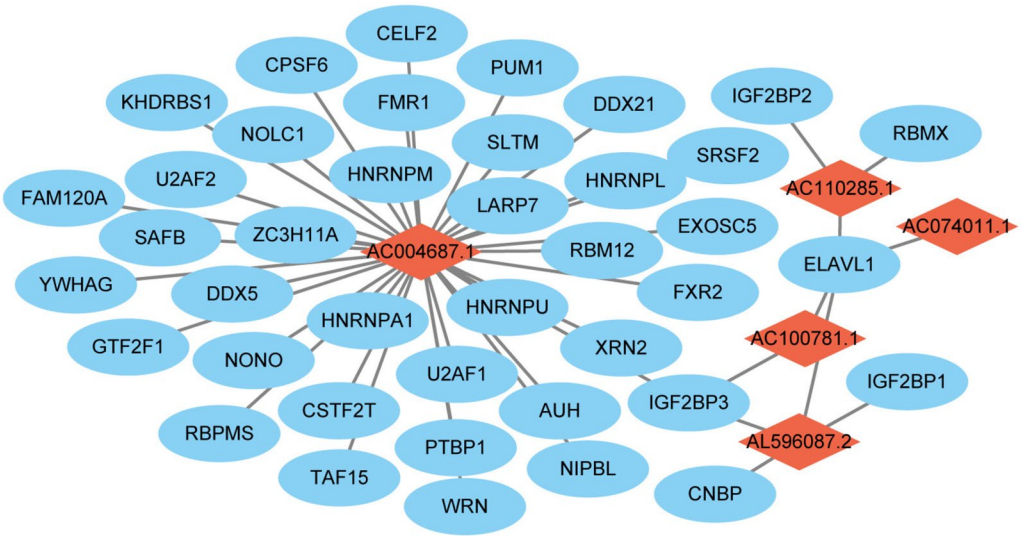

Fig. 7 Construction of a lncRNA-RBP interaction network. LncRNA, long non-coding RNA; RBP, RNA-binding protein

cells, effector memory CD4 T cells, effector memory CD8 T cells, eosinophil, γδ T cells, immature B cells, immature dendritic cells, macrophages, mast cells, myeloid-derived suppressor cells, monocytes, natural killer cells, natural killer T cells, neutrophils, plasmacytoid dendritic cells, T follicular helper cells, type 1 T helper cells and type 2 T helper cells) were notably varied between the LUAD and matched normal tissue groups (Fig. 9A). Possible correlations between the 23 immune cells and other immune cells were subsequently determined. The findings revealed that the infiltration abundance of the 23 immune cells was positively correlated with the other immune cells (Fig. 9B). Meanwhile, the correlations between the infiltration abundance of the 23 immune cells and the 9 LRHGs levels were investigated. It was showed that expression of 3 of the LRHGs (AC110285.1, AP004609.1 and HOXD-AS2) was observably negative correlation with the infiltration abundance of the 23 immune cells (Fig. 9C), while AC004687.1 expression was positively correlated (Fig. 9C). However, the expression of AP000322.2 was not significantly associated with the infiltration abundance of the 23 immune cells (data not shown). To further certify the aforementioned findings, the correlations between the infiltration abundance of immune cells and expression of the LRHGs were investigated by the CIBERSORT algorithm. It was demonstrated that several immune cells infiltration occurred in the tumor group, including resting dendritic cells, eosinophils, M0 macrophages, M1 macrophages, M2 macrophages, resting mast cells, monocytes, neutrophils, resting NK cells, plasma cells, activated memory CD4 T cells, resting memory CD4 T cells, T follicular helper cell and regulatory T cells (Fig. 9D). However, the infiltration abundance of the 15 immune cells was negatively associated with the other immune cells, and expression of the 9 LRHGs was positively or negatively linked to the infiltration abundance of the 15 immune cells (Fig. 9E and F). These results demonstrated that the LRHGs were notably correlated with the tumor immune microenvironment in LUAD.

Validation of LRHG expression

To verify the expression levels of the LRHGs in LUAD, the mRNA levels of the LRHGs in 21 paired LUAD and matched normal lung tissues were quantified using RT-qPCR. Consistent with the expression pattern of the LRHGs in TCGA datasets, it was observed that HOXD-AS2 was highly expressed in LUAD tissues compared with the matched normal tissues ($P < 0.001$; Fig. 10A and B). However, the expression levels of the other LRHGs

(See figure on next page.)

Fig. 8 Identification of LUAD-associated subtypes. **A** Consistent clustering graph (k = 2). **B** CDF curve of consistent clustering. **C** Delta curve of the CDF. **D** Heatmap of the LRHGs expression in clusters 1 and 2. **E** PCA of the LRHGs in clusters 1 and 2. **F** Comparison of the LRHGs expression in clusters 1 and 2. **G** Stromal score. **H** Immune score. **I** Estimate score. **J** Tumor purity. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ns, not significant. TCGA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma; LRHGs, long non-coding RNA-related hub genes; CDF, Cumulative Distribution Function; PCA: Principal Component Analysis

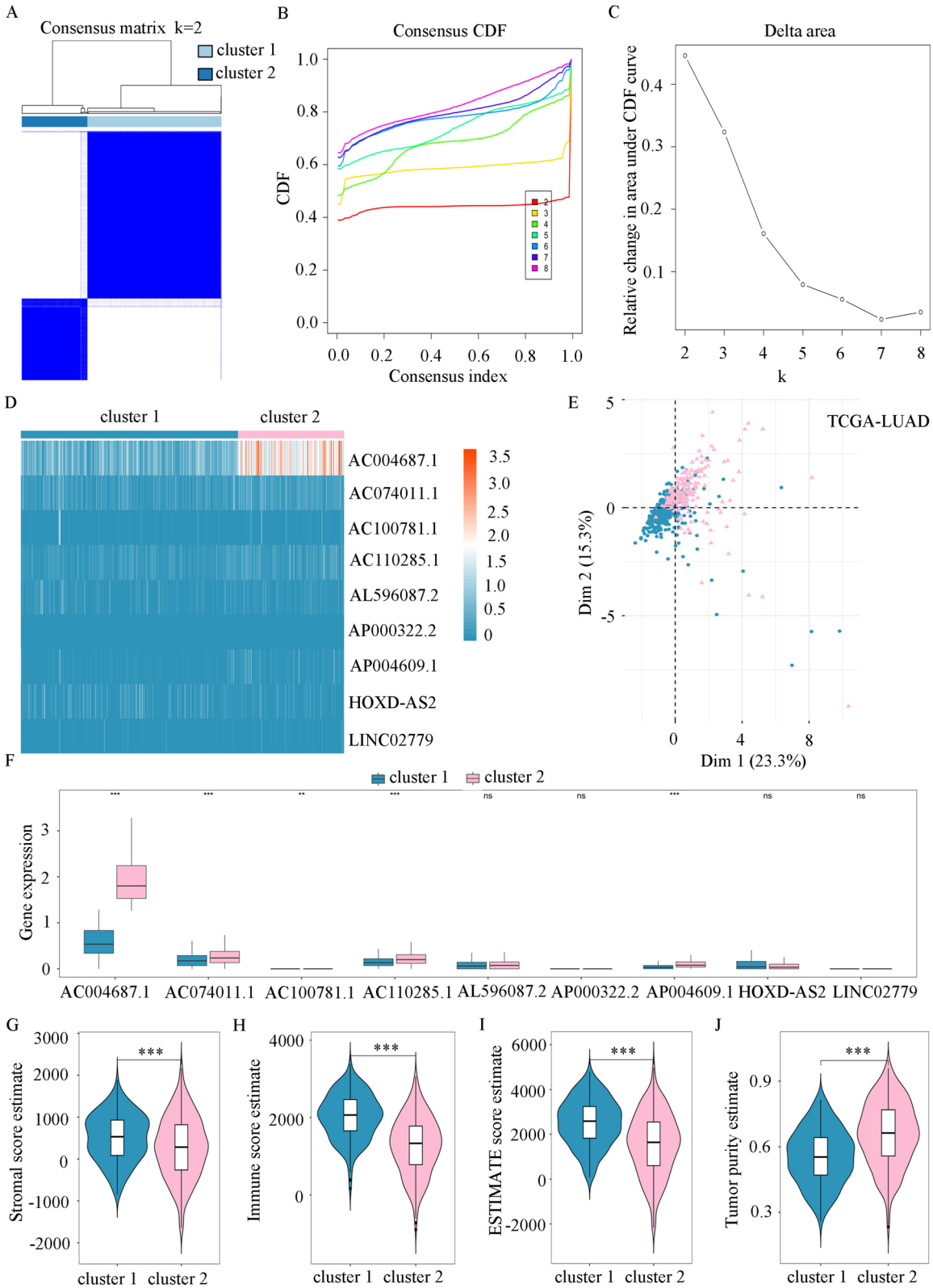

Fig. 8 (See legend on previous page.)

(AC100781.1, AC004687.1, AC110285.1, AC074011.1, AL596087.2, AP000322.2, AP004609.1 and LINC02779) were not different between the LUAD tumor and matched normal tissues (Fig. 10A). Further analysis indicated that patients with LUAD with high HOXD-AS2 expression had a shorter OS [n = 601; hazard ration (HR) = 1.44; P = 0.023] and recurrence-free survival (n = 366; HR = 1.6; P = 0.038) time than patients with low HOXD-AS2 expression (Fig. 10C and D).

Diagnostic role of HOXD-AS2 in LUAD

To explore the diagnostic role of HOXD-AS2, the serum HOXD-AS2 levels of patients with LUAD and healthy donors were firstly detected. It was demonstrated that serum HOXD-AS2 expression was significantly upregulated in patients with LUAD compared with in the healthy controls (P < 0.001; Fig. 11A). Additionally, the associations between serum HOXD-AS2 levels and the clinicopathological characteristics of patients with LUAD were investigated. The results revealed that the serum HOXD-AS2 level was closely associated with the TNM stage, lymph node metastasis and tumor differentiation (P < 0.01; Fig. 11B-D). However, there were no associations with sex, age, smoking history and tumor size (Table 2). Finally, ROC curve analysis revealed that the AUC of HOXD-AS2 (0.825) was higher than that of CYFRA21-1 (0.618), and the sensitivity and specificity of HOXD-AS2 were 0.765 and 0.867, respectively (Fig. 11E). However, the sensitivity and specificity of CYFRA21-1 were only 0.632 and 0.511, respectively. These data suggested that serum HOXD-AS2 could serve as a novel biomarker for the early diagnosis of LUAD.

CeRNA analysis of HOXD-AS2 in LUAD

Based on the competing endogenous RNA (ceRNA) mechanism [38], we categorized the TCGA-LUAD samples into high- and low-HOXD-AS2 expression groups based on the median expression. Differential analysis was subsequently conducted to identify differentially expressed miRNAs and mRNAs. miRNA with $\log_2$FC < 0 and P < 0.05 was classified as a low expression miRNA (LEmiRNA), while mRNA with $\log_2$FC > 0 and P.adj < 0.05 was used as a high expression mRNA (HEmRNA). Ultimately, three eligible LEmiRNAs were identified, including miR-93, miR-4538, and miR-6859–4. Pearson correlation analysis was used to investigate the associations between HOXD-AS2 and those three miRNAs. It was indicated that miR-4538 showed notably negative correlation with HOXD-AS2 (Fig. 12A), but miR-93 and miR-6859–4 had not correlation (data not shown). Subsequently, the targeted mRNAs of miR-4538 were conducted using the TargetScan [34] and miRWalk [35] databases, integrating the HEmRNAs and a list of genes associated with necrosis to create a Venn diagram (Fig. 12B). CAMK2B, a gene related to necroptosis, was identified and exhibited a strongly positive correlation with HOXD-AS2 (Fig. 12C). These data implied that a potential regulatory network of HOXD-AS2/miR-4538/CAMK2B in LUAD progression, providing a novel therapeutic target for this disease.

Discussion

LUAD is the most common histopathological subtype of lung cancer that threatens younger female individuals who do not smoke, and its incidence of morbidity and mortality is persistently increasing yearly [1]. At present, although the clinical treatment for LUAD has achieved encouraging progress, treatment still fails to improve the outcome of patients with LUAD. Therefore, identifying novel potential biomarkers for LUAD is important for the early diagnosis, prognosis management and therapy development. Increasing evidence has revealed that disease-related lncRNA signatures have a critical role in the diagnosis, prognosis, progression and tumor immune microenvironment in LUAD, providing a new perspective for prognostic prediction, personalized immune intervention and therapeutic regimens in LUAD [39–42]. However, the function of LRHGs in LUAD has not been fully understood. In the present study, the DELRs associated with LUAD progression were evaluated, and a comprehensive analysis was conducted to screen candidate LRHGs. Then, a corresponding lncRNA-RBP interaction network was constructed to study the potential underlying mechanisms of the LRHGs, and a LRHG-associated diagnostic and prognostic model were also built, which showed to provide a trustworthy diagnostic and prognostic model

(See figure on next page.)

Fig. 9 Immune infiltration analysis. **A** Immune infiltration analysis using single sample Gene Set Enrichment Analysis. **B** Correlation analysis between the infiltration abundance of 28 immune cells and other immune cells. **C** Correlation analysis between the infiltration abundance of 28 immune cells and expression of the LRHGs. **D** Immune infiltration analysis using CIBERSORT. **E** Correlation analysis between the infiltration abundance of 15 immune cells and other immune cells. **F** Correlation analysis between the infiltration abundance of 15 immune cells and expression of the LRHGs. *P < 0.05, **P < 0.01, ***P < 0.001, ns, not significant. LRHGs, long non-coding RNA-related hub genes; CIBERSORT, cell type Identification By Estimating Relative Subtypes Of RNA Transcripts

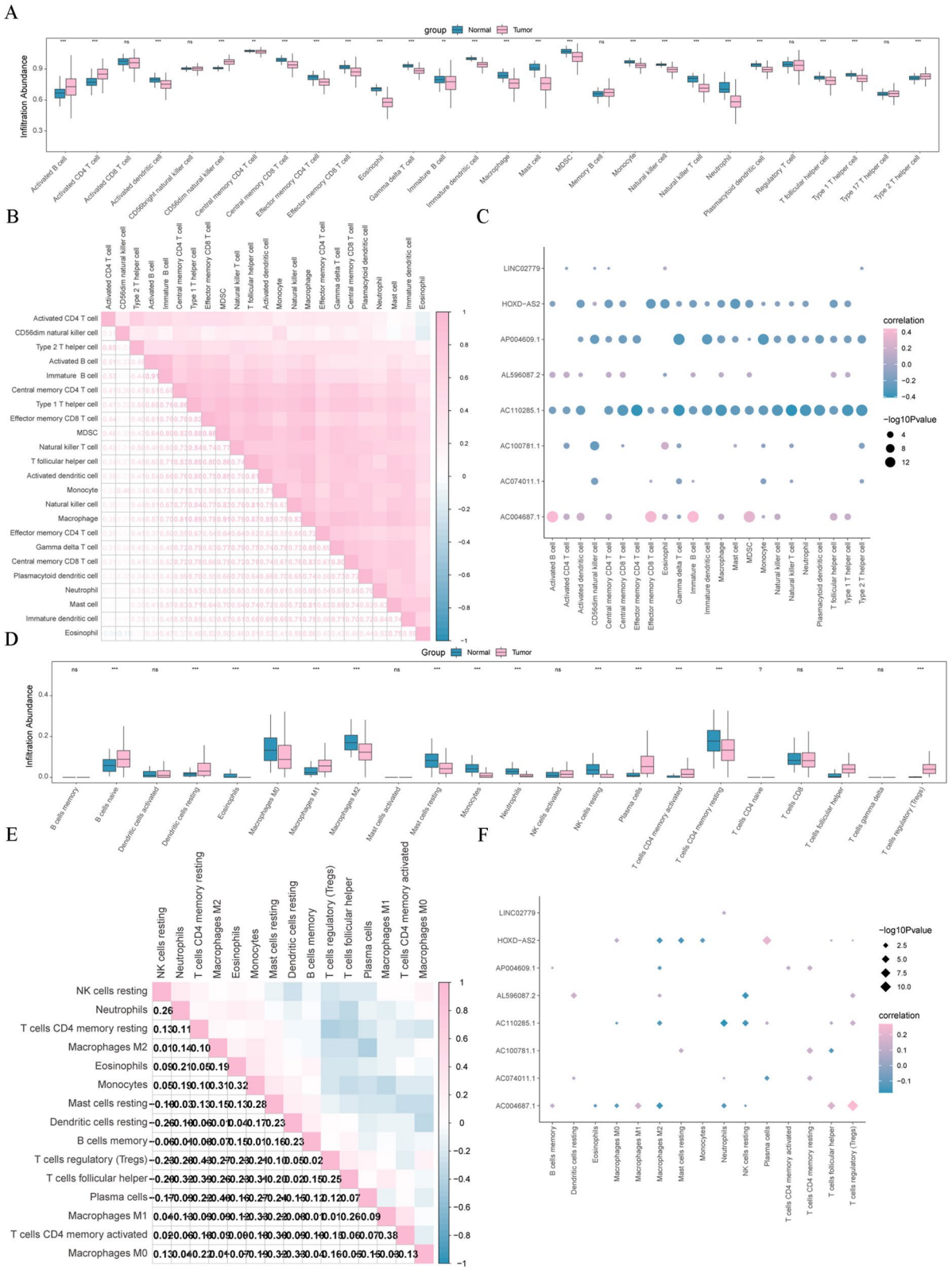

Fig. 9 (See legend on previous page.)

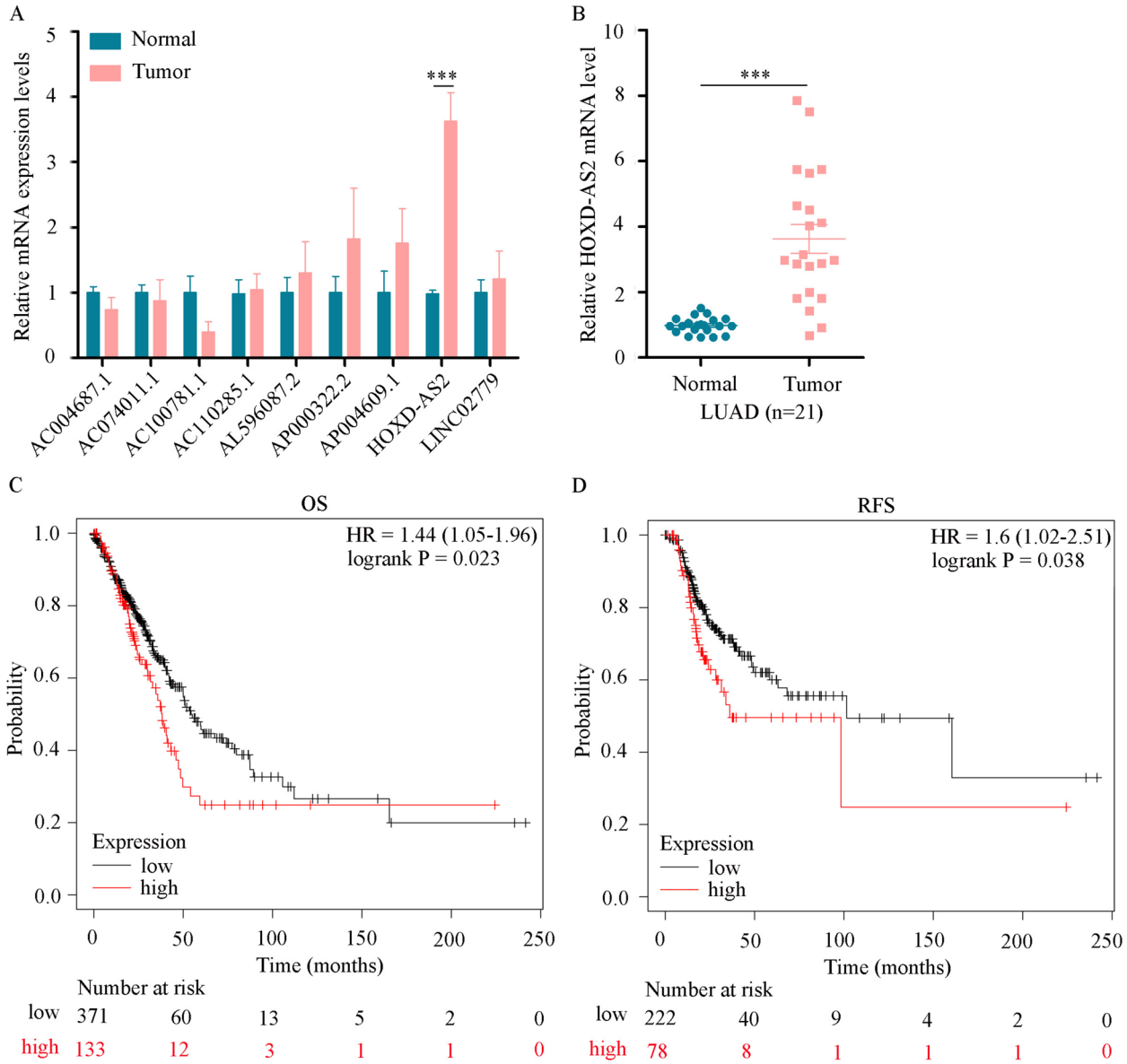

Fig. 10 Verification of the LRHGs expression level. **A** RT-qPCR analysis of the mRNA levels of the LRHGs in 21 paired LUAD tumor and normal tissues. **B** RT-qPCR analysis of the HOXD-AS2 mRNA level in 21 paired LUAD tumor and normal tissues. **C** KM analysis of the overall survival data from TCGA-LUAD datasets containing 601 patients. **D** KM analysis of the recurrence-free survival data from TCGA-LUAD datasets containing 366 patients. ***$P < 0.001$. KM, Kaplan–Meier; LUAD, lung adenocarcinoma; HOXD-AS2, HOXD cluster antisense RNA 2; TCGA, The Cancer Genome Atlas; LRHGs, long non-coding RNA-related hub genes; RT-qPCR, reverse transcription-quantitative PCR

for patients with LUAD. Moreover, LRHG- associated networks based on GSEA and GSVA were identified to certify their affected signaling pathways, revealing promising mechanisms linked to LUAD progression.

Studies have previously demonstrated that a disease-related lncRNA model was constructed using LASSO algorithm for predicting the prognosis, diagnosis, clinical outcome, immune landscape and progression based on TCGA datasets [41, 43, 44]. The 16 tryptophan metabolism-associated lncRNA prognostic model was constructed based on TCGA-LUAD datasets (the training cohort) and Gene Expression Omnibus (GEO) database (the validation cohort) [41]. Jiang et al. a novel anoikis-related lncRNA signature was developed using LASSO algorithm for predicting the prognosis and immune response in pancreatic adenocarcinoma. In the present study, a novel LRHGs model was constructed for predicting the diagnosis and prognosis of LUAD.

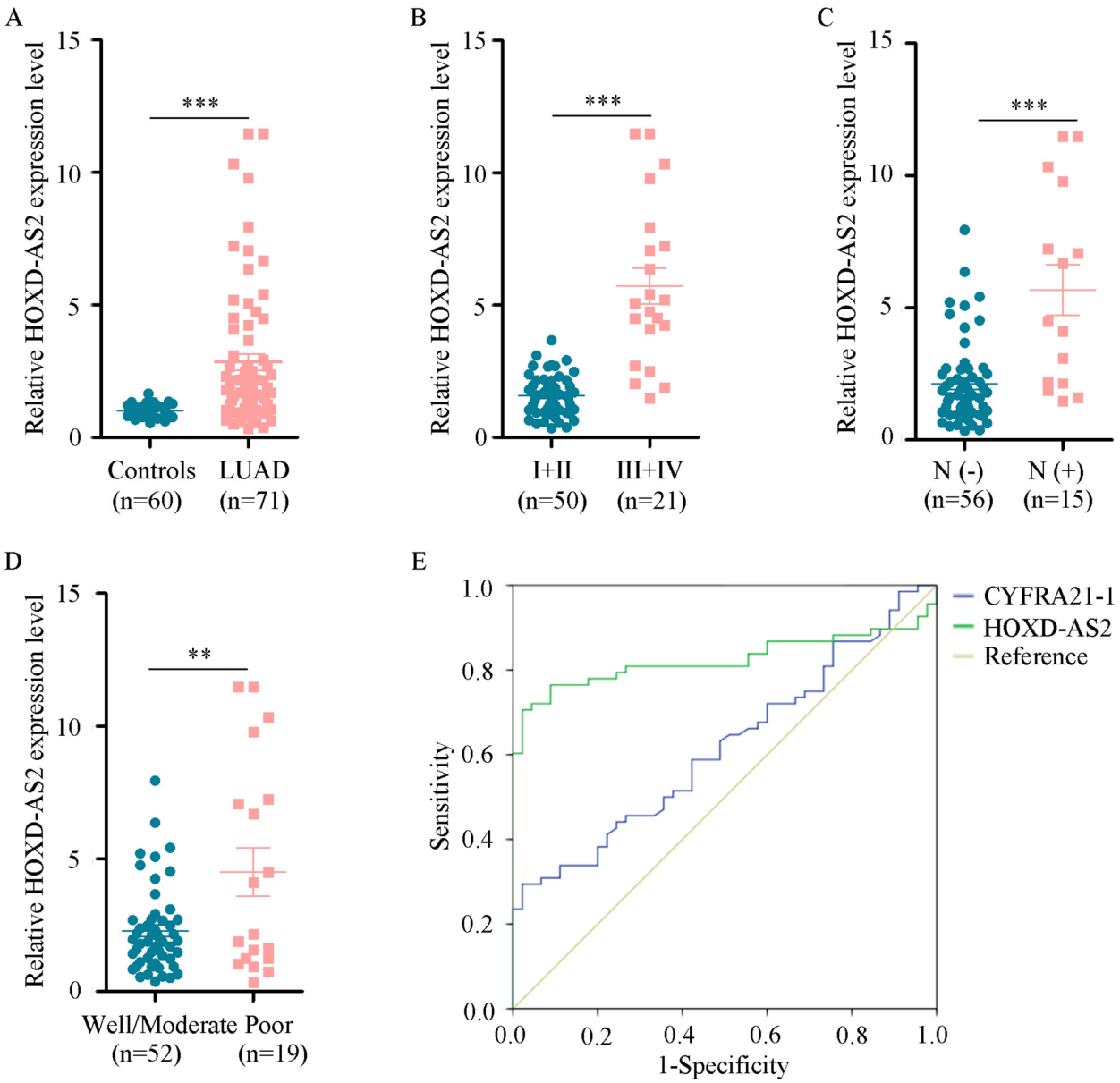

Fig. 11 Diagnostic value of serum HOXD-AS2 in LUAD. **A** Reverse transcription-quantitative PCR analysis of the serum HOXD-AS2 mRNA level in LUAD and healthy controls samples. Associations between the serum HOXD-AS2 level and the (**B**) TNM stage, (**C**) lymph node metastasis and (**D**) tumor differentiation of patients with LUAD. **E** The diagnostic value of serum HOXD-AS2 and CYFRA21-1 in LUAD was measured by receiver operating curve analysis. $^{**}P < 0.01$, $^{***}P < 0.001$. LUAD, lung adenocarcinoma; HOXD-AS2, HOXD cluster antisense RNA 2. CYFRA21-1, cytokeratin-19 fragment

The univariate and multivariate Cox regression analyses indicated that 9 LRHGs were closely associated with the prognosis of patients with LUAD. Previous study indicated that AC004687.1 was related with tumor progression and served as a prognostic marker to predict the prognosis and tumor immunity of patients with LUAD [45]. Xu et al. [46] found that AC074011.1 was related with tumor immunity and constructed a immune-related lncRNAs prognostic risk signature for predicting the prognosis of patients with colon cancer. These findings suggest that the identified 9 LRHGs may be promising prognostic and diagnostic markers for LUAD, providing novel targets for the research and development commercial reagent kits.

The enrichment analyses of the 9 LRHGs revealed that integrin 4, advanced glycosylation end product receptor,

Table 2 Association between HOXD-AS2 level and clinicopathological characteristics of patients with lung adenocarcinoma

Features	n	HOXD-AS2 (mean ± SEM)	*t*	*P*-value
Sex			1.368	0.176
Male	37	2.47 ± 0.41		
Female	34	3.32 ± 0.48		
Age (years)			0.050	0.960
> 65	35	2.89 ± 0.44		
≤ 65	36	2.86 ± 0.44		
Smoking history			1.057	0.294
Never	33	3.23 ± 0.46		
Ever	38	2.57 ± 0.43		
Tumor size (cm)			0.320	0.750
> 3	26	3.00 ± 0.59		
≤ 3	45	2.80 ± 0.36		
Tumor differentiation			3.368	0.001
Well/Moderate	52	2.28 ± 0.22		
Poor	19	4.50 ± 0.92		
TNM stages			8.942	< 0.001
I + II	50	1.58 ± 0.11		
III + IV	21	5.72 ± 0.67		
Lymphovascular invasion			5.523	< 0.001
Present	15	2.12 ± 0.21		
Absent	56	5.67 ± 0.96		

mitochondrial translation and fatty acid metabolism signaling pathways were identified. Past researches had revealed that LRHGs levels were tightly associated with these tumourigenesis and immune-related pathways [47, 48]. Among them, the integrin pathway has a notable role in a number of biological functions, such as metastasis, tumor growth and immune escape [49, 50]. Wei et al. [51] found that lncRNA PVT1 expression was positively correlated with integrin β-8, and that knockdown of PVT1 suppressed NSCLC progression by regulating the miR-145-5p/integrin β-8 axis and inhibiting the MEK/ERK signaling pathway, implying the mechanism of PVT1 acts as a ceRNA network in NSCLC progression. In addition, mitochondrial translation and fatty acid metabolism are essential to the progression of LUAD. Numerous studies had demonstrated that high rates of oxidative phosphorylation and fatty acid oxidation exist in tumors [52–54]. Therefore, the aforementioned results indicated that the identified 9 LRHGs promoted the progression of LUAD by influencing the aforementioned 4 signaling pathways.

In the present study, patients with LUAD were assigned to a high or low-risk group based on the median risk score. The constructed nomogram implied that the LRHGs had a particularly marked predictive efficacy. Additionally, the difference of the 9 LRHGs levels was analyzed and verified in two LUAD clusters and our collected samples. It was observed that the LRHGs levels were differentially expressed in the two LUAD clusters. Further tumor microenvironment analysis indicated that the stromal, immune and estimate scores were higher in cluster 1 than in cluster 2, suggesting that antitumor immunotherapy may have a more optimal therapeutic effect in patients in cluster 1. Moreover, the results of the ssGSEA algorithm also implied that the infiltration abundance of 23 immune cells was positively associated with other immune cells. Studies have previously revealed the critical roles of the tumor microenvironment in modulating immune escape and tumor malignancies [55, 56]. Immune cells within the tumor microenvironment can promote an immune response through interacting with other immune cells, contributing to an extensive impact

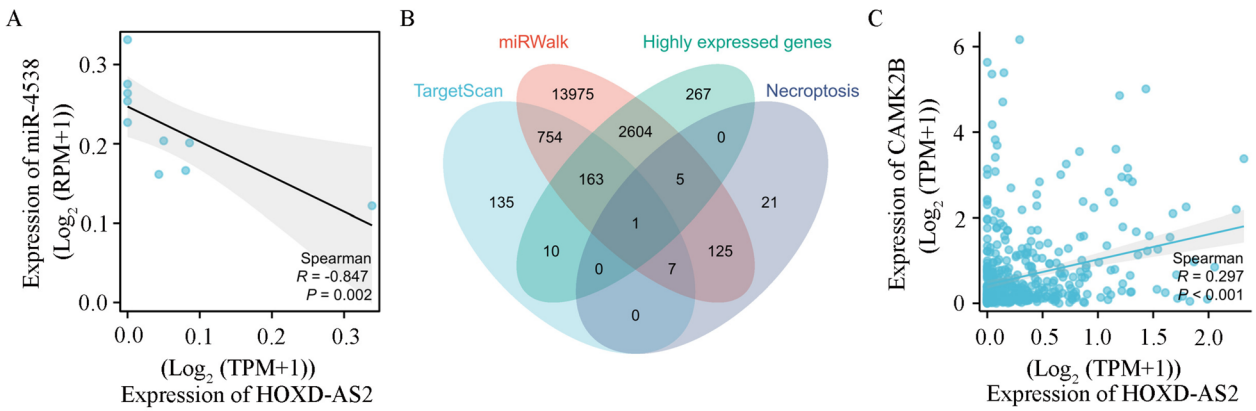

Fig. 12 CeRNA analysis of HOXD-AS2 in LUAD. **A** Pearson correlation analysis was used to investigate the association between miR-4538 and HOXD-AS2. **B** A Wayne diagram conducted in conjunction with the TargetScan database, the miRWalk database, the HEmRNAs and the necroptosis-related gene list. **C** Pearson correlation analysis was used to investigate the association between CAMK2B and HOXD-AS2

on tumor progression and metastasis [57]. Liu et al. [58] found that the six disulfidptosis-related lncRNA signatures were tightly associated with immune escape by influencing immune cell proportion, including T follicular helper cells and M0 macrophages. Overall, these findings implied that the constructed LRHGs model has provided a potential guidance for the administration of immune checkpoint inhibitors.

Furthermore, to explore the potential mechanism of HOXD-AS2 via a ceRNA network in LUAD [38], it was indicated that a potential regulatory network involving miR-4538 and CAMK2B was identified. The negative correlation between HOXD-AS2 and miR-4538 suggested that HOXD-AS2 was used as a molecular sponge through sequestering miR-4538 and thereby modulating the target gene expressions, underscoring the complex regulatory mechanisms between non-coding RNAs and their target gene expressions in cancer [59]. CAMK2B, identification as a necroptosis-related gene, was positively correlated with HOXD-AS2 and played a critical role in calcium signaling pathways in various cancers [60, 61]. These data suggest that the HOXD-AS2/miR-4538/CAMK2B axis presents a promising avenue for further research and potential therapeutic intervention in LUAD.

Taken together, the discovery of the 9 LRHGs has contributed to the investigation of a more resultful and personalized therapeutic strategy for patients with LUAD. There is no doubt that this research has several limitations. For instance, findings were based on the TCGA database without large clinical samples validation. Additionally, the TCGA-LUAD samples were mainly from North America and therefore potential biases existed in the patient selection. More diverse and multi-regional samples will be collected in the future. Although we conducted some experimental validations, there is a lack of more in-depth experimental validations. Further gain or loss of function experiments should be conducted to explore the correlations between LRHGs levels and associated signaling pathway as well as immune cell infiltration, especially focusing on HOXD-AS2. Combined with single-cell RNA sequencing, we will provide more intuitive cell-specific data, contributing to understand the role of HOXD-AS2 in tumor evolution and immune escape. Lastly, the role and particular mechanism of HOXD-AS2/miR-4538/CAMK2B axis in LUAD progression remain to be further investigated.

In summary, the present study constructed a LRHGs model using integrated bioinformatics analyses to serve as an independent prognostic and diagnostic predictor for patients with LUAD. Underscoring the HOXD-AS2/miR-4538/CAMK2B axis as the potential mechanism in LUAD progression, offering a new perspective for prognosis monitoring and targeted therapy in LUAD.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-024-13144-2.

Supplementary Material 1.

Acknowledgements
Not applicable.

Authors' contributions
Data curation, Yan Jiang and Yanqin Hu; Formal analysis, Yong Li and Qiuxia Liu; Funding acquisition, Yong Li, Yanping Xie and Xilin Zhang; Investigation, Danfei Shi, Yan Jiang, Yanqin Hu and Qiuxia Liu; Methodology, Danfei Shi; Project administration, Yanping Xie and Xilin Zhang; Writing – original draft, Yong Li; Writing – review & editing, Yanping Xie and Xilin Zhang.

Funding
This research was kindly supported by Zhejiang Provincial Natural Science Foundation of China under Grant Nos. LY22H160026 and LTGY23H160004, The Scientific Technology Projects of Health and Medicine of Zhejiang Province under Grant nos. 2023KY1180 and 2023KY318, and Huzhou Science and Technology Fund under Grant no. 2021GZB03.

Data availability
Data is provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate
This study was approved by the Medical Research and Clinical Trials Ethics Committee of the First Affiliated Hospital of Huzhou University in accordance with the Declaration of Helsinki (approved number: 2023KYLL014). Informed consent was obtained from all patients.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Laboratory Medicine, First Affiliated Hospital of Huzhou University, Huzhou 313000, Zhejiang, China. [2]Department of Pathology, First Affiliated Hospital of Huzhou University, Huzhou 313000, Zhejiang, China. [3]Department of Respiratory Medicine, First Affiliated Hospital of Huzhou University, No. 158 Guangchang Back Road, Huzhou, Zhejiang 313000, P. R. China. [4]Central Laboratory, Huzhou Key Laboratory of Translational Medicine, First Affiliated Hospital of Huzhou University, No. 158 Guangchang Back Road, Huzhou, Zhejiang 313000, P. R. China.

Received: 3 August 2024 Accepted: 4 November 2024
Published online: 09 November 2024

References

1. Siegel RL, Miller KD, Wagle NS and Jemal A: Cancer statistics, 2023. CA: a cancer journal for clinicians 73: 17–48, 2023.
2. Sung H, Ferlay J, Siegel RL, et al.: Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: a cancer journal for clinicians 71: 209–249, 2021.

3. Yuan J, Sun Y, Wang K, et al. Development and validation of reassigned CEA, CYFRA21-1 and NSE-based models for lung cancer diagnosis and prognosis prediction. BMC Cancer. 2022;22:686.
4. Zhou X, Li Y, Wu L, Tian C, Wu X. Upregulated lncRNA LINC01128 in colorectal cancer accelerates cell growth and predicts malignant prognosis through sponging miR-363-3p. J Cancer Res Clin Oncol. 2024;150:276.
5. Wang M, Niu X, Wang M, et al.: Long non-coding RNA RP11–197K6.1 as ceRNA promotes colorectal cancer progression via miR-135a-5p/DLX5 axis. J Transl Med 22: 469, 2024.
6. He XY, Fan X, Qu L, et al. LncRNA modulates Hippo-YAP signaling to reprogram iron metabolism. Nat Commun. 2023;14:2253.
7. Zhao C, Xu H and Liu C: Identification of Novel Prognostic Long Noncoding RNAs in Lung Adenocarcinoma Using WGCNA Analysis. Biochemical genetics, 2023.
8. Wang P, Wang Z, Lin Y, et al.: Development of a Novel Pyroptosis-Associated lncRNA Biomarker Signature in Lung Adenocarcinoma. Mol Biotechnol, 2023.
9. Xu Y, Lin G, Liu Y, et al. An integrated analysis of the competing endogenous RNA network associated of prognosis of stage I lung adenocarcinoma. BMC Cancer. 2022;22:188.
10. Li S, Xie Y, Zhou W, et al. Association of long noncoding RNA MALAT1 with the radiosensitivity of lung adenocarcinoma cells via the miR-140/PD-L1 axis. Heliyon. 2023;9: e16868.
11. Li ZX, Zhu QN, Zhang HB, Hu Y, Wang G, Zhu YS. MALAT1: a potential biomarker in cancer. Cancer Manag Res. 2018;10:6757–68.
12. Fan C, Wang Q, Kuipers TB, et al. LncRNA LITATS1 suppresses TGF-beta-induced EMT and cancer cell plasticity by potentiating TbetaRI degradation. EMBO J. 2023;42: e112806.
13. Li Y, Huang S, Wei Z, Yang B. A putative competing endogenous RNA network in cisplatin-resistant lung adenocarcinoma cells identifying potentially rewarding research targets. Oncol Lett. 2020;19:4040–52.
14. Wang XJ, Gao J, Wang Z, Yu Q. Identification of a Potentially Functional microRNA-mRNA Regulatory Network in Lung Adenocarcinoma Using a Bioinformatics Analysis. Front Cell Dev Biol. 2021;9: 641840.
15. Colaprico A, Silva TC, Olsen C, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44: e71.
16. Goldman MJ, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38:675–8.
17. Tien FM, Yao CY, Tsai XC, et al. Dysregulated immune and metabolic pathways are associated with poor survival in adult acute myeloid leukemia with CEBPA bZIP in-frame mutations. Blood Cancer J. 2024;14:15.
18. Cheng F, Li Q, Wang J, Zeng F, Wang K, Zhang Y. Identification of Differential Intestinal Mucosa Transcriptomic Biomarkers for Ulcerative Colitis by Bioinformatics Analysis. Dis Markers. 2020;2020:8876565.
19. Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin. Epigenetics. 2019;11:123.
20. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2010;5:1315–6.
21. Park SY. Nomogram: An analogue tool to deliver digital knowledge. J Thorac Cardiovasc Surg. 2018;155:1793.
22. Perkins NJ, Weck J, Mumford SL, et al. Combining Biomarker Calibration Data to Reduce Measurement Error. Epidemiology. 2019;30(Suppl 2):S3–9.
23. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. Eur Urol. 2018;74:796–804.
24. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50.
25. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P and Mesirov JP: Molecular signatures database (MSigDB) 3.0. Bioinformatics 27: 1739–1740, 2011.
26. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:7.
27. Wu D, Huo C, Jiang S, et al. Exostosin1 as a novel prognostic and predictive biomarker for squamous cell lung carcinoma: A study based on bioinformatics analysis. Cancer Med. 2021;10:2787–801.
28. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.
29. Briere G, Darbo E, Thebault P, Uricaru R. Consensus clustering applied to multi-omics disease subtyping. BMC Bioinformatics. 2021;22:361.
30. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26:1572–3.
31. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013;39:782–95.
32. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–7.
33. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25:402–8.
34. McGeary SE, Lin KS, Shi CY, et al.: The biochemical basis of microRNA targeting efficacy. Science 366, 2019.
35. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. PLoS ONE. 2018;13: e0206239.
36. Liu F, Wei T, Liu L, et al. Role of Necroptosis and Immune Infiltration in Human Stanford Type A Aortic Dissection: Novel Insights from Bioinformatics Analyses. Oxid Med Cell Longev. 2022;2022:6184802.
37. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med. 1997;16:385–95.
38. Kamali MJ, Salehi M, Mostafavi M, et al. Hijacking and rewiring of host CircRNA/miRNA/mRNA competitive endogenous RNA (ceRNA) regulatory networks by oncoviruses during development of viral cancers. Rev Med Virol. 2024;34: e2530.
39. Ma C, Gu Z, Yang Y. Development of m6A/m5C/m1A regulated lncRNA signature for prognostic prediction, personalized immune intervention and drug selection in LUAD. J Cell Mol Med. 2024;28: e18282.
40. Chen W, Liao C, Xiang X, et al. A novel tumor mutation-related long non-coding RNA signature for predicting overall survival and immunotherapy response in lung adenocarcinoma. Heliyon. 2024;10: e28670.
41. Gao M, Wang M, Chen Y, et al. Identification and validation of tryptophan metabolism-related lncRNAs in lung adenocarcinoma prognosis and immune response. J Cancer Res Clin Oncol. 2024;150:171.
42. Ren Y, Wu R, Li C, et al. Single-cell RNA sequencing integrated with bulk RNA sequencing analysis identifies a tumor immune microenvironment-related lncRNA signature in lung adenocarcinoma. BMC Biol. 2024;22:69.
43. Jiang Y, Ye Y, Huang Y, et al. Identification and validation of a novel anoikis-related long non-coding RNA signature for pancreatic adenocarcinoma to predict the prognosis and immune response. J Cancer Res Clin Oncol. 2023;149:15069–83.
44. Zhao Q, Ye Y, Zhang Q, et al. PANoptosis-related long non-coding RNA signature to predict the prognosis and immune landscapes of pancreatic adenocarcinoma. Biochem Biophys Rep. 2024;37: 101600.
45. Chen H, Zhou C, Hu Z, et al. Construction of an algorithm based on oncosis-related LncRNAs comprising the molecular subtypes and a risk assessment model in lung adenocarcinoma. J Clin Lab Anal. 2022;36: e24461.
46. Xu M, Li Q, Zhang J, Xie H. Identification of Immune-Related lncRNA Pairs and Construction and Validation of a New Prognostic Signature of Colon Cancer. Can J Gastroenterol Hepatol. 2022;2022:5827544.
47. Wu G, Wang Q, Zhu T, et al. Identification and Validation of Immune-Related LncRNA Prognostic Signature for Lung Adenocarcinoma. Front Genet. 2021;12: 681277.
48. Li H, Zhang H, Huang G, et al. Loss of RPS27a expression regulates the cell cycle, apoptosis, and proliferation via the RPL11-MDM2-p53 pathway in lung adenocarcinoma cells. J Exp Clin Cancer Res. 2022;41:33.
49. Di X, Gao X, Peng L, et al. Cellular mechanotransduction in health and diseases: from molecular mechanism to therapeutic targets. Signal Transduct Target Ther. 2023;8:282.
50. He C, Xiu W, Chen Q, et al.: Gut-licensed beta7(+) CD4(+) T cells contribute to progressive retinal ganglion cell damage in glaucoma. Science translational medicine 15: eadg1656, 2023.
51. Wei CM, Zhao XF, Qiu HB, Ming Z, Liu K, Yan J. The long non-coding RNA PVT1/miR-145-5p/ITGB8 axis regulates cell proliferation, apoptosis, migration and invasion in non-small cell lung cancer cells. Neoplasma. 2020;67:802–12.
52. Han M, Bushong EA, Segawa M, et al. Spatial mapping of mitochondrial networks and bioenergetics in lung cancer. Nature. 2023;615:712–9.

53. Feng J, Lian Z, Xia X, et al. Targeting metabolic vulnerability in mitochondria conquers MEK inhibitor resistance in KRAS-mutant lung cancer. Acta Pharm Sin B. 2023;13:1145–63.
54. Chen Y, Zhao Y, Deng Y, Yang Y, Xu L, Fu J. FATP2 regulates non-small cell lung cancer by mediating lipid metabolism through ACSL1. Tissue Cell. 2023;82: 102105.
55. Ye W, Li M and Luo K: Therapies Targeting Immune Cells in Tumor Microenvironment for Non-Small Cell Lung Cancer. Pharmaceutics 15, 2023.
56. Spaas M, Sundahl N, Kruse V, et al.: Checkpoint Inhibitors in Combination With Stereotactic Body Radiotherapy in Patients With Advanced Solid Tumors: The CHEERS Phase 2 Randomized Clinical Trial. JAMA oncology, 2023.
57. Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. Cancer Res. 2019;79:4557–66.
58. Liu S, Wang S, Guo J, et al. Crosstalk among disulfidptosis-related lncRNAs in lung adenocarcinoma reveals a correlation with immune profile and clinical prognosis. Noncoding RNA Res. 2024;9:772–81.
59. Ding D, Zhang J, Luo Z, et al. Analysis of the lncRNA-miRNA-mRNA Network Reveals a Potential Regulatory Mechanism of EGFR-TKI Resistance in NSCLC. Front Genet. 2022;13: 851391.
60. Zhang J, Liu Z, Chen W, Liu H. Identification and validation of a necroptosis-related gene prognostic signature for colon adenocarcinoma. Transl Cancer Res. 2023;12:2239–55.
61. Bhowmick C, Rahaman M, Bhattacharya S, et al. Identification of hub genes to determine drug-disease correlation in breast carcinomas. Med Oncol. 2023;41:36.

Publisher's Note