Construction of an m6A-related lncRNA model for predicting prognosis and immunotherapy in patients with lung adenocarcinoma

Jin-hai Tang (jhtang@njmu.edu.cn)
Nanjing Medical University https://orcid.org/0000-0003-2016-4216
Hong-yu Shen
Nanjing Medical University
Jin Zhang
Nanjing Medical University
Di Xu
Nanjing Medical University
Zheng Xu
Nanjing Medical University
Mingxing Liang
Nanjing Medical University
Wen-quan Chen
Nanjing Medical University
Wen-jia Xia
Nanjing Medical University

Research Article

Keywords: m6A , lncRNA , LUAD , prognosis , immunotherapy

Posted Date: February 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1345012/v1

Construction of an m6A-related lncRNA model for predicting prognosis and immunotherapy in patients with lung adenocarcinoma

Hong-yu Shen[1,2*], Jin Zhang[3*], Di Xu[1], Zheng Xu[1], Ming-xing Liang[1], Wen-quan Chen[1], Wen-jia Xia[4#], Jin-hai Tang[1,2#]

[1]Department of General Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China
[2]Department of Oncology, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou, China
[3]Department of General Practice, Geriatric Hospital of Nanjing Medical University, Jiangsu Province Geriatric Hospital, Nanjing, Jiangsu, China
[4]Department of Thoracic Surgery, Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, China

[*]Hong-yu Shen, and Jin Zhang contributed equally to this work.

Correspondence: Jin-hai Tang, jhtang@njmu.edu.cn; Wen-jia Xia, xwj09010810@163.com

Abstract

Purpose: This study aimed to explore the role of N6-methyladenosine (m6A)-related lncRNAs in lung adenocarcinoma (LUAD).

Methods: Gene expression data and clinical data of LUAD patients were acquired from The Cancer Genome Atlas (TCGA) Database. Combined with clinical information, the prognostic m6A-related lncRNAs were identified through differential lncRNA expression analysis and Spearman correlation analysis. Next, the least absolute shrinkage and selection operator (LASSO) regression was used to establish the prognostic risk model. We evaluated and validated the predictive performance of this model via survival analysis and receiver operating characteristic (ROC) curve analysis. The expression of immune checkpoints, immune cell infiltration and drug sensitivity of patients in different risk groups were analyzed separately.

Results: A total of 19 prognostic m6A-related lncRNAs were identified and then the prognostic risk model was well established. The patients were divided into high- and low-risk group based on the median value of the risk scores. Compared with the patients in the low-risk group, the prognosis of the patients in the high-risk group was relatively poor. The ROC curves showed that this model had excellent sensitivity and specificity. Multivariate Cox regression analysis indicated that the risk score could be used as an independent prognostic risk factor. The expression levels of immune checkpoint *CD276, PVR*, and *VTCN1* were significantly increased in the high-risk group. Finally, we found that the risk scores were correlated with immune cell infiltration and drug sensitivity.

Conclusion: We constructed a prognostic risk model in LUAD patients based on

m6A-related lncRNAs. This model was also associated with the expression of immune checkpoints, immune cell infiltration and drug sensitivity, which will provide new insights into immunotherapy for LUAD patients in the future.
Keywords: m6A · lncRNA · LUAD · prognosis · immunotherapy

Background

Lung adenocarcinoma (LUAD) accounts for 40% of all lung cancer patients, becoming the main subtype of non-small cell lung cancer (NSCLC) (Denisenko et al. 2018). LUAD occurs in people of all ages, and is not correlated with gender or smoking (Couraud et al. 2012). Although most LUAD patients have received systematic clinical techniques, such as surgery, chemotherapy, radiotherapy and targeted therapy, their 5-year survival rate is less than 15% (Ettinger et al. 2015). Unfortunately, there are still a large number of patients who were diagnosed at advanced stages, thus losing the opportunity for surgery, having limited treatment options and causing low survival rates (Rodriguez-Canales et al. 2016). Therefore, uncovering new diagnostic and prognostic indicators is urgently needed for the early detection and treatment of LUAD.

N6-Methyladenosine (m6A), first reported in the 1970s (Desrosiers et al. 1974), is the most abundant modification in eukaryotic messenger RNAs (mRNAs), microRNAs (miRNAs) and long non-coding RNAs (lncRNAs), playing an indispensable role in RNA splicing, stability, export and translation (Dai et al. 2018). The m6A methylation is a dynamic and reversible process, which is intimately mediated by regulators, including “writers” (m6Amethylases), “erasers” (m6A demethylases), and “readers” (signaltransducers) (Zaccara et al. 2019). Abnormal m6A methylation occurs frequently in both normal biological processes and tumorigenesis (Chen et al. 2018b). Recent studies have shown that m6A modification regulates lung tumorigenesis. For example, downregulation of the m6A reader *YT521*-Bhomology domain containing 2 (*YTHDC2*) in LUAD is associated with poor clinical outcome. *YTHDC2* decreases tumorigenesis both in LUAD mouse models and in human LUAD cells (Ma et al. 2021). Additionally, it is identified that the m6A writer methyltransferase-like 3 (*METTL3*) promotes *YAP* translation and increases *YAP* activity through the*MALAT1-miR-1914-3p-YAP* axis to induce drug resistance and metastasis in NSCLC (Jin et al. 2019). Moreover, studies have also revealed that m6A modifications are involved in cell infiltration, tumor microenvironment, immuneactivity, and response to immune checkpoint inhibitors (Zhou et al. 2021). However, the relationship of m6A modification to cancer progression and immune response remains elusive in LUAD.

LncRNAs have received great attention in elucidating the complex mechanisms in malignant tumors such as occurrence, recurrence, metastasis and drug resistance (Statello et al. 2021). Increasing studies proposed a new regulatory mechanism in which lncRNAs may act as competing endogenous RNAs (ceRNAs) and engage in crosstalk with mRNAs by using miRNA response elements (MREs) to competitively sponge their miRNAs (Lei et al. 2019). For instance, lncRNA *DGCR5* promotes LUAD progression via inhibition of hsa-mir-22-3p (Dong et al. 2018). Likewise,

FOXP4-AS1 post-transcriptionally regulated *FOXP4* by sponging to miR-3184-5p in prostate cancer (Wu et al. 2019). Meanwhile, other studies have also suggested a crucial role of m6A modifications in the dysregulation of lncRNAs during the development and progression of tumors (Dai et al. 2020). Nevertheless, the specific role and the mechanisms underlying m6Amodification in lncRNAs in LUAD have not been clarified.

The emergence of immune checkpoint inhibitors (ICIs) has driven a revolutionary shift in traditional cancer treatment. Resent preclinical and clinical trials have supported that ICIs are also promising approaches for the treatment of LUAD (Marinelli et al. 2020). For example, in the global KEYNOTE-042 study, compared with chemotherapy, pembrolizumab significantly improved overall survival (OS) in patients with previously untreated programmed death ligand 1 (*PD-L1*) -positive locally advanced/ metastatic NSCLC without *EGFR/ALK* alterations (Wu et al. 2021). Depending on the results of the KEYNOTE-189 study, pembrolizumab combined with pemetrexed and platinumis currently NMPA-approved for previously untreated metastatic nonsquamous NSCLC (Gadgeel et al. 2020). In the present study, we aimed to describe the landscape of m6A-related lncRNAs in the TCGA LUAD cohort. Furthermore, we constructed an m6A-related lncRNA prognostic model via the least absolute shrinkage and selection operator (LASSO) regression. The expression of immune checkpoints, immune cell infiltration and drug sensitivity were also systematically analyzed in the high-risk and low-risk subgroups.

Materials and methods

LUAD dataset acquisition

RNA sequencing data of 526 LUAD samples and 59 normal paired tissues were acquired from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). Clinical data (including age, sex, differentiation grade and TNM stage) of patients were directly downloaded from TCGA. The lncRNAs in the TCGA dataset were identified based on the annotation of Genome Reference Consortium Human Build 38 (GRCh38), and a total of 5,606 lncRNAs were obtained in the TCGA-LUAD transcriptome matrix for further analysis.

Differential lncRNA expression analysis

We employed the differential lncRNA expression analysis based on EdgeR (Robinson et al. 2010) package to retrieve differentially expressed lncRNAs (DElncRNAs) from TCGA-LUAD. DElncRNAs were considered as significantly different if they meet the conditions of fold change > 1.5 and $p<0.05$.

Extraction of m6A-related lncRNAs

According to previous publications (Han et al. 2021; Wang et al. 2021), the expression levels of 23 m6A-related genes were extracted from the TCGA-LUAD, including expression data on writers (*CBLL1, METTL3, METTL14, WTAP, VIRMA, RBM15, RBM15B*, and *ZC3H13*), readers (*ELAVL1, FMR1, YTHDC1, YTHDC2,YTHDF1, YTHDF2, YTHDF3, HNRNPC, HNRNPA2B1, IGF2BP1, IGF2BP2, IGF2BP3*, and *LRPPRC*), and erasers (*FTO* and *ALKBH5*). The Spearman correlation analysis was applied to screen m6A-related lncRNAs with the criterion of $p<0.05$ in the TCGA-LUAD dataset. Combined with the DElncRNAs selected above,

totally 742 differentially expressed m6A-related lncRNAs (DEm6A-lncRNAs) were identified. The network of m6A-related genes and m6A-related lncRNAs was visualized by Cytoscape software 3.6.0 (Shannon et al. 2003). Based on LUAD clinical data, we conducted receiver operating characteristic (ROC) curves with "survivalROC" package (Heagerty and Zheng 2005) and calculated the area under the curve (AUC) value of each lncRNA at 7 years. Finally, a total of 26 DEm6A-lncRNAs with AUC>0.6 were selected for subsequent analysis.

Construction and validation of the m6A-related lncRNAs prognostic risk model

First, the TCGA LUAD patients with complete survival information were randomly divided into training cohort (n=251) and testing cohort (n=251) using the R caret package (Hengl et al. 2017), with each cohort consisting of 50% of cases. The training cohort was utilized to construct an m6A-related lncRNA prognostic model, and the entire cohort (n=502) and the testing cohort were applied to validate this established model. Subsequently, totally 26 DEm6A-lncRNAs selected above and LUAD clinical data were further analyzed through the LASSO regression analysis using the minimum lambda value with the R glmnet package (Friedman et al. 2010). Finally, the regression coefficients corresponding to each gene were calculated, and the marked 19 lncRNAs were determined according to the criterion of coefficient $\neq 0$. We calculated the risk score of each sample with the formula: risk score = $\sum_{i=1}^{n}$(coefi $\times$ βi), where coef represents the regression coefficient, and β represents the m6A-related lncRNA expression value. Then, the patients were divided into high- and low-risk group using the median value of the risk scores. Univariate Cox regression analysis was performed to sort the m6A-related lncRNAs with significant prognostic value ($p < 0.05$). The R survival package (Heagerty and Zheng 2005) was used for Kaplan–Meier analysis based on the risk score and clinical characteristics (T stage, N stage, M stage, and TNM stage). The log-rank test was used to compare the overall survival rates between the high- and low-risk group. ROC curve was generated to analyze the 7 year survival rate of patients and assess the accuracy of survival prediction of the gene signature. Univariate and multivariate Cox regression analysis were used to evaluate whether the risk score is an independent prognostic factor.

Correlation analysis between m6A-related lncRNAs and immune cell infiltration

LncRNAs are involved in the regulation of the tumor microenvironment (TME) and signaling transduction in tumor cells (Li et al. 2019). To better understand the correlation between m6A-related lncRNAs and TME, the fraction of 22 immune cell types for each sample was scored through cell type identification by estimating relative subsets of RNA transcripts. We used the CIBERSORT (Chen et al. 2018a) algorithm to evaluate the immune cell infiltration in different LUAD clusters. A comparison of the fraction of 22 immune cells in the high- and low-risk group was produced with "limma" (Ritchie et al. 2015) packages in R.

Correlation analysis between m6A-related lncRNAs and immune checkpoints

To explore the correlation between m6A-related lncRNAs and immune checkpoints, we selected 21 critical immune checkpoints (*HAVCR2, CD274, CD86, LAG3, LAIR1, PVR, IDO1, CD80, CTLA4, PDCD1, TIGIT, CD200R1, CEACAM1, CD276, CD200, KIR3DL1, BTLA, ADORA2A, LGALS3, VTCN1* and *CLEC4G)*, which are associated

with currently used tumor immune checkpoint inhibitors (Hu et al. 2020).The expression of checkpoint members were estimated in the high- and low-risk group using the "limma" (Ritchie et al. 2015) package and Wilcoxon test.

Correlation analysis between m6A-related lncRNAs and therapeutic drugs

We predicted the therapeutic response for each sample based on the largest publicly available pharmacogenomics database [the Genomics of Drug Sensitivity in Cancer(GDSC), https://www.cancerrxgene.org/] (Jiang et al. 2021). The half-maximal inhibitory concentration (IC50) of therapeutic drugs was estimated using the R package pRRophetic (Geeleher et al. 2014), and the differences of IC50 values between the high- and low- risk group were then compared.

Tissue samples and real-time quantitative PCR (RT-qPCR) Assay

A total of 20 tumor tissues and adjacent normal tissues were collected from LUAD patients who underwent surgery at Geriatric Hospital of Nanjing Medical University. According to the TNM Staging System for LUAD and histopathological evaluation, all patients were comfirmed. No neoadjuvant therapy was performed on these patients before surgery. Prior to RT-qPCR analysis, all tissues were stored at −80°C. All human tissue samples were obtained with written informed consent from all subjects, and this project was approved by the Research Ethics Committee of Geriatric Hospital of Nanjing Medical University.

Total RNA was isolated using TRIzol reagent (Invitrogen) from tissues. Total RNA (1 μg) was reverse transcribed by using BuSuperScript RT Kit (Biouniquer Technology, Nanjing, China) following the manufacturer's instruction. For detection of lncRNA expression levels, RT-qPCR was then performed on LightCycler®480 (Roche, Switzerland) with SYBR Green PCR Master Mix (Roche, Australia). Before calculation using the ΔΔCt method, the levels of GAPDH were used to normalize the relative expression levels of lncRNA.

Statistical Analysis

All analyses were performed on R software 3.6.3. The related R codes and original data were uploaded to the GitHub page (https://github.com/njmushy/LUAD-lncRNA). The significance of the two groups of samples passed the Wilcoxon test, and the significance of the three groups and above passed the Kruskal-Wallis test. Kaplan-Meier curves were plotted to compare the OS among various subgroups and the logrank $p< 0.05$ was recognized as statistically significant. Univariate Cox regression analysis and multivariate Cox regression analysis were conducted to evaluate the predictive performance of the risk model based on m6A-related lncRNAs. $p < 0.05$ was considered statistically significant.

Results

Screening for significant DElncRNAs in LUAD

We downloaded the RNA-seq data of 585 samples (tumor samples, 526; normal samples, 59) from TCGA, and obtained 5,606 lncRNAs according to the annotation of GRCh38 from GENCODE. The study flowchart was shown in Figure 1. The expression of totally 5,606 lncRNAs was displayed in Figure 2A and Figure 2B. Differential lncRNA expression analysis was performed by EdgeR package, with the

criteria of fold change > 1.5 and p < 0.05. As shown in Figure 2C, a total of 1,770 up-regulated significant DElncRNAs and 376 down-regulated significant DElncRNAs in LUAD were confirmed. Figure 2D showed the expression of the top 20 up-regulated and down-regulated DElncRNAs through heatmap. Subsequently, these significant DElncRNAs were subject to further analysis.

Identification of m6A-related lncRNAs associated with LUAD

A total of 5,606 lncRNAs and expression matrixes of 23 m6A-related genes were extracted from TCGA-LUAD RNA sequencing dataset. Then Spearman correlation analysis was conducted for preliminarily selecting m6A-related lncRNAs. With a threshold of $p<0.05$, totally 2,322 lncRNAs were significantly correlated with m6A-related genes, called m6A-lncRNAs. We further searched some significant DEm6A-lncRNAs which were simultaneously appeared in the DElncRNAs and m6A-lncRNAs datasets. Figure 2E showed that 742 DEm6A-lncRNAs were observed by Venn diagrams, selecting for in-depth analysis. Clinical data of LUAD patients were also downloaded from TCGA. To further narrow the DEm6A-lncRNAs, we conducted ROC curves and calculated the AUC value of each lncRNA at 7 years. Finally, a total of 26 DEm6A-lncRNAs with the criterion of AUC>0.6 was selected for subsequent analysis (Table S1). The co-expression network between DEm6A-lncRNAs and m6A-related genes was visualized by Cytoscape in Figure 2F.

Construction and validation of the prognostic risk model for patients with LUAD based on m6A-related lncRNAs

The LUAD patients in the TCGA database were randomly divided into a training cohort (n=251) and a testing cohort (n=251). There was no significant difference in the clinical baseline characteristics between the training cohort and the testing cohort ($p>0.05$; Table S2). To identify the most powerful prognostic m6A-related lncRNAs, 26 DEm6A-lncRNAs and coefficient of each lncRNA were finally identified through LASSO regression analysis using the minimum lambda criterion based on the training cohort (Figure 3A, B). Eliminating lncRNAs with regression coefficient = 0, 19 DEm6A-lncRNAs were obtained: *ACOXL-AS1, AF131215.2, EXOSC10-AS1, HCG18, LARGE-AS1, LINC01011, LINC01224, LINC01266, LINC02147, LINC02321, LINC02802, LMO7DN-IT1, MIR133A1HG, NKILA, OGFRP1, PPP1R14B-AS1, STRA6LP, ZNF252P-AS1* and *ZNF516-DT* (Table S3). These lncRNAs were used to construct a prognostic risk model according to the formula mentioned in the Materials and methods. Among them, 3 up-regulated lncRNAs and 1 down-regulated lncRNA were verified with the most significantly differential expression, and then proceeded to RT-qPCR experimental verification. Compared with the paired adjacent normal tissues, the expression of PPP1R14B-AS1, LINC01224 and LINC02321 in LUAD tissue samples were increased, on the contrary, the expression of AF131215.2 was decreased (Figure S1). These results were consistent with their expression in TCGA, and also in agreement with some reported literatures (Yang et al. 2020; Xiao et al. 2021). The distributions of risk score and survival status of all LUAD patients were shown in Figure 3C and 3D. The figures indicate that as the risk score increased, mortality increased and survival time decreased. As assessed through the univariate Cox regression analysis, 10 of the 19 DEm6A-lncRNAs were found to be

significantly associated with prognosis (Figure 3E).

Based on the median value of the risk scores, patients with LUAD were divided into high- and low-risk group for further analysis. Next, we performed Kaplan–Meier analysis to evaluate the prognostic value of the prognostic model. In the training cohort, Figure 4A showed that the overall survival of patients in the high-risk group was lower than that of patients in the low-risk group ($p<0.001$). Similar results were displayed in the testing cohort and the entire cohort ($p<0.001$, Figure 4B and 4C). The predictive performance of the risk model was evaluated with ROC curves. The AUCs at 7 years for OS were 0.73 in the training cohort (Figure 4D), 0.713 in the testing cohort (Figure 4E), and 0.735 in the entire cohort (Figure 4F), respectively. These results suggest that the prognostic model has excellent sensitivity and specificity. In summary, the risk scores based on 19 m6A-related lncRNAs have the best predictive power for the prognosis of LUAD patients.

Prognostic risk score correlates with clinicopathological characteristics

We assessed the association between the risk model and clinical factors such as age, sex, T stage, N stage, M stage and pathological stage by univariate and multivariate Cox regression analyses. Univariate Cox regression analysis indicated that T stage, N stage, pathological stage, and risk score were correlated with the prognosis of LUAD patients (Figure 5A). Multivariate Cox regression analysis suggested that the pathological stage and risk score could be used as independent prognostic risk factors (Figure 5B).

Next, we performed Wilcoxon test to evaluate the differences between the risk scores and other clinicopathological characteristics of the prognostic model. Risk scores for patients in T2 vs T1, T3 vs T1, N1 vs N0, N2 vs N0, Stage II vs I, Stage III vs I, and Stage IV vs II were statistically significant ($p<0.05$, Figure 5C–F). The worse the clinical pathological characteristic was, the higher the risk score was. However, there was no statistical significance among the remaining T/N/M/Stage groups. Furthermore, we assessed ROC curves of the risk score and different clinicopathological characteristics. The AUC of the risk score was also higher than the AUCs of clinicopathological characteristics (age, sex, T, N, M and pathological features) (Figure 5G), indicating that the risk score could better predict the occurrence and development of LUAD and this prognostic model of the 19 m6A-related lncRNAs may be relatively reliable.

Prognostic risk score partly correlates with immune cell infiltration

In order to determine whether our risk model is related to the immune microenvironment and provide guidance for immunotherapy response, we analyzed the relationship of risk score to immune cell infiltration. We used the CIBERSORT algorithm to calculate the tumor infiltrating cells in different LUAD groups. Figure 6A displayed the fraction of 22 immune cell types in all LUAD patients. The differences of immune cell infiltration between the high- and low-risk group were shown using a heatmap in Figure 6B. In addition, the fraction of 22 immune cell types between the high- and low-risk group was also analyzed using Wilcoxon test. The high-risk group showed significant higher infiltration levels of macrophages M1, activated mast cells，neutrophils, resting NK cells, activated memory CD4 T cells, and

Tregs, whereas the low-risk group was more correlated with resting mast cells and monocytes infiltration ($p<0.05$, Figure 6C).

Prognostic risk score correlates with the expression of immune checkpoints

Due to the importance of checkpoint inhibitor-based immunotherapy, the differences in immune checkpoint gene expression between the two groups were also explored. Here we systematically investigated the expression of 21 immune checkpoint genes, including *HAVCR2, CD274, CD86, LAG3, LAIR1, PVR, IDO1, CD80, CTLA4, PDCD1, TIGIT, CD200R1, CEACAM1, CD276, CD200, KIR3DL1, BTLA, ADORA2A, LGALS3, VTCN1* and *CLEC4G*. Among them, the expression of 3 immune checkpoints, that is, *CD276, PVR*, and *VTCN1*, were significantly increased in the high-risk group (Figure 7A, B, C). The others were decreased or not statistically significant (Figure S2). The results indicate that immune checkpoint inhibitors may be sensitive in LUAD patients receiving a specific immunotherapy.

Prognostic risk score correlates with the sensitivity of therapeutic drugs

To determine potential drugs to treat LUAD patients based on our lncRNA model, we used the pRRophetic algorithm to estimate the IC50 from the GDSC database. Table S4 reflected that 64 of the 94 drugs showed significant differences between the high- and low-risk group. Figure 7D displays the top 50 therapeutic drugs with significantly different IC50s between the two groups, and the high-risk group is more sensitive to the drugs marked in blue in the figure.

Discussion

LUAD is universally acknowledged as the leading cause of cancer-related deaths worldwide. Due to considerably high morbidity and mortality,an increasing attention is paid to the field of LUAD annually (Tang et al. 2020). M6A modification, the most common epigenetic methylation modification in mammalian mRNA and lncRNA, participates in almost every process of RNA metabolism (Dai et al. 2018). Convincing evidence indicated that m6Amodification of lncRNAs plays a critical regulatory role in the development and progression of lung cancer (Teng et al. 2021). For example, *ALKBH5*-mediated m6A demethylation of lncRNA *RMRP* promotes tumorigenesis of LUAD in vitro and in vivo (Yu and Zhang 2021). Moreover, m6A modification mediated by *METTL3* leads to the upregulation of *LCAT3* and promotes the proliferation and invasion of LUAD through the *LCAT3-FUBP1-MYC* axis (Qian et al. 2021). Integrating above evidence, we were confident that m6A modification of lncRNAs could affect both occurrence and progression of LUAD. However, whether and how m6A-modified lncRNAs function in the prognosis and immunity of LUAD have yet to be fully explored.

In our present study, we acquired 59 paracancerous and 526 cancerous LUAD samples from TCGA dataset. Combined differential lncRNA expression analysis and the Spearman correlation analysis with m6A-related genes, a total of 742 DEm6A-lncRNAs were identified. With the criterion of the AUC of ROC of 7 years>0.6, 26 DEm6A-lncRNAs were finally determined to be related to the prognosis. Through the LASSO regression analysis, the m6A-related lncRNAs prognostic risk model was constructed based on 19 m6A-related prognostic lncRNAs.

Several of the 19 lncRNAs were reported to be associated with lung cancer progression. For instance, LINC01224 can enhance tumor progression and cisplatin resistance via sponging miR-2467 and promote irradiation resistance by regulating the stem cell-like properties of NSCLC cells through the transcription of *ZNF91* (Fu et al. 2021; Xiao et al. 2021). In contrast, *NKILA* inhibits proliferation or migration of NSCLC through *IL-11/STAT3*pathway or *NF-kappaB/Snail* pathway (Lu et al. 2017; Liu and Shi 2019). *HCG18*, as an oncogene in LUAD, contributes to LUAD progression via miR-34a-5p/*HMMR* axis (Li et al. 2020). Similarly, *OGFRP1* accelerates progression of NSCLC partly through upregulating *LYPD3* expression by sponging miR-124-3p (Tang et al. 2018). Nevertheless, investigations on how lncRNAs interact with m6A-related genes are rare. Therefore, we focused on their interactions to explore potential prognostic biomarkers or therapeutic targets for LUAD.

The 19 m6A-related lncRNAs prognostic risk model was well established and then further prognostic analysis was conducted based on the risk score of each patient. We found that the overall survival of patients in the high-risk group was lower than that of patients in the low-risk group in both the training cohort and the testing cohort through Kaplan–Meier curves, suggesting that poor prognosis might be related to the high risk score. The AUCs of ROC of 7 years were 0.73 and 0.713, respectively in the training cohort and the testing cohort, which indicated that this model had superior accuracy. Clinically, pathological stage is a key factor for the prognosis of LUAD (Jurisic et al. 2018). However, due to tumor heterogeneity, the clinical outcomes of patients with the same stage are always different, indicating that the current staging system for predicting prognosis is not accurate (Arbour et al. 2018). Furthermore, we assessed the relationship of the risk model to clinical factors. On the one hand, multivariate Cox regression analysis showed that the pathological stage and the risk score could be used as independent prognostic risk factors. On the other hand, the AUC of the risk score was also higher than that of clinicopathological characteristics, indicating that the risk score could better predict LUAD tumorigenesis.

Emerging evidence has demonstrated that several components of the TME, including immune cells, chemokines, cytokines, inhibitory receptors and ligands, play an irreplaceable role in tumorigenesis and progression (Bujak et al. 2020). Multitudes of reports have also verified that m6A modification is of undeniable importance in the immune system (Shulman and Stern-Ginossar 2020). However, whether m6A regulators and lncRNAs can synergistically affect the immune landscape in LUAD has not yet been recognized. In the present study, the risk score based on the 19 m6A-related lncRNAs was significantly correlated with the expression level of immune checkpoints (*CD276, PVR*, and *VTCN1*), suggesting that poor prognosis might be associated with the induction of high immune checkpoint expression. This finding was consistent with a previous study, which indicated that the prognosis of patients with high *PD-L1* level is worse than that of patients with low *PD-L1* level (Boscolo et al. 2020). Risk scores were negatively correlated with the infiltration levels of resting mast cells and monocytes, but were positively correlated with the infiltration levels of macrophages M1, activated mast cells, neutrophils, resting NK

cells, activated memory CD4 Tcells, and Tregs. These findings indicated that m6A-related lncRNAs are involved in TME regulation of LUAD to a certain extent. In addition, the drugs marked in blue in Figure 7D differed between high-risk and low-risk groups and may have better efficacy in high-risk groups. Taken together, given the superior ability of our m6A-related lncRNA risk model in predicting therapeutic effect, it might be better applied to immunotherapy.

In conclusion, our study systematically identified m6A-related lncRNAs and established a new prognostic model of LUAD based on these lncRNAs. Meanwhile, we also explored the association between this model and the immune microenvironment status of LUAD, providing a new direction for the immunotherapy of LUAD patients. Supplementally, we investigated the expression of 4 of 19 significant DElncRNAs through preliminary RT-qPCR validation in LUAD tissue samples. However, our current research has some limitations. First, the risk model constructed in this study is based on the public clinical database TCGA. The results were confirmed in the TCGA cohort, but external verification was lacking. Therefore, we should include more prospective real-world data to verify the clinical utility of this model we have established. Second, the current study is insufficiency in experimental validation. Although we performed RT-qPCR assay to confirm 4 identified m6A-related lncRNAs in LUAD tissue samples, the other m6A-related lncRNAs remain unexplored. Third, nor have there been studies to evaluate how m6A modifications function on the identified m6A-related lncRNAs. Fourth, although multiple clinical trials have confirmed the role of immunotherapy in lung cancer, this model is only associated with some immune infiltrating cells and immune checkpoints. Overall, more experimental and clinical studies are needed in the future to fully elucidate the specific regulatory mechanism of m6A-related lncRNAs, which will help to develop better strategies for the diagnosis and therapy of LUAD.

Funding

This work was supported by grant from the National Key Research and Development Program of China (No. 2016YFC0905900), National Natural Science Foundation of China (No. 81872365, No.81802902) and Jiangsu Provincial Key Research Development Program (No. BE2019731).

Conflict of Interest

The authors declare that they have no conflict of interest.

Abbreviations:

LUAD, lung adenocarcinoma; m6A, N6-methyladenosine; TCGA, The Cancer Genome Atlas; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; NSCLC, non-small cell lung cancer; mRNAs, messenger RNAs; miRNAs, microRNAs; lncRNAs, long non-coding RNAs; YTHDC2, YT521-B homology domain containing 2; METTL3, methyltransferase-like 3; ceRNAs, competing endogenous RNAs; MREs, miRNA response elements; ICIs, immune checkpoint inhibitors; PD-L1, programmed death ligand 1; GRCh38, Genome Reference Consortium Human Build 38; DElncRNAs, differentially expressed lncRNAs; DEm6A-lncRNAs, differentially expressed m6A-related lncRNAs; AUC, the area under the curve; TME, tumor

microenvironment; GDSC, the Genomics of Drug Sensitivity in Cancer; IC50, the half-maximal inhibitory concentration; OS, overall survival.

References:

Arbour K C, Jordan E, Kim H R et al (2018) Effects of Co-occurring Genomic Alterations on Outcomes in Patients with KRAS-Mutant Non-Small Cell Lung Cancer. Clin Cancer Res 24(2)**:** 334-340

Boscolo A, Fortarezza F, Lunardi F et al (2020) Combined Immunoscore for Prognostic Stratification of Early Stage Non-Small-Cell Lung Cancer. Front Oncol 10**:** 564915

Bujak J K, Szopa I M, Pingwara R et al (2020) The Expression of Selected Factors Related to T Lymphocyte Activity in Canine Mammary Tumors. Int J Mol Sci 21(7)

Chen B, Khodadoust M S, Liu C L et al (2018a) Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol 1711**:** 243-259

Chen M, Wei L, Law C T et al (2018b) RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2. Hepatology 67(6)**:** 2254-2270

Couraud S, Zalcman G, Milleron B et al (2012) Lung cancer in never smokers--a review. Eur J Cancer 48(9)**:** 1299-1311

Dai D, Wang H, Zhu L et al (2018) N6-methyladenosine links RNA metabolism to cancer progression. Cell Death Dis 9(2)**:** 124

Dai F, Wu Y, Lu Y et al (2020) Crosstalk between RNA m(6)A Modification and Non-coding RNA Contributes to Cancer Growth and Progression. Mol Ther Nucleic Acids 22**:** 62-71

Denisenko T V, Budkevich I N, and Zhivotovsky B (2018) Cell death-based treatment of lung adenocarcinoma. Cell Death Dis 9(2)**:** 117

Desrosiers R, Friderici K, and Rottman F (1974) Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci U S A 71(10)**:** 3971-3975

Dong H X, Wang R, Jin X Y et al (2018) LncRNA DGCR5 promotes lung adenocarcinoma (LUAD) progression via inhibiting hsa-mir-22-3p. J Cell Physiol 233(5)**:** 4126-4136

Ettinger D S, Wood D E, Akerley W et al (2015) Non-Small Cell Lung Cancer, Version 6.2015. J Natl Compr Canc Netw 13(5)**:** 515-524

Friedman J, Hastie T, and Tibshirani R (2010) Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw 33(1)**:** 1-22

Fu W, Zhao J, Hu W et al (2021) LINC01224/ZNF91 Promote Stem Cell-Like Properties and Drive Radioresistance in Non-Small Cell Lung Cancer. Cancer Manag Res 13**:** 5671-5681

Gadgeel S, Rodriguez-Abreu D, Speranza G et al (2020) Updated Analysis From KEYNOTE-189: Pembrolizumab or Placebo Plus Pemetrexed and Platinum

for Previously Untreated Metastatic Nonsquamous Non-Small-Cell Lung Cancer. J Clin Oncol 38(14): 1505-1517

Geeleher P, Cox N, and Huang R S (2014) pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One 9(9): e107468

Han T, Xu D, Zhu J et al (2021) Identification of a robust signature for clinical outcomes and immunotherapy response in gastric cancer: based on N6-methyladenosine related long noncoding RNAs. Cancer Cell Int 21(1): 432

Heagerty P J, and Zheng Y (2005) Survival model predictive accuracy and ROC curves. Biometrics 61(1): 92-105

Hengl T, Mendes de Jesus J, Heuvelink G B et al (2017) SoilGrids250m: Global gridded soil information based on machine learning. PLoS One 12(2): e0169748

Hu W, Wang G, Chen Y et al (2020) Coupled immune stratification and identification of therapeutic candidates in patients with lung adenocarcinoma. Aging (Albany NY) 12(16): 16514-16538

Jiang Q, Sun J, Chen H et al (2021) Establishment of an Immune Cell Infiltration Score to Help Predict the Prognosis and Chemotherapy Responsiveness of Gastric Cancer Patients. Front Oncol 11: 650673

Jin D, Guo J, Wu Y et al (2019) m(6)A mRNA methylation initiated by METTL3 directly promotes YAP translation and increases YAP activity by regulating the MALAT1-miR-1914-3p-YAP axis to induce NSCLC drug resistance and metastasis. J Hematol Oncol 12(1): 135

Jurisic V, Obradovic J, Pavlovic S et al (2018) Epidermal Growth Factor Receptor Gene in Non-Small-Cell Lung Cancer: The Importance of Promoter Polymorphism Investigation. Anal Cell Pathol (Amst) 2018: 6192187

Lei F, Zhang H, and Xie X (2019) Comprehensive analysis of an lncRNA-miRNA-mRNA competing endogenous RNA network in pulpitis. PeerJ 7: e7135

Li W, Pan T, Jiang W et al (2020) HCG18/miR-34a-5p/HMMR axis accelerates the progression of lung adenocarcinoma. Biomed Pharmacother 129: 110217

Li Z, Zhang J, Zheng H et al (2019) Modulating lncRNA SNHG15/CDK6/miR-627 circuit by palbociclib, overcomes temozolomide resistance and reduces M2-polarization of glioma associated microglia in glioblastoma multiforme. J Exp Clin Cancer Res 38(1): 380

Liu D, and Shi X (2019) Long non-coding RNA NKILA inhibits proliferation and migration of lung cancer via IL-11/STAT3 signaling. Int J Clin Exp Pathol 12(7): 2595-2603

Lu Z, Li Y, Wang J et al (2017) Long non-coding RNA NKILA inhibits migration and invasion of non-small cell lung cancer via NF-kappaB/Snail pathway. J Exp Clin Cancer Res 36(1): 54

Ma L, Chen T, Zhang X et al (2021) The m(6)A reader YTHDC2 inhibits lung adenocarcinoma tumorigenesis by suppressing SLC7A11-dependent antioxidant function. Redox Biol 38**:** 101801

Marinelli D, Mazzotta M, Scalera S et al (2020) KEAP1-driven co-mutations in lung adenocarcinoma unresponsive to immunotherapy despite high tumor mutational burden. Ann Oncol 31(12)**:** 1746-1754

Qian X, Yang J, Qiu Q et al (2021) LCAT3, a novel m6A-regulated long non-coding RNA, plays an oncogenic role in lung cancer via binding with FUBP1 to activate c-MYC. J Hematol Oncol 14(1)**:** 112

Ritchie M E, Phipson B, Wu D et al (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43(7)**:** e47

Robinson M D, McCarthy D J, and Smyth G K (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1)**:** 139-140

Rodriguez-Canales J, Parra-Cuentas E, and Wistuba, II (2016) Diagnosis and Molecular Classification of Lung Cancer. Cancer Treat Res 170**:** 25-46

Shannon P, Markiel A, Ozier O et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11)**:** 2498-2504

Shulman Z, and Stern-Ginossar N (2020) The RNA modification N(6)-methyladenosine as a novel regulator of the immune system. Nat Immunol 21(5)**:** 501-512

Statello L, Guo C J, Chen L L et al (2021) Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol 22(2)**:** 96-118

Tang L X, Chen G H, Li H et al (2018) Long non-coding RNA OGFRP1 regulates LYPD3 expression by sponging miR-124-3p and promotes non-small cell lung cancer progression. Biochem Biophys Res Commun 505(2)**:** 578-585

Tang Q, Li W, Zheng X et al (2020) MELK is an oncogenic kinase essential for metastasis, mitotic progression, and programmed death in lung carcinoma. Signal Transduct Target Ther 5(1)**:** 279

Teng P C, Liang Y, Yarmishyn A A et al (2021) RNA Modifications and Epigenetics in Modulation of Lung Cancer and Pulmonary Diseases. Int J Mol Sci 22(19)

Wang Y, Li N, Tian D et al (2021) Analysis of m6A-Related lncRNAs for Prognosis Value and Response to Immune Checkpoint Inhibitors Therapy in Hepatocellular Carcinoma. Cancer Manag Res 13**:** 6451-6471

Wu X, Xiao Y, Zhou Y et al (2019) LncRNA FOXP4-AS1 is activated by PAX5 and promotes the growth of prostate cancer by sequestering miR-3184-5p to upregulate FOXP4. Cell Death Dis 10(7)**:** 472

Wu Y L, Zhang L, Fan Y et al (2021) Randomized clinical trial of pembrolizumab vs chemotherapy for previously untreated Chinese patients with PD-L1-positive locally advanced or metastatic non-small-cell lung cancer: KEYNOTE-042 China Study. Int J Cancer 148(9)**:** 2313-2320

Xiao S, Sun L, Ruan B et al (2021) Long non-coding RNA LINC01224 promotes progression and cisplatin resistance in non-small lung cancer by sponging miR-2467. Pulm Pharmacol Ther 70**:** 102070

Yang Y, Zhang Y, Miao L et al (2020) LncRNA PPP1R14B-AS1 Promotes Tumor Cell Proliferation and Migration via the Enhancement of Mitochondrial Respiration. Front Genet 11**:** 557614

Yu H, and Zhang Z (2021) ALKBH5-mediated m6A demethylation of lncRNA RMRP plays an oncogenic role in lung adenocarcinoma. Mamm Genome 32(3)**:** 195-203

Zaccara S, Ries R J, and Jaffrey S R (2019) Reading, writing and erasing mRNA methylation. Nat Rev Mol Cell Biol 20(10)**:** 608-624

Zhou Z, Zhang J, Xu C et al (2021) An integrated model of N6-methyladenosine regulators to predict tumor aggressiveness and immune evasion in pancreatic cancer. EBioMedicine 65**:** 103271

Figure legends

Fig. 1 The flowchart of the whole study based on TCGA database.

Fig. 2 Screening for DEm6A-lncRNAs in LUAD. (A-B) The heatmap and scatter plot of totally 5,606 lncRNAs obtained from TCGA. (C) Volcano plots from TCGA. X axis represents the average gene expression differences between LUAD samples and normal samples, and Y axis represents the p-value of the logarithmic transformation.Red dots and blue dots represent the up-regulated and down-regulated lncRNAs in LUAD samples, respectively; black dots represent genes that are not differentially expressed between LUAD samples and controls. Fold change > 1.5 and $p < 0.05$ were set as the cut-off criteria. (D) The heatmap of the top 20 up-regulated and down-regulated DElncRNAs. (E) Venn diagrams of DEm6A-lncRNAs from the DElncRNAs and m6A-lncRNAs datasets. (F) The co-expression network between DEm6A-lncRNAs and m6A-related genes.

Fig. 3 Construction of the prognostic risk model for patients with LUAD based on m6A-related lncRNAs. (A-B) LASSO regression was performed.The penalization coefficient λ in the LASSO model was tuned using 10-fold cross-validation and minimum criterion for the selection of m6A-related lncRNAs. (C-D) The distributions of risk score and survival status of all LUAD patients. (E)Univariate Cox regression analysis of the lncRNAs in the risk model. $p< 0.05$ was considered statistically significant.

Fig. 4 Validation of the prognostic risk model for patients with LUAD based on m6A-related lncRNAs. (A-C) Kaplan-Meier analysis showed that the high-risk group exhibited worse survival outcome than the low-risk group in the training (A), testing (B) and entire (C) cohort (HR =2.84, 95% confidence interval (CI) = 1.86 – 4.33; HR =2.17, 95% CI= 1.38 – 3.41; HR =2.62, 95% CI = 1.92 – 3.58, respectively). $p< 0.05$ was considered statistically significant. (D-F) Receiver operating characteristic (ROC) curves of the risk model for predicting prognosis in the training (D), testing (E) and entire (F) cohort.

Fig. 5 Prognostic risk score correlates with clinicopathological characteristics. (A-B) Univariate and multivariate Cox regression analyses between the risk model and clinical factors showed that the risk score calculated by LASSO model was an independent prognostic predictor. (C-F) Wilcoxon test was performed to evaluate the differences between the risk scores and clinicopathological characteristics (T stage, N stage, M stage, and TNM stage). (G) ROC curves of the risk score and different clinicopathological characteristics showed that the risk score could better predict prognosis of LUAD. $p< 0.05$ was considered statistically significant.

Fig. 6 Prognostic risk score correlates with immune cell infiltration. (A) The barplot of the fraction of 22 immune cell types in all LUAD patients. (B) The heatmap of immune cell infiltration between the high- and low-risk group. (C) Wilcoxon test was performed to analyze the difference in the fraction of 22 immune cell types between the high- and low-risk group. $p< 0.05$ was considered statistically significant.

Fig. 7 Prognostic risk score correlates with the expression of immune checkpoints and the sensitivity of therapeutic drugs. (A-C) The expression of 3 immune checkpoints (*CD276, PVR,* and *VTCN1*) were significantly increased in the high-risk group compared with the low-risk group. $p< 0.05$ was considered statistically significant. (D) The top 50 therapeutic drugs with significantly different IC50s between the high- and low-risk group.

Fig. S1 Preliminary validation of the lncRNA expression in LUAD tissue samples. The expression of totally 4 lncRNAs in LUAD tissues (n = 20) compared with normal tissues (n = 20) were examined by qRT-PCR and normalized to GAPDH expression. The expression of PPP1R14B-AS1, LINC01224 and LINC02321 were up-regulated, but the expression of AF131215.2 was down-regulated. $p< 0.05$ was considered statistically significant.

Fig. S2 The expression of other 18 immune checkpoint genes were decreased or not statistically significant.

Figures

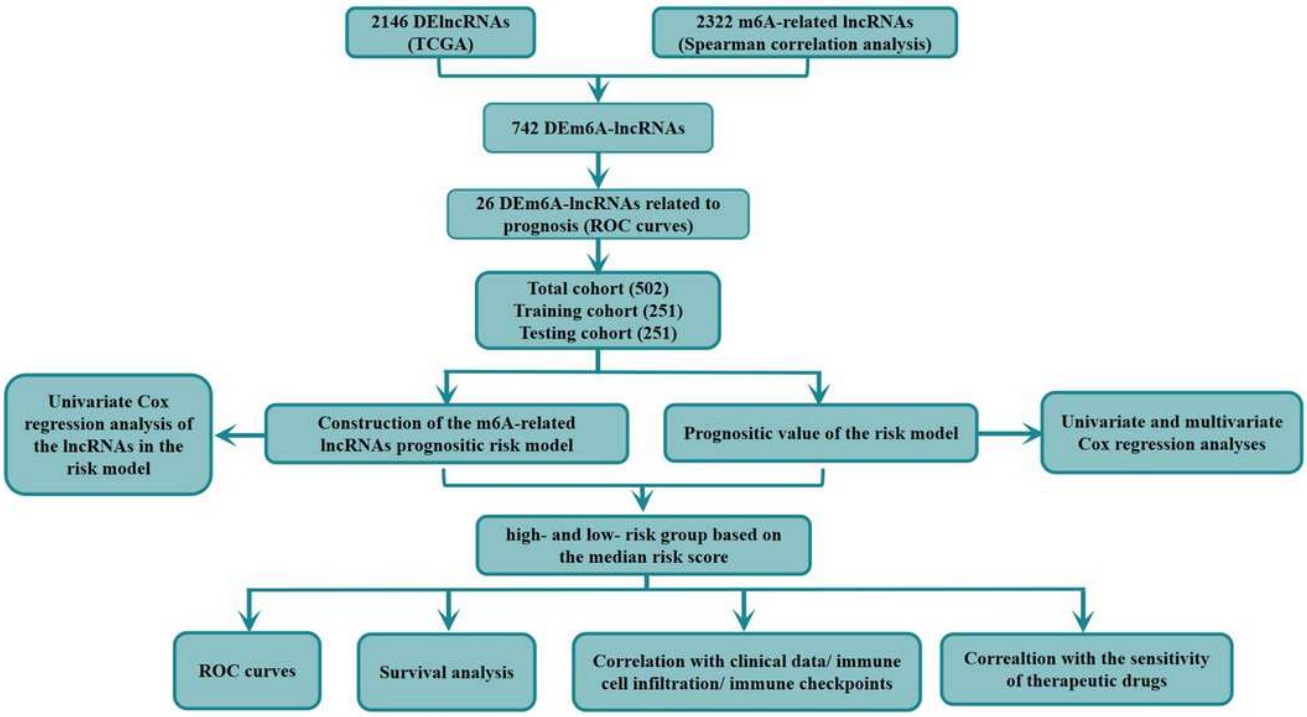

Figure 1

The flowchart of the whole study based on TCGA database

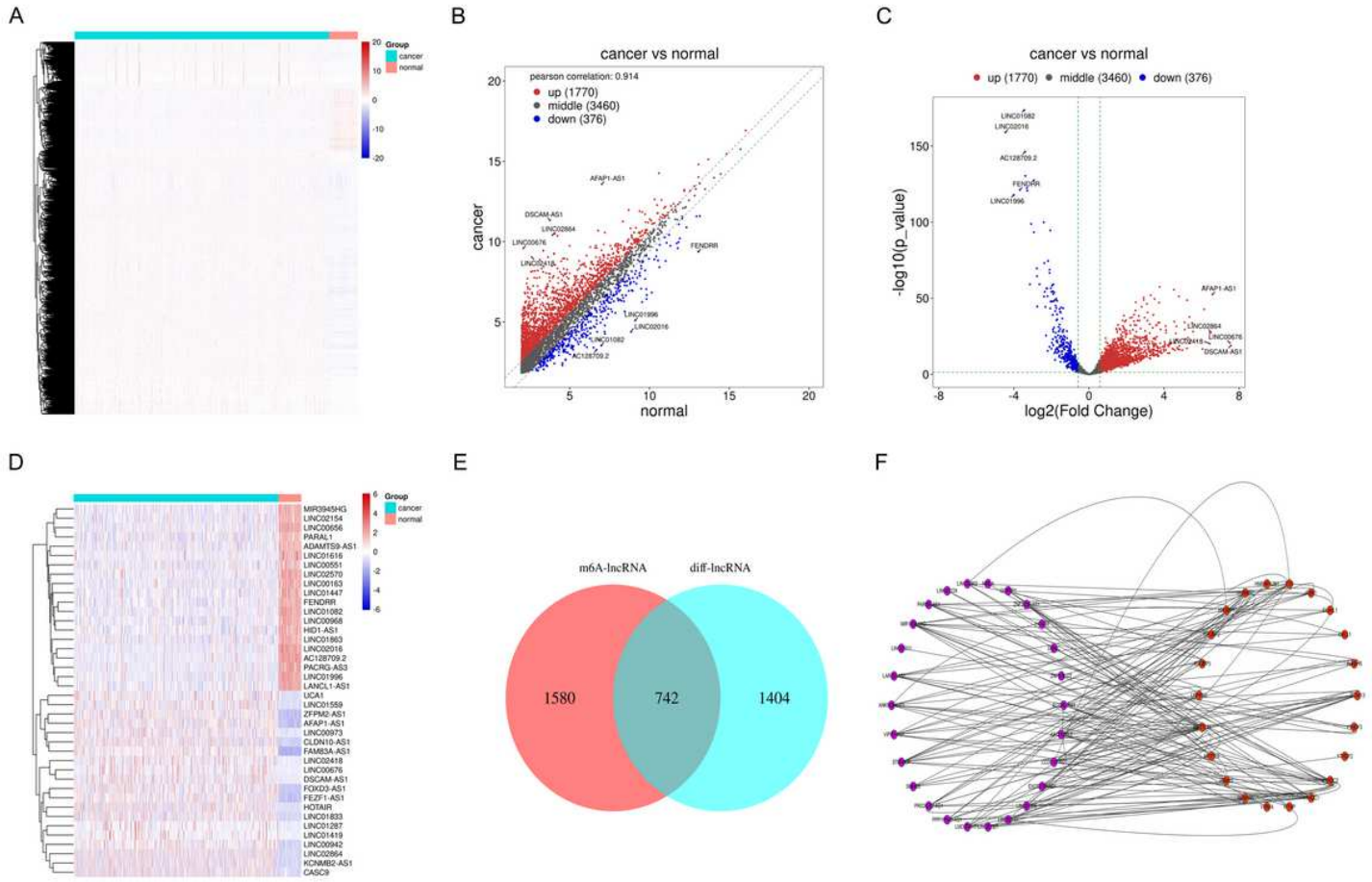

Figure 2

Screening for DEm6A-lncRNAs in LUAD. (A-B) The heatmap and scatter plot of totally 5,606 lncRNAs obtained from TCGA. (C) Volcano plots from TCGA. X axis represents the average gene expression differences between LUAD samples and normal samples, and Y axis represents the p-value of the logarithmic transformation.Red dots and blue dots represent the up-regulated and down-regulated lncRNAs in LUAD samples, respectively; black dots represent genes that are not differentially expressed between LUAD samples and controls. Fold change > 1.5 and $p < 0.05$ were set as the cut-off criteria. (D) The heatmap of the top 20 up-regulated and down-regulated DElncRNAs. (E) Venn diagrams of DEm6A-lncRNAs from the DElncRNAs and m6A-lncRNAs datasets. (F) The co-expression network between DEm6A-lncRNAs and m6A-related genes.

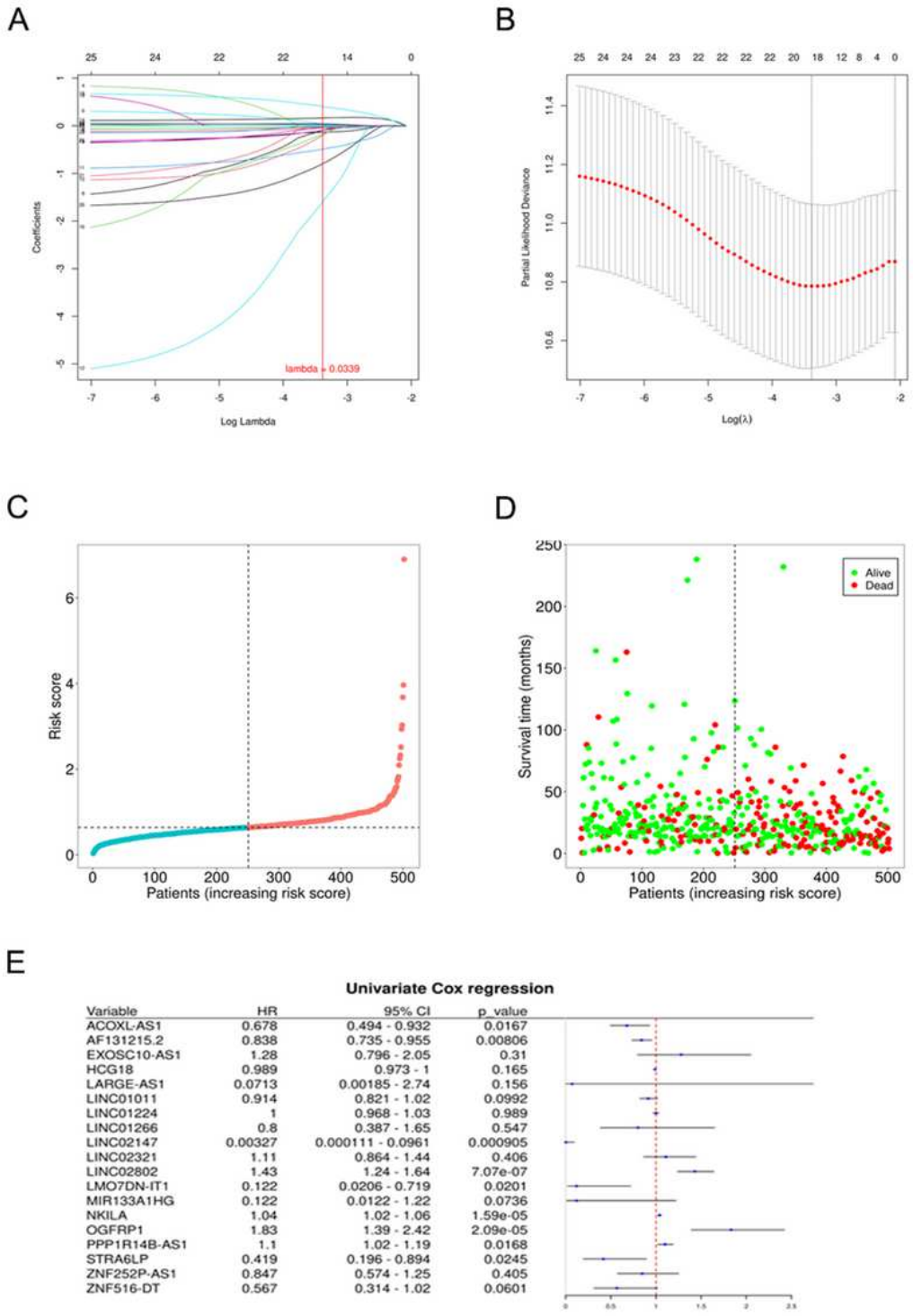

Figure 3

Construction of the prognostic risk model for patients with LUAD based 31 on m6A-related lncRNAs. (A-B) LASSO regression was performed.The penalization coefficient λ in the LASSO model was tuned using 10-fold cross-validation and minimum criterion for the selection of m6A-related lncRNAs. (C-D) The distributions of risk score and survival status of all LUAD patients. (E)Univariate Cox regression analysis of the lncRNAs in the risk model. $p < 0.05$ was considered statistically significant.

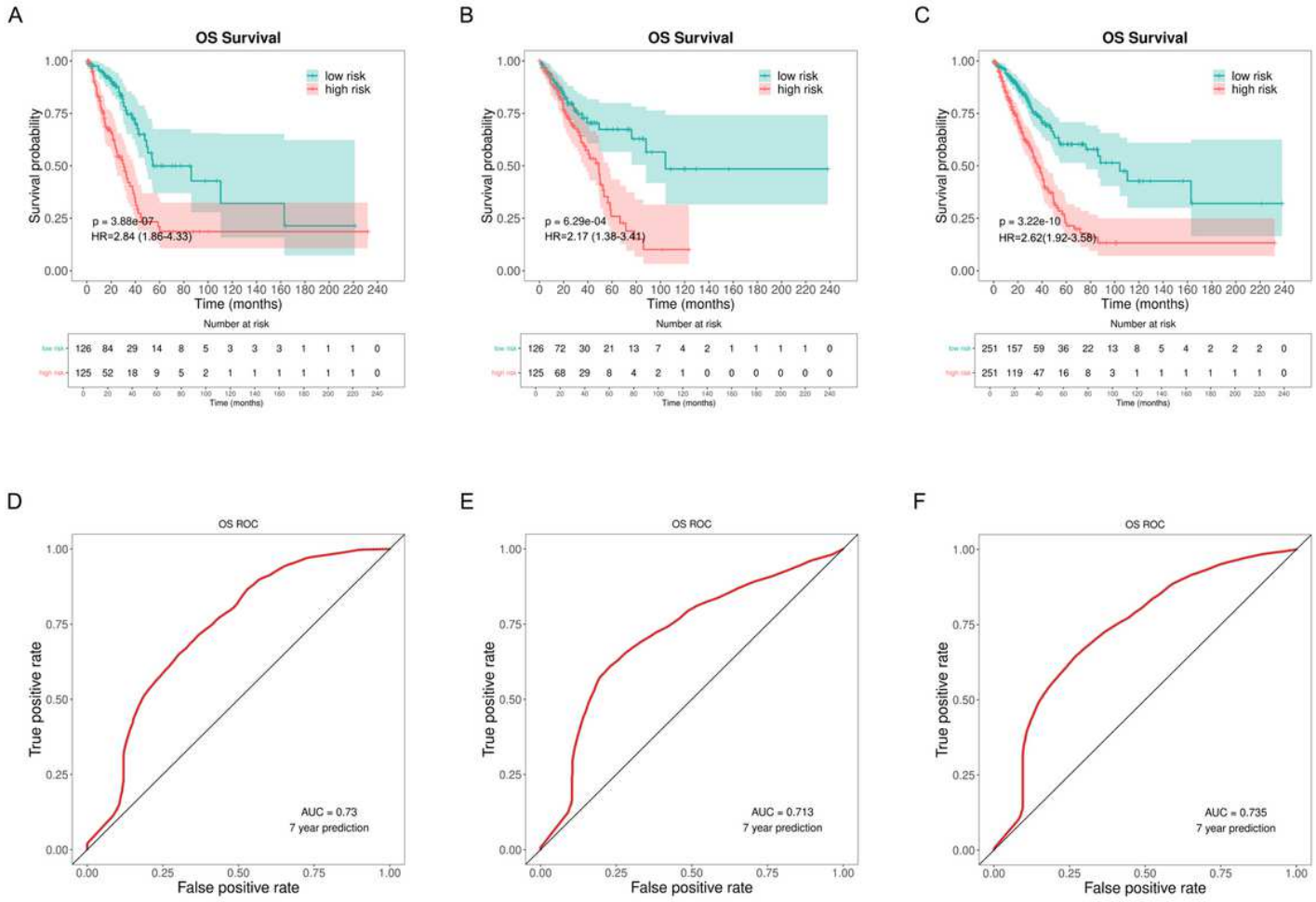

Figure 4

Validation of the prognostic risk model for patients with LUAD based on m6A-related lncRNAs. (A-C) Kaplan-Meier analysis showed that the high-risk group exhibited worse survival outcome than the low-risk group in the training (A), testing (B) and entire (C) cohort (HR =2.84, 95% confidence interval (CI) = 1.86–4.33; HR =2.17, 95% CI= 1.38–3.41; HR =2.62, 95% CI = 1.92–3.58, respectively). p< 0.05 was considered statistically significant. (D-F) Receiver operating characteristic (ROC) curves of the risk model for predicting prognosis in the training (D), testing (E) and entire (F) cohort

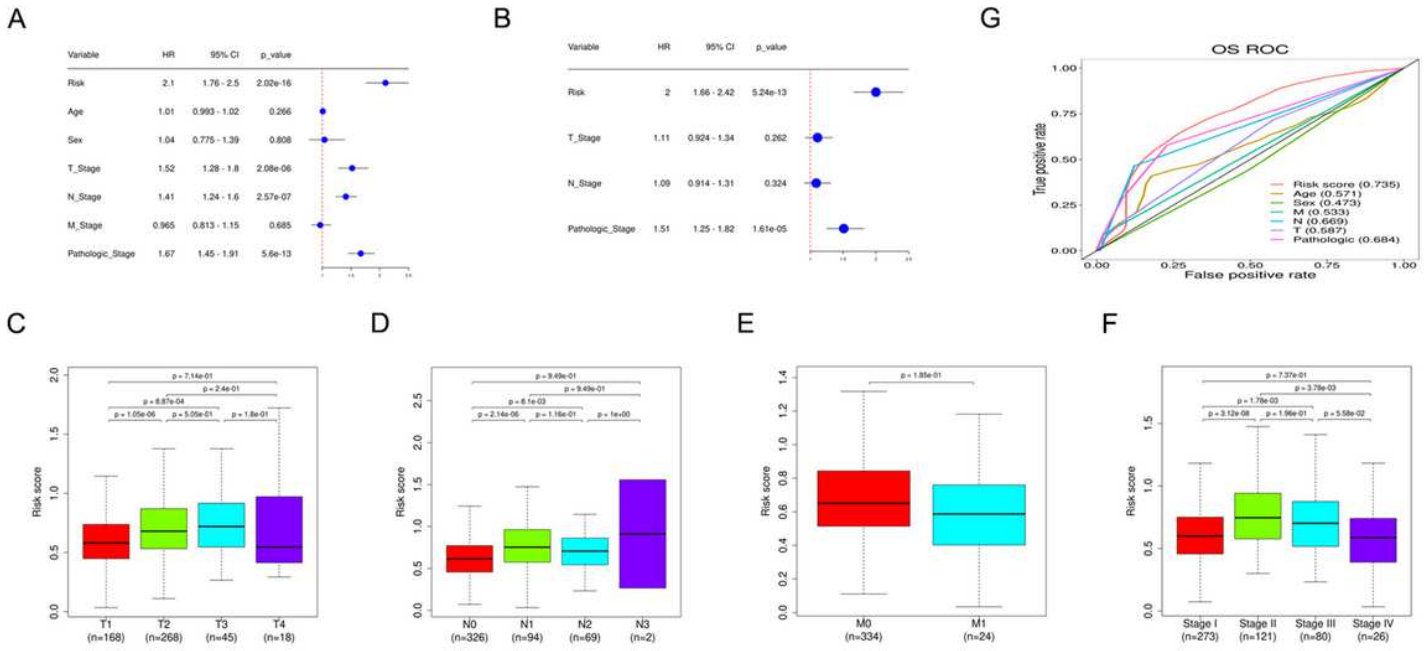

Figure 5

Prognostic risk score correlates with clinicopathological characteristics. (A-B) Univariate and multivariate Cox regression analyses between the risk model and clinical factors showed that the risk score calculated by LASSO model was an independent prognostic predictor. (C-F) Wilcoxon test was performed to evaluate the differences between the risk scores and clinicopathological characteristics (T stage, N stage, M stage, and TNM stage). (G) ROC curves of the risk score and different clinicopathological characteristics showed that the risk score could better predict prognosis of LUAD. p< 0.05 was considered statistically significa

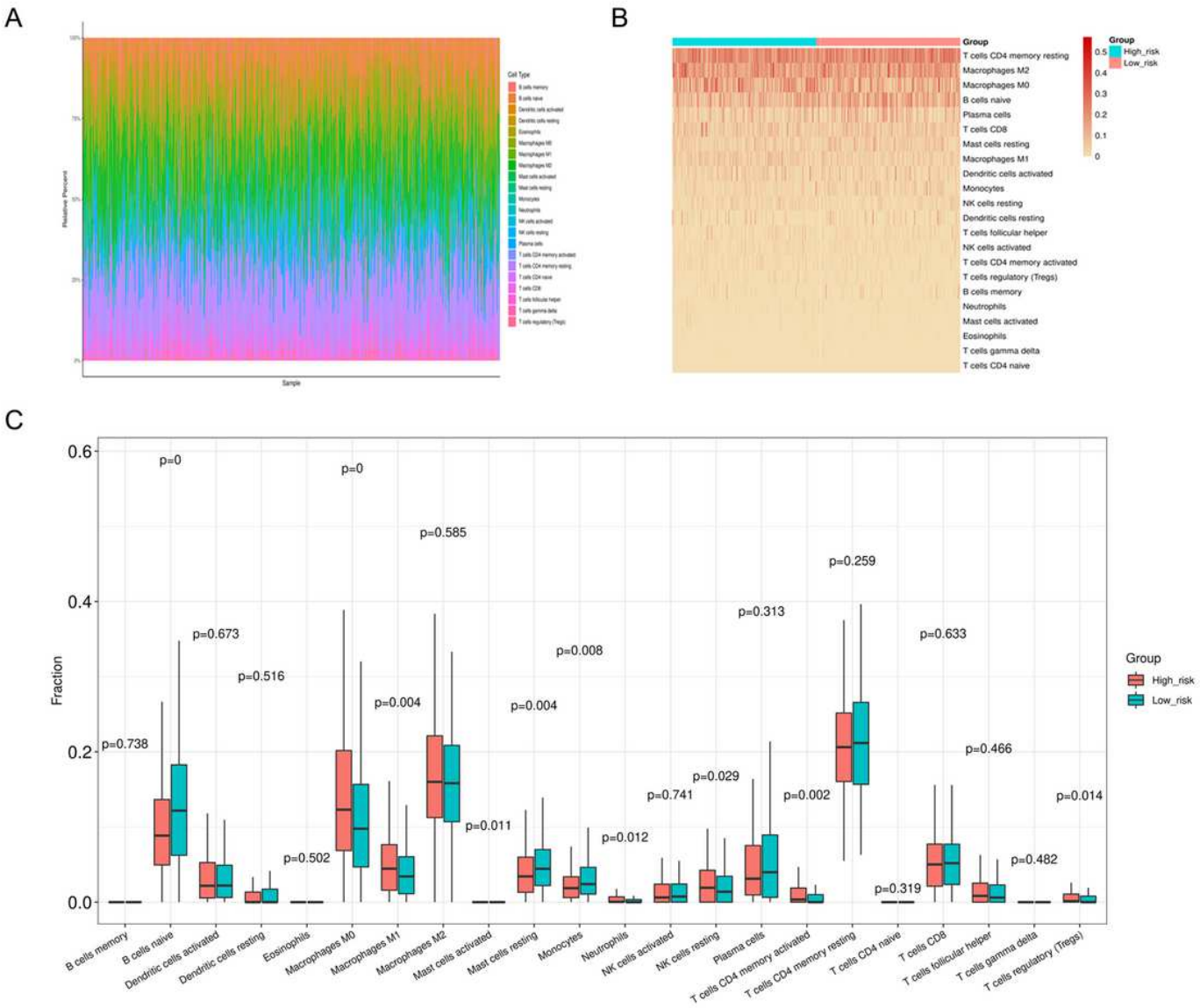

Figure 6

Prognostic risk score correlates with immune cell infiltration. (A) The barplot of the fraction of 22 immune cell types in all LUAD patients. (B) The heatmap of immune cell infiltration between the high- and low-risk group. (C) Wilcoxon test was performed to analyze the difference in the fraction of 22 immune cell types between the high- and low-risk group. p< 0.05 was considered statistically s

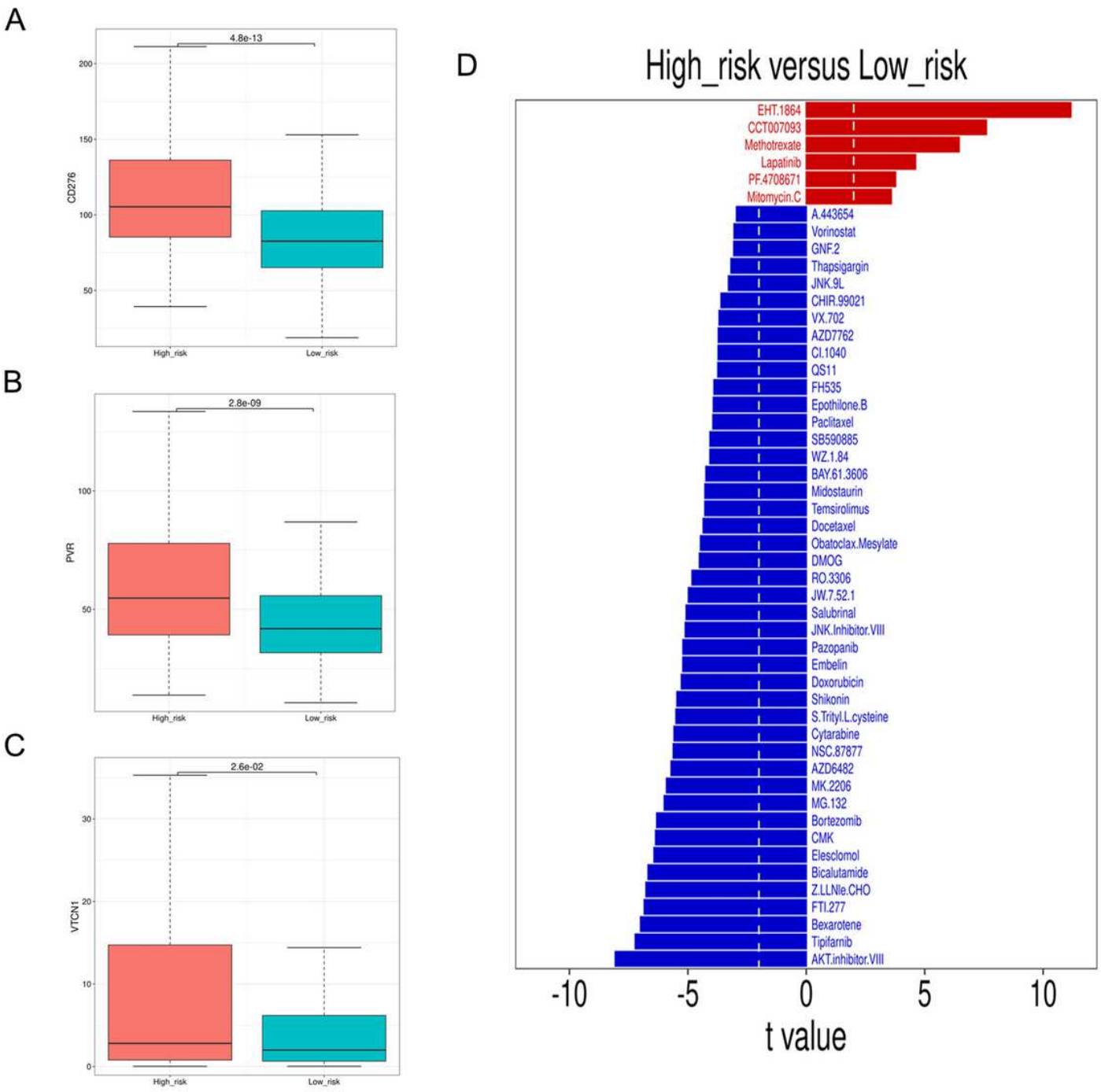

Figure 7

Prognostic risk score correlates with the expression of immune checkpoints and the sensitivity of therapeutic drugs. (A-C) The expression of 3 immune checkpoints (CD276, PVR, and VTCN1) were significantly increased in the high‐risk group compared with the low-risk group. p< 0.05 was considered statistically significant. (D) The top 50 therapeutic drugs with significantly different IC50s between the high- and low-risk group.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS1.tif
- FigS2.tif
- TableS1.docx
- TableS2.docx
- TableS3.docx
- TableS4.docx