Molecular Therapy
Nucleic Acids
Original Article

m^{6}A-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with LUAD

Feng Xu,[1,2,5] Xiaoling Huang,[1,2,5] Yangyi Li,[1,2,5] Yongsong Chen,[2,3] and Ling Lin[2,4]

[1]Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong 515041, China; [2]Clinical Research Center, The First Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong 515041, China; [3]Department of Endocrinology, The First Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong 515041, China; [4]Department of Rheumatology, The First Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong 515041, China

Lung adenocarcinoma (LUAD) is the most frequent subtype of lung cancer worldwide. However, the survival rate of LUAD patients remains low. N6-methyladenosine (m^{6}A) and long noncoding RNAs (lncRNAs) play vital roles in the prognostic value and the immunotherapeutic response of LUAD. Thus, discerning lncRNAs associated with m^{6}A in LUAD patients is critical. In this study, m^{6}A-related lncRNAs were analyzed and obtained by coexpression. Univariate, least absolute shrinkage and selection operator (LASSO), and multivariate Cox regression analyses were conducted to construct an m^{6}A-related lncRNA model. Kaplan-Meier analysis, principal-component analysis (PCA), functional enrichment annotation, and nomogram were used to analyze the risk model. Finally, the potential immunotherapeutic signatures and drug sensitivity prediction targeting this model were also discussed. The risk model comprising 12 m^{6}A-related lncRNAs was identified as an independent predictor of prognoses. By regrouping the patients with this model, we can distinguish between them more effectively in terms of the immunotherapeutic response. Finally, candidate compounds aimed at LUAD subtype differentiation were identified. This risk model based on the m^{6}A-based lncRNAs may be promising for the clinical prediction of prognoses and immunotherapeutic responses in LUAD patients.

INTRODUCTION

Lung adenocarcinoma (LUAD) is the most frequent subtype of lung cancer worldwide.[1–3] With advances in diagnosis, surgery, radiotherapy, and molecular therapeutics, the clinical outcome of LUAD patients has significantly improved. However, the 5-year survival rate of LUAD patients is still at a low level.[4,5] Presently, evidence has shown that the discovery and application of molecular biomarkers can provide prognostic value.[6]

N^6-methyladenosine (m^{6}A), the most abundant RNA modification of eukaryotic cells, is vital in various biological processes and mRNA metabolism, such as RNA processing, transport, and stability.[7,8] m^{6}A modification comprises methyltransferase, signal transducers, and demethylase, which are also called writers, readers and erasers, respectively. Additionally, m^{6}A modification is also a reversible RNA epigenetic process.[9] Changes in the RNA structure can affect various cell processes; therefore, the effect of m^{6}A-regulated long noncoding RNAs (lncRNAs) may be crucial for the proliferation and migration of cancer cells.[10]

Recent studies have shown that m^{6}A modification regulates oncogenesis and tumor development. For example, FEZF1-AS1 influenced by m^{6}A modification regulates the ITGA11/miR-516b-5p axis and finally is upregulated in non-small cell lung cancer (NSCLC).[11] Additionally, m^{6}A methyltransferase-like 3 (METTL3)-induced lncRNA ABHD11-AS1 is upregulated in NSCLC, and its ectopic expression is closely related to the poor prognosis of patients with NSCLC.[12] Recently, several bioinformatics studies have shown that the dysregulation of m^{6}A regulators is involved in LUAD.[13,14] The specific role of m^{6}A regulators in lncRNAs remains unclear; therefore, understanding the mechanism of m^{6}A-related lncRNAs in the development of LUAD may be useful for prognostic targets.

In our study, we abstracted the expression profiles of 14,142 lncRNAs and 21 m^{6}A genes from The Cancer Genome Atlas (TCGA) dataset. Next, we identified the m^{6}A-associated lncRNAs using Pearson's correlation analysis. This model is based on a novel prognostic model of m^{6}A, which was developed to forecast the overall survival (OS) of patients with LUAD. Next, using the publicly available drug sensitivity database, we discovered candidate drugs targeting this m^{6}A-related

Received 7 February 2021; accepted 3 April 2021;
https://doi.org/10.1016/j.omtn.2021.04.003.

[5]These authors contributed equally

Correspondence: Ling Lin, Department of Rheumatology, The First Affiliated Hospital of Shantou University Medical College, No. 57 Changping Road, Shantou, Guangdong 515041, China.
E-mail: llinc@163.net
Correspondence: Yongsong Chen, Department of Endocrinology, The First Affiliated Hospital of Shantou University Medical College, No. 57 Changping Road, Shantou, Guangdong 515041, China.
E-mail: yongsongchen@126.com

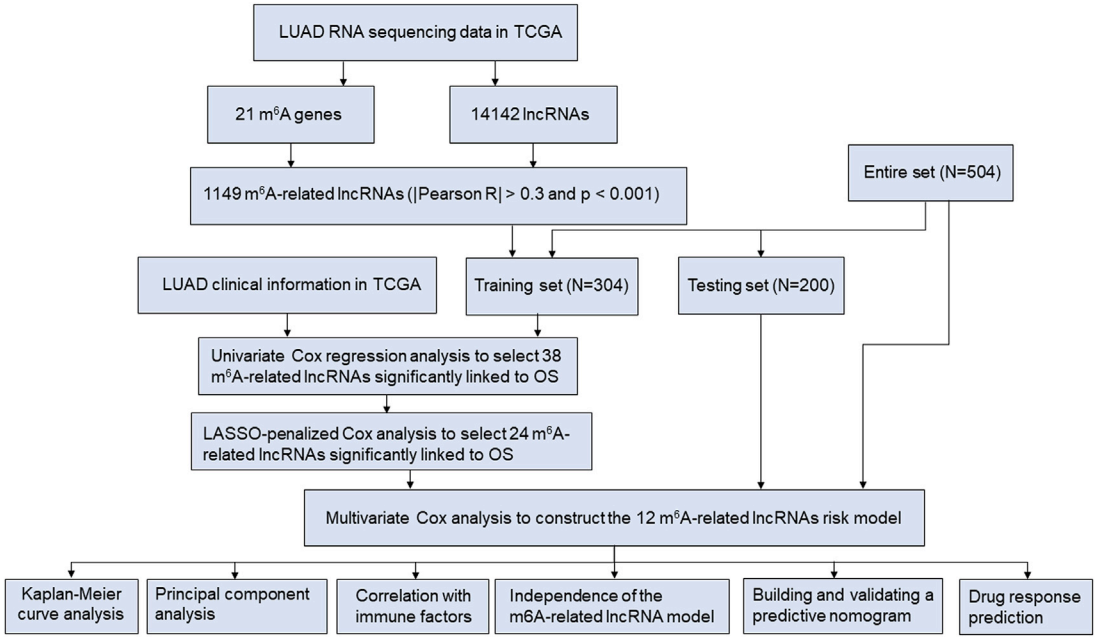

Figure 1. Flow chart of this study

lncRNA signature. Additionally, we explored the relationship association with immunotherapy responses. Finally, we established a nomogram to predict the OS of patients with LUAD.

RESULTS

Identification of m^{6}A-related lncRNAs in patients with LUAD

The detailed workflow for risk model construction and subsequent analyses is shown in Figure 1. The matrix expression of 21 m^{6}A genes and 14,142 lncRNAs was abstracted from the TCGA database. We defined m^{6}A-related lncRNAs as lncRNAs that were significantly related to greater than or equal to one of the 21 m^{6}A genes (|Pearson R| > 0.3 and p < 0.001). Finally, the m^{6}A-lncRNA coexpression network was visualized using the Sankey diagram in Figure 2A, and 1,149 m^{6}A-related lncRNAs were discerned as m^{6}A-related lncRNAs. The correlation between m^{6}A genes and m^{6}A-related lncRNAs in the TCGA entire set is shown in Figure 2B.

Construction and validation of a risk model according to m^{6}A-related lncRNAs in LUAD patients

We screened m^{6}A-associated prognostic lncRNAs from 1,149 m^{6}A-associated lncRNAs in the TCGA training set using univariate Cox regression analysis. Thirty-eight m^{6}A-related lncRNAs in the TCGA dataset were significantly correlated with OS (Figure 3A). LASSO-penalized Cox analysis is a common method of multiple regression analysis. The application of this method not only enhances the forecast accuracy and explainability of the statistical model but also makes variable options and regularization simultaneously. This method is extensively applied for the optimal choice of characteristics in high-dimensional data with an inferior correlation and prominent forecasted value to avoid overfitting. Consequently, this method can effectively discern the most available forecast markers and produce a prognostic indicator to predict clinical results. The dashed perpendicular line illustrates the first-rank value of log λ with the minimum segment likelihood bias. Hence, 24 m^{6}A-related lncRNAs were selected for the subsequent multivariate analysis (Figures 3B and 3C). Next, we used multivariate Cox ratio hazard regression analysis to distinguish autocephalous prognostic proteins. Twelve m^{6}A-related lncRNAs were prognostic proteins independently correlated with OS in the training queue and were used to construct a risk model to assess the prognostic risk of patients with LUAD (Figure 3D). LUAD samples were categorized into low- and high-risk groups based on the median value of the prognostic risk grade. The distribution of risk grades between the low-risk and high-risk groups is depicted in Figure 4A, and the survival status and survival time of patients in the two different risk groups are shown in Figure 4B. The relative expression standards of the 12 m^{6}A-related lncRNAs for each patient are shown in Figure 4C. The survival analysis demonstrated that the OS of the low-risk group was longer than that of the high-risk group (p < 0.001) (Figure 4D).

To test the prognostic capability of this established model, we calculated risk scores for every patient in the test set and entire set using the uniform formula. Figure 5 depicts the distribution of risk grades, pattern of survival status and survival time, and expression of the m^{6}A-related lncRNAs in the testing set (Figures 5A–5C) and entire set (Figures 5E–5G). Kaplan-Meier survival analyses performed on the testing set and entire set showed no differences in the outcomes in the TCGA training set: the OS of LUAD patients with higher risk scores was worse than that of patients with lower risk scores (Figures 5D and 5H). To further predict the ability of the prognostic

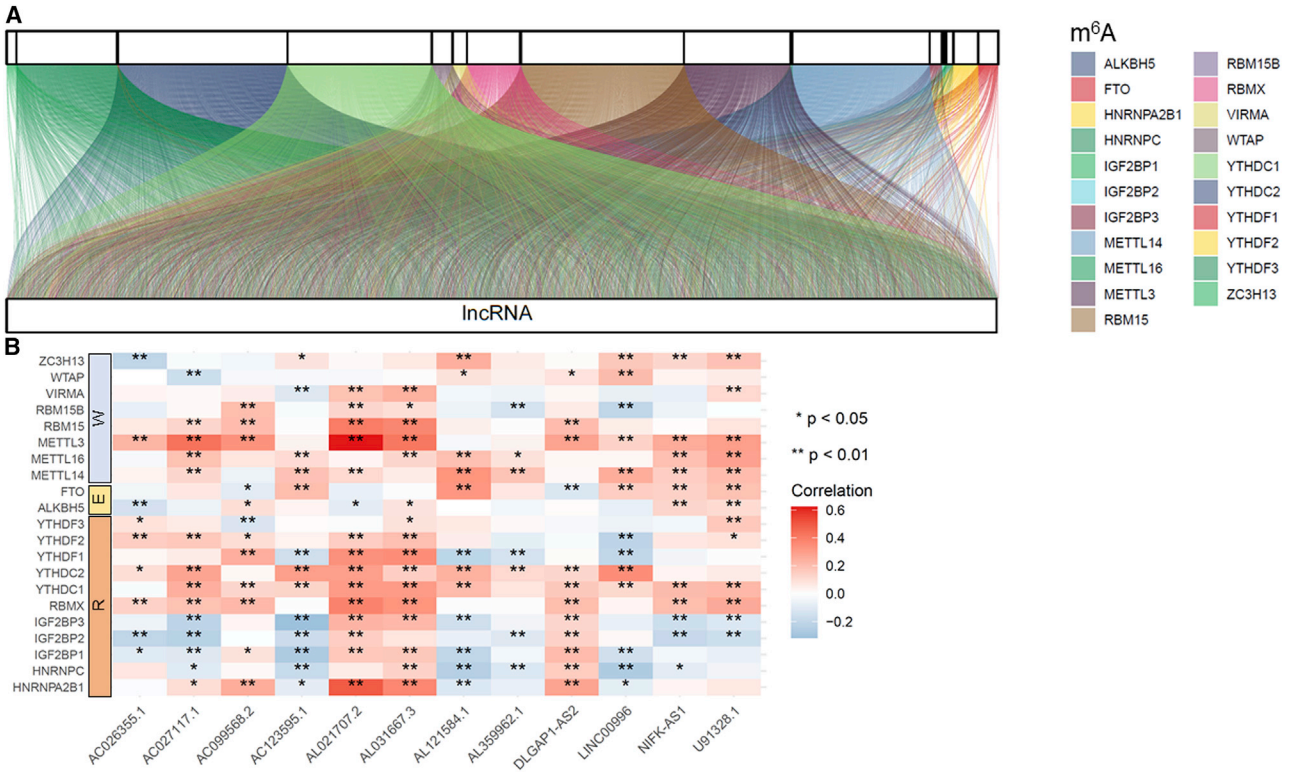

Figure 2. Identification of m^{6}A-related lncRNAs in LUAD patients

(A) Sankey relational diagram for 21 m^{6}A genes and m^{6}A-related lncRNAs. (B) Heatmap for the correlations between 21 m^{6}A genes and the 12 prognostic m^{6}A-related lncRNAs.

model, the disease-free interval (DFI), progression-free interval (PFI), and disease-specific survival (DSS) were explored to distinguish high- and low-risk LUAD patients. As predicted, DFI, PFI, and DSS were different between the low- and high-risk groups, indicating that the m^{6}A-related lncRNA model level affected the prognosis of LUAD patients (Figures S1A–S1C).

The discrepancies in OS stratified by the universal clinicopathologic characteristics were analyzed between the low- and high-risk groups in the TCGA entire set. According to the subgroups classified by gender, age, stage, or tumor stage, the OS of the low-risk group continued to be superior to that of the high-risk group (Figure 6).

Principal-component analysis (PCA) further verifies the grouping ability of the m^{6}A-related lncRNA model

PCA was conducted to test the difference between the low-risk and high-risk groups based on the entire gene expression profiles, 21 m^{6}A genes, 12 m^{6}A-related lncRNAs, and risk model classified by the expression profiles of the 12 m^{6}A-related lncRNAs (Figures 7A–7D). Figures 7A–7C display that the distributions of the high- and low-risk groups were relatively scattered. However, the results obtained based on our model illustrated that the low- and high-risk groups had different distributions (Figure 7D). These results suggest that the prognostic signature can distinguish between the low- and high-risk groups.

Estimation of the tumor immune microenvironment and cancer immunotherapy response using the m^{6}A-related lncRNA model

The enrichment level and activity of several immune cells, pathways, or functions in LUAD were further analyzed based on the m^{6}A-related lncRNA model from 504 LUAD samples. The low-risk and high-risk groups showed prominent differences in the expression of immune indicators (Figure 8A). To explore the underlying molecular mechanisms of the m^{6}A-based model, we performed Gene Ontology (GO) enrichment analysis, which revealed the involvement of many immune-related biological processes (Figure 8B). We next investigated the correlations between the m^{6}A-related lncRNA model and immunotherapeutic biomarkers. Unsurprisingly, we discovered that the high-risk group was more likely to respond to immunotherapy than the low-risk group, indicating that this m^{6}A-based classifier index might serve as an indicator for predicting Tumor Immune Dysfunction and Exclusion (TIDE) (Figure 8C). Using the R package maftools, the mutation data were analyzed and summarized. The mutations were stratified based on the variant effect predictor. The top 20 driver genes with the highest alteration frequency between the high- and low-risk subgroups are shown in Figures 8D and 8E. We then calculated TMB scores based on the TGCA somatic mutation data. The TMB in the low-risk group exceeded that in the high-risk group, showing that the m^{6}A-based classifier index had a high correlation with TMB (Figure 8F). TP53 mutations are correlated with a worse

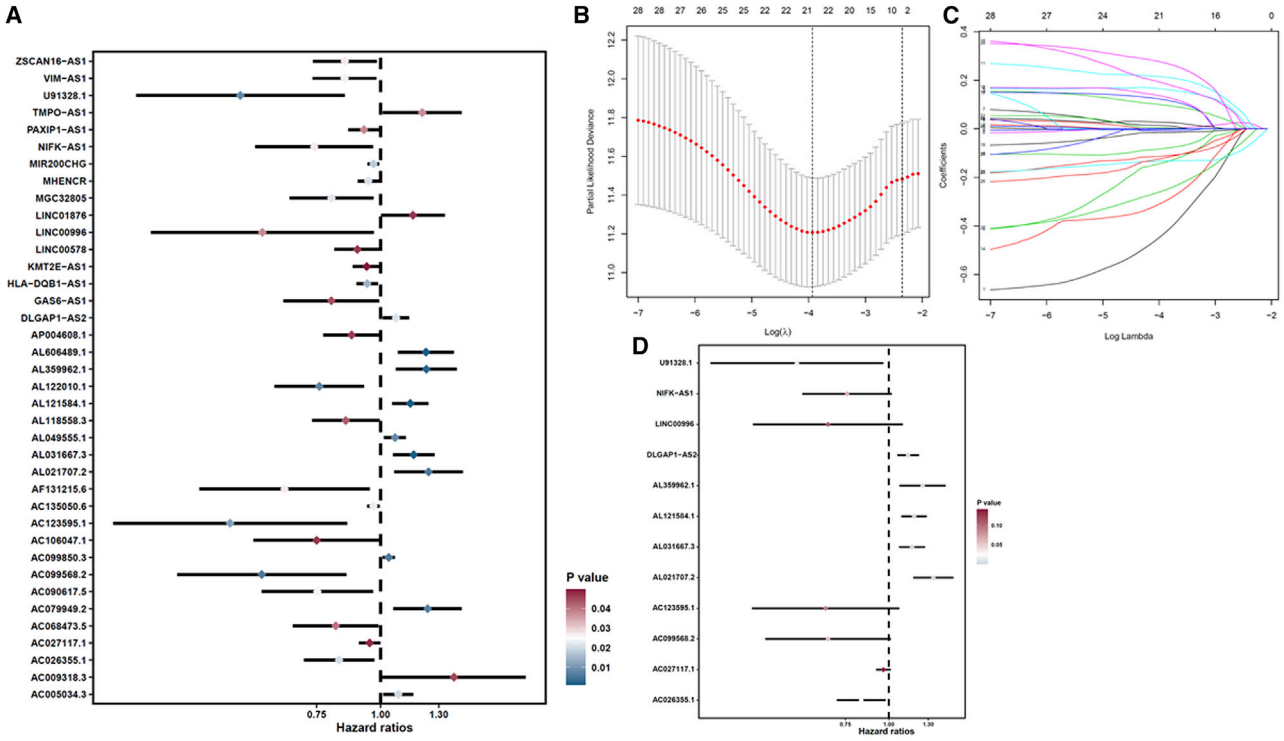

Figure 3. Risk model for LUAD patients based on m^{6}A-related lncRNAs

(A) Univariate Cox regression analysis revealed that the selected lncRNAs significantly correlated with clinical prognosis. (B) The tuning parameters (log λ) of OS-related proteins were selected to cross-verify the error curve. According to the minimal criterion and 1-se criterion, perpendicular imaginary lines were drawn at the optimal value. (C) The LASSO coefficient profile of 24 OS-related lncRNAs and perpendicular imaginary line were drawn at the value chosen by 10-fold cross-validation. (D) Multivariate Cox regression analysis showed 12 independent prognostic lncRNAs.

survival and can be used as a prognostic marker in lung cancer. Therefore, we tested whether the m^{6}A-related lncRNA model could predict the OS outcome better than TP53 mutation status. Patients with TP53 mutation and wild-type TP53 in the high-risk groups (defined as TP53 mutation/high and TP53 wild/high, respectively) presented a worse OS than patients with TP53 mutation and wild-type TP53 in the low-risk groups (TP53 mutation/low and TP53 wild/low, respectively) (Figure 8G). Interestingly, patients with wild-type TP53 in the high-risk group (TP53 wild/high) had worse survival outcomes than patients with TP53 mutation in the low-risk group (TP53 mutation/low). The survival curve of patients with TP53 mutation in the high-risk group (TP53 mutation/high) was similar to that of patients with wild-type TP53 in the high-risk group (TP53 wild/high), indicating that the TP53 mutation status failed to distinguish the survival rate in the high-risk group. Thus, these findings indicate that the m^{6}A-related lncRNA model may have greater prognostic significance than the TP53 mutation status.

Identification of novel candidate compounds targeting the m^{6}A-related lncRNA model

To identify potential drugs targeting our lncRNA model for treating LUAD patients, we used the pRRophetic algorithm to estimate the therapeutic response based on the half-maximal inhibitory concentration (IC_{50}) available in the Genomics of Drug Sensitivity in Cancer (GDSC) database for each sample. We found that 78 compounds were screened out for significant differences in the estimated IC_{50} between these two groups, and the high group was more sensitive to all of these compounds. Figure S2 displays the top 20 compounds that might be used for further analysis in patients with LUAD.

Evaluation of the prognostic risk model of m^{6}A-related lncRNAs and clinical features of LUAD

We performed univariate and multivariate Cox regression analyses to evaluate whether the risk model of 12 m^{6}A-related lncRNAs had independent prognostic characteristics for LUAD. The HR of the risk score and 95% confidence interval (CI) were 1.06 and 1.04–1.08 ($p < 0.001$) in univariate Cox regression analysis, respectively. In multivariate Cox regression analysis, the HR was 1.07 and 95% CI was 1.05–1.09 ($p < 0.001$) (Figure 9A), indicating that the risk model of the 12 m^{6}A-related lncRNAs was unrelated to clinicopathological parameters, such as gender, age, tumor/node/metastasis (TNM) stage, and tobacco smoking history. The conformance index of the risk score and the area under the ROC curve (AUC) were assessed. This process was conducted to better assess the uniqueness and susceptibility of risk scores in predicting outcomes in patients with LUAD. With increasing time, the concordance index of the risk score was

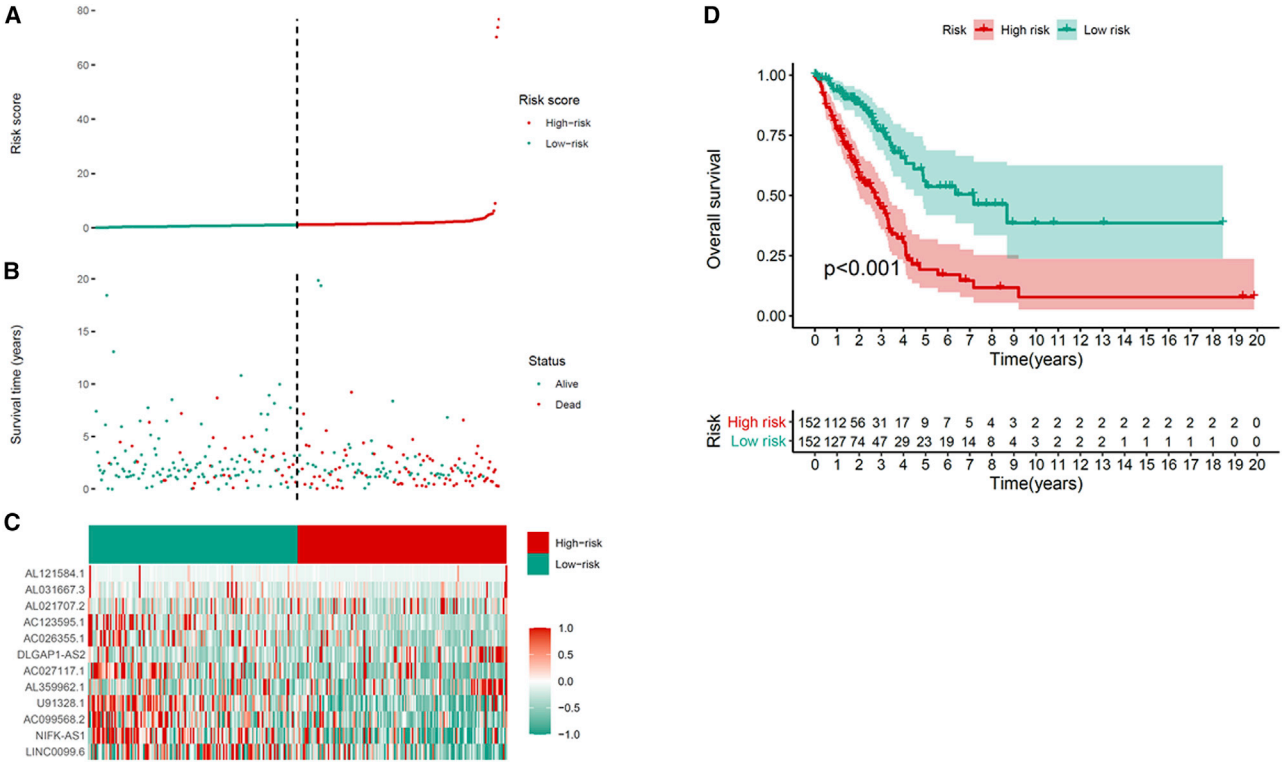

Figure 4. Prognostic value of the risk patterns of the 12 m^{6}A-related lncRNAs in the TCGA training set

(A) Distribution of m^{6}A-related lncRNA model-based risk score. (B) Different patterns of survival status and survival time between the high- and low-risk groups. (C) Clustering analysis heatmap shows the expression standards of the 12 prognostic lncRNAs for each patient. (D) Kaplan-Meier survival curves of the OS of patients in the high- and low-risk groups.

always greater than that of other clinical factors, suggesting that the risk grade could better forecast the prognosis of LUAD (Figure 9B). The AUC of the risk grade was also higher than the AUCs of other clinicopathological characteristics, showing that the prognostic risk model of the 12 m^{6}A-related lncRNAs for LUAD was comparatively dependable (Figure 9C).

Construction and evaluation of the prognostic nomogram

The nomogram comprising the risk grade and clinical risk characteristics was fabricated to predict the 1-, 2-, and 3-year OS incidences. By comparison with clinical factors, the risk grade of the prognostic model showed predominant predictive ability in the nomogram (Figure 10A). Correlation charts displayed that the observed versus predicted rates of the 1-, 2-, and 3-year OS revealed ideal consistency (Figures 10B–10D).

DISCUSSION

As the most common subtype of lung cancer,[15] many medical researchers have focused on studying the occurrence, development, and treatment of LUAD in recent years.[9,16,17] Accumulating studies have shown that different lung cancer subtypes have distinct clinical characteristics and clinical outcomes; thus, an increasing number of studies focusing on identifying signatures with noncoding RNAs have been conducted to predict the survival and immunotherapeutic response in patients with LUAD.[18–20]

As the most abundant posttranscriptional modification in eukaryotic mRNAs and lncRNAs, m^{6}A has extensive regulatory usefulness in regulating mRNA transcription, splicing, and translation and influencing the structure and effect of lncRNAs.[10] Additionally, lncRNA-related studies have drawn attention in numerous cancer fields.[19] m^{6}A regulators can modify specific lncRNAs to maintain malignant n6-methyladenosine and lncRNAs in various tumors.[20] Studies have shown that m^{6}A modification of lncRNAs can affect the occurrence and development of tumors, and lncRNAs may target m^{6}A regulators as competitive endogenous RNAs, affecting tumor invasive progression.[10] m^{6}A modification can adjust lncRNA function by supplying binding sites to m^{6}A reader proteins. It can also regulate local RNA structure to permit concrete RNA-binding proteins to enter the surrounding m^{6}A residue. Additionally, m^{6}A modification can affect the formation of the RNA-DNA triple helix, in which a lncRNA binds to the series through the Hoogsteen base pair in the main groove of double-stranded DNA.[21] Furthermore, m^{6}A may impact the reciprocity site between lncRNAs and specific DNA.[22] Both m^{6}A and lncRNAs are important regulators of LUAD tumorigenesis.[23] However, studies on the pathological role of m^{6}A

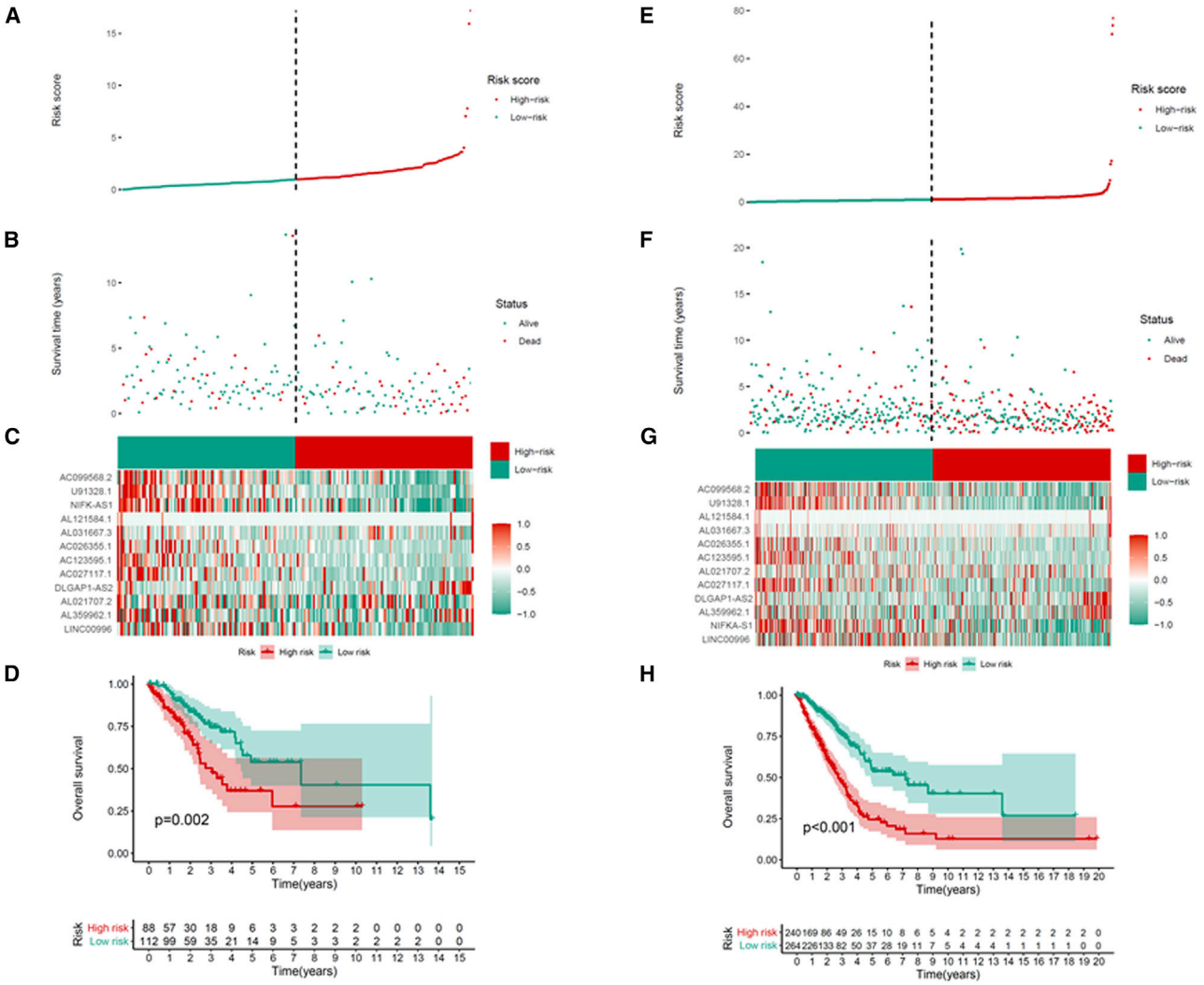

Figure 5. Prognostic value of the risk model of the 12 m^{6}A-related lncRNAs in the TCGA testing and entire sets

(A) Distribution of m^{6}A-related lncRNA model-based risk score for the testing set. (B) Patterns of the survival time and survival status between the high- and low-risk groups for the testing set. (C) Clustering analysis heatmap shows the display levels of the 12 prognostic lncRNAs for each patient in the testing set. (D) Kaplan-Meier survival curves of the OS of patients in the high- and low-risk groups for the testing set. (E) Distribution of the m^{6}A-related lncRNA model-based risk score for the entire set. (F) Patterns of the survival time and survival status between the high- and low-risk groups for the entire set. (G) Clustering analysis heatmap shows the expression levels of the 12 prognostic lncRNAs for each patient for the entire set. (H) Kaplan-Meier survival curves of OS of patients in the low- and high-risk groups for the entire set.

and lncRNAs in LUAD progression remain limited, and studies on the biological mechanisms and prognostic biomarkers of LUAD concerning m^{6}A-related lncRNAs are still lacking.[24,25] In the present study, we were inspired by the function of m^{6}A and lncRNAs in LUAD; thus, we attempted to construct an independent model based on m^{6}A-related lncRNAs.

In our study, 1,149 m^{6}A-related lncRNAs from the TCGA dataset were identified in this paper to explore the prognostic function of m^{6}A-related lncRNAs. The TCGA dataset confirmed the prognostic value of 24 m^{6}A-related lncRNAs, and 12 of them were applied to construct an m^{6}A-related lncRNA model to predict the OS of patients with LUAD. Among them, DLGAP1-AS2 contributes to glioma cell proliferation, migration, and apoptosis by upregulating the expression of the downstream target YAP1.[26] Additionally, as immune-related lncRNAs, AC123595.1 and AC026355.1 can increase the predicted value of LUAD.[27] Additionally, other lncRNAs were revealed for the first time. Subsequently, LUAD patients were separated into high- and low-risk groups based on the intermediate risk score, and the high-risk group had apparently poor clinical results. Multivariate Cox regression analysis showed that the m^{6}A-related lncRNA model was an autocephalous risk element of OS. ROC analysis showed that the model was superior to conventional clinical features in the survival prediction for LUAD. We also established a nomogram

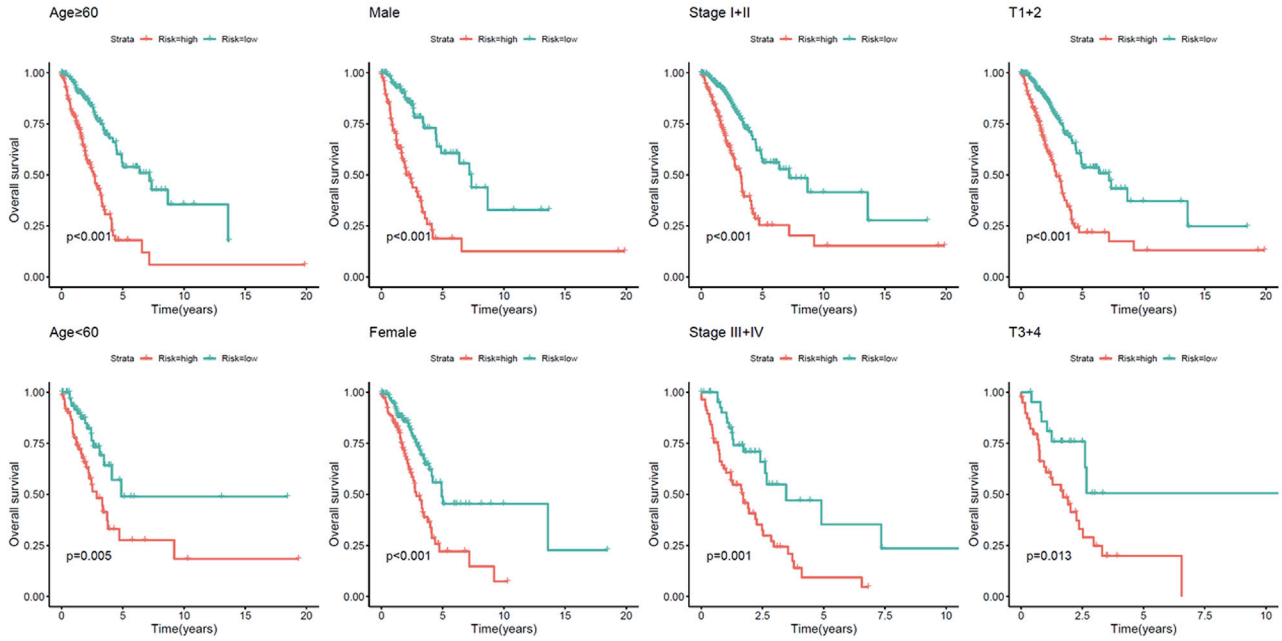

Figure 6. Kaplan-Meier curves of OS differences stratified by gender, age, tumor grade, or TNM stage between the high- and low-risk groups in the TCGA entire set

showing perfect consistency between the observed and predicted rates for the 1-year, 3-year, and 5-year OS. Finally, the observed versus 1-year, 3-year, and 5-year OS forecasting rates displayed excellent consistency. The risk model based on 12 m^{6}A-related lncRNAs that were independently associated with OS was fairly accurate, and this prediction model could identify novel biomarkers for subsequent studies.

The TMB is the total number of somatic coding mutations and is related to the emergence of neoantigens that trigger antitumor immunity.[28] Recent studies revealed the TMB as a valid biomarker to predict the response to PD-L1 treatment.[29] We found that the TMB in the low-risk group exceeded that the in high-risk group. In addition, an increasing number of studies have used the TIDE prediction score, which is a computational framework developed for immunotherapeutic prediction,[30] and its predictive function has been validated successfully. In our study, the prediction of the TIDE algorithm suggests that patients with the high-risk subtype have a superior response to immunotherapy. According to the above results, we infer that this prediction model may provide dependable immune biomarkers for oncotherapy. Additionally, our study provides new insight into the molecular biological mechanism of m^{6}A-related lncRNAs in LUAD.

In the clinic, the pathological stage is the decisive factor in the prognosis of LUAD.[31] However, LUAD patients at the same stage always have different clinical outcomes, suggesting that the present periodization systems in providing dependable predictions and reflecting the heterogeneity of LUAD are inaccurate.[32] Therefore, latent predictive and therapeutic biomarkers should be explored. The established m^{6}A-related lncRNA model provides a new method for prognostic prediction in LUAD patients. The results also provide insight for future studies on the process and mechanism of m^{6}A modification of lncRNAs. In our study, several methods were used to confirm this novel model, and we might choose the optimal model and utilize it evenly. We assumed that the prediction model was acceptable without the external data validation. We are also aware of some shortcomings and limitations in this study. External validation by other clinical datasets would be beneficial, and the biological mechanism of m^{6}A-related lncRNAs has not been fully elucidated. Thus, we will recollect clinical samples and expand the sample size. Additionally, we will attempt to validate the accuracy of this model via more external experiments to explore the role of lncRNAs and their interaction with m^{6}A-related genes in our following work.

In conclusion, our study provides clues for prognostic prediction in patients with LUAD and may help elucidate the process and mechanism of m^{6}A-regulated lncRNAs. In addition, the prediction model shows sensitivity in identifying LUAD patients who may respond well to immunotherapy.

MATERIALS AND METHODS

Acquisition of information of patients with LUAD

Using VarScan software, we obtained RNA sequence transcriptome data, relevant clinical information, and mutation data of LUAD patients from the TCGA (https://cancergenome.nih.gov/) database. To reduce statistical bias in this analysis, LUAD patients with missing OS values were excluded.

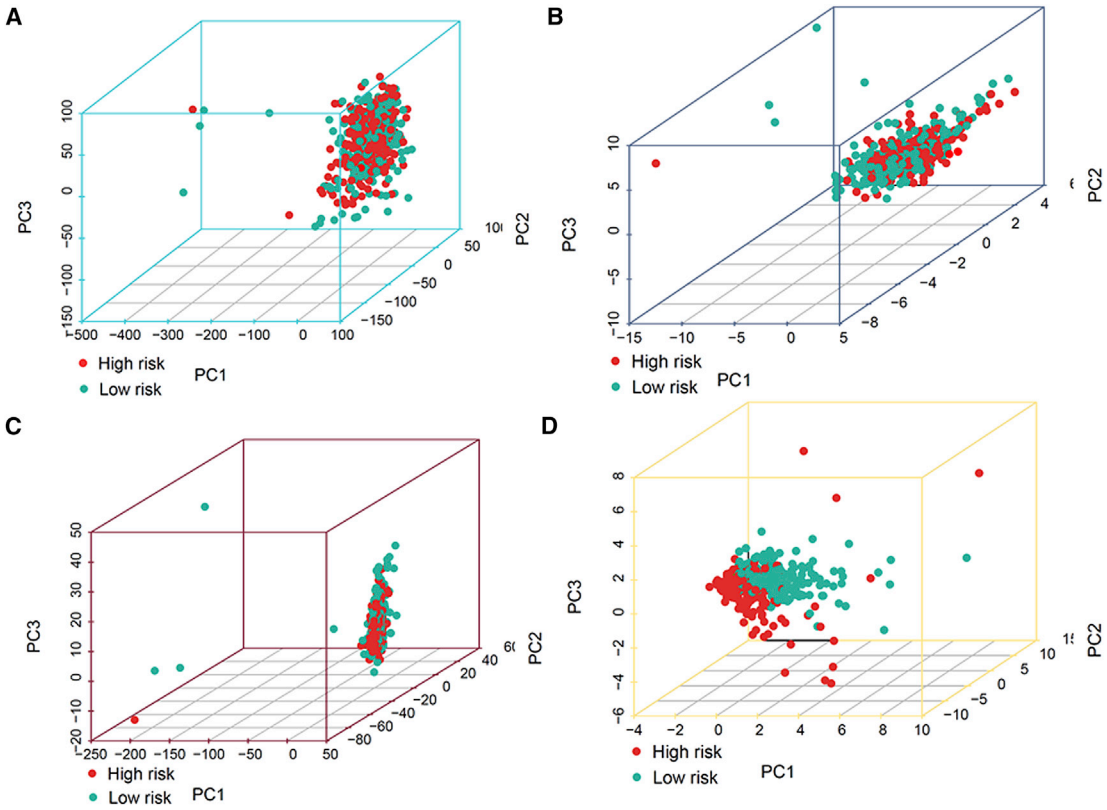

Figure 7. Principal component analysis between the high- and low-risk groups based on

(A) entire gene expression profiles, (B) 21 m^{6}A genes, (C) 12 m^{6}A-related lncRNAs, and (D) risk model based on the representation profiles of the 12 m^{6}A-related lncRNAs in the TCGA entire set

Selection of m^{6}A genes and m^{6}A-related lncRNAs

We obtained the profiles of lncRNAs and m^{6}A genes from the TCGA database. According to previous studies, the expression matrixes of 21 m^{6}A genes were retrieved from the TCGA, including the expression data of writers (METTL3, METTL14, METTL16, VIRMA, RBM15, RBM15B, ZC3H13, and WTAP), readers (IGF2BP1, IGF2BP2, IGF2BP3, YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, HNRNPA2B1, HNRNPC, and RBMX), and erasers (ALKBH5 and FTO).[9] We screened m^{6}A-related lncRNAs by Pearson's correlation analysis, and 1,149 m^{6}A-related lncRNAs were identified. The process used the criteria of |Pearson R| >0.3 and p <0.001.

Establishment and validation of the risk signature

The entire TCGA set was randomized as a training set and the testing set. The training set was utilized to construct an m^{6}A-related lncRNA model, and the entire set and testing set were applied to validate this established model. Table S1 shows the baseline characteristics of these two sets. No significant differences in clinical properties were observed between the two datasets ($p > 0.05$). Combined with LUAD survival information in TCGA, we screened the prognosis of m^{6}A-related lncRNAs from 1,149 m^{6}A-related lncRNAs in the TCGA dataset ($p < 0.05$), and univariate Cox regression was used in this study.[33] Using the R package glmnet to conduct LASSO Cox regression (using the penalty parameter estimated by 10-fold cross-validation), we found that 24 m^{6}A-related lncRNAs were distinctly related to the OS of LUAD patients from TCGA datasets.[34] Multifactor Cox regression was applied to analyze the 24 m^{6}A-related lncRNAs, and a 12-m^{6}A-related lncRNA risk model was ultimately established.[3] The following formula was used to calculate the risk score: Risk score = coef (lncRNA1) × expr (lncRNA1) + coef (lncRNA2) × expr (lncRNA2) + …… + coef (lncRNAn) × expr (lncRNAn), where coef indicates the coefficients, coef (lncRNAn) was the coefficient of lncRNAs correlated with survival, and expr (lncRNAn) was the expression of lncRNAs. According to the median risk score, subgroups including low- and high- risk groups were established.[35]

Functional analysis

We performed GO analysis to identify the differentially expressed genes. This process utilizes the R package clusterProfiler. The analysis threshold was determined by the p value, and p <0.05 indicates that the functional comment is significantly enriched.[16,36]

Exploration of the model in the immunotherapeutic treatment

We used the R package maftools to evaluate and sum the mutation data. The TMB was measured according to tumor-specific mutated

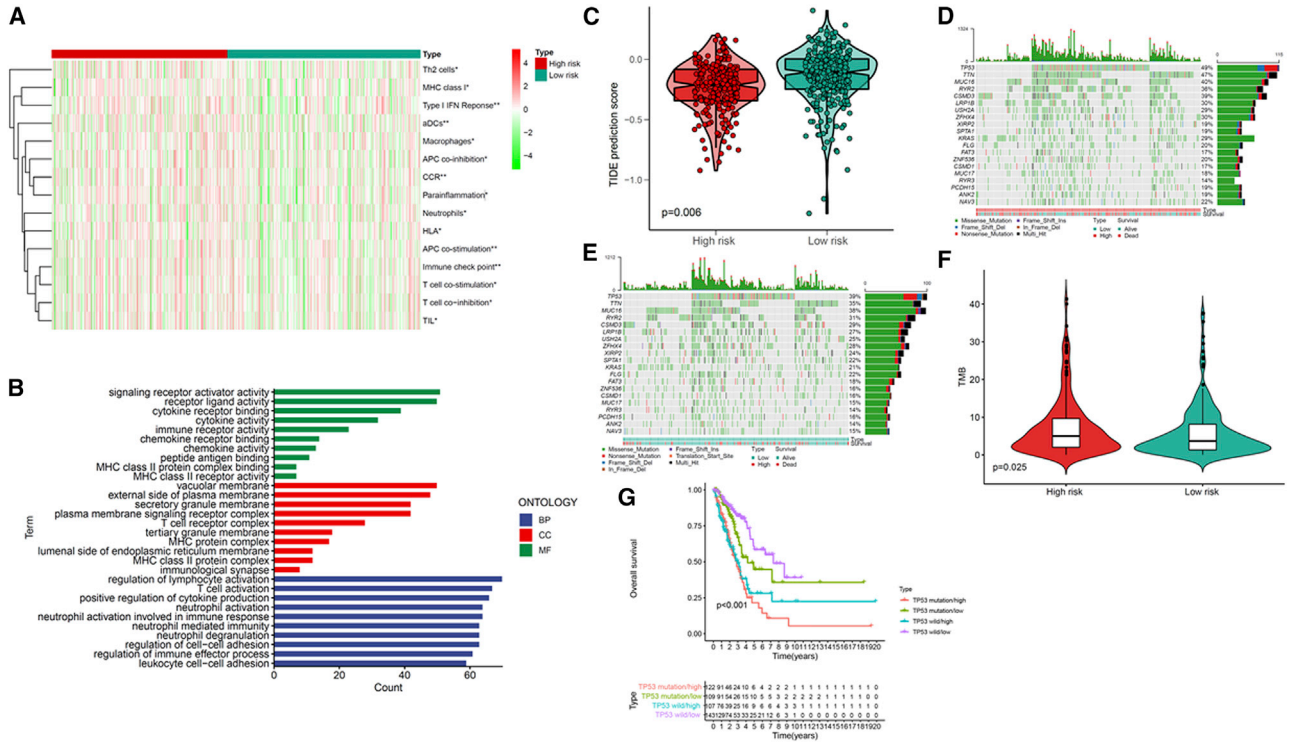

Figure 8. Estimation of the tumor immune microenvironment and cancer immunotherapy response using the m^{6}A-related lncRNA model in the TCGA entire set

(A) The indicated standards of the immunity index for each patient. (B) GO enrichment analysis. (C) TIDE prediction difference in the high- and low-risk patients. (D and E) Waterfall plot displays mutation information of the genes with high mutation frequencies in the high-risk group (D) and low-risk group (E). (F) TMB difference in the high- and low-risk patients. (G) Kaplan-Meier curve analysis of OS is shown for patients classified according to the TP53 mutation status and m^{6}A-related lncRNA model.

genes.[37] We used the TIDE algorithm to predict the likelihood of the immunotherapeutic response.[33]

PCA and Kaplan-Meier survival analysis

PCA was used for effective dimensionality reduction, model identification, and grouping visualization of high-dimensional data of the entire gene expression profiles, 21 m^{6}A genes, 12 m^{6}A-related lncRNAs, and risk model according to the expression patterns of the 12 m^{6}A-related lncRNAs.[38] We used Kaplan-Meier survival analysis to appraise diversities in the OS between the high-risk and low-risk groups. The R packages survMiner and survival were tools to enable this process.[3]

Exploration of potential compounds targeting m^{6}A-related lncRNA model in clinical treatment

To obtain potential compounds in the clinic for LUAD treatment, we calculated the IC_{50} of compounds obtained from the GDSC website in the TCGA project of the LUAD dataset. The R package pRRophetic was used to predict the IC_{50} of compounds obtained from the GDSC website in patients with LUAD.

Independence of the m^{6}A-related lncRNA model

Multivariate and univariate Cox regression analyses were conducted to test whether the prognostic pattern was an independent variable considering other clinical characteristics (gender, age, TNM stage, T stage, N stage, M stage, and tobacco smoking history) in the patients with LUAD.[16]

Establishing and proving a predictive nomogram

The predictive ability of the nomogram and other predictors (age, gender, risk score, TNM stage, T stage, N stage, M stage, and tobacco smoking history) for the 1-, 3-, and 5-year OS was set up. Correction curves based on the Hosmer-Lemeshow test were applied to illustrate the uniformity between the practical outcome and model prediction outcome.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2021.04.003.

ACKNOWLEDGMENTS

This study was supported by grants from the National Natural Science Foundation of China (81672640); the Grant for Key Disciplinary Project of Clinical Medicine under the Guangdong High-level University Development Program; the 2020 Li Ka Shing Foundation Cross-Disciplinary Research Grant (2020LKSFG04A and 2020LKSFG10A); the Dengfeng Project for the construction

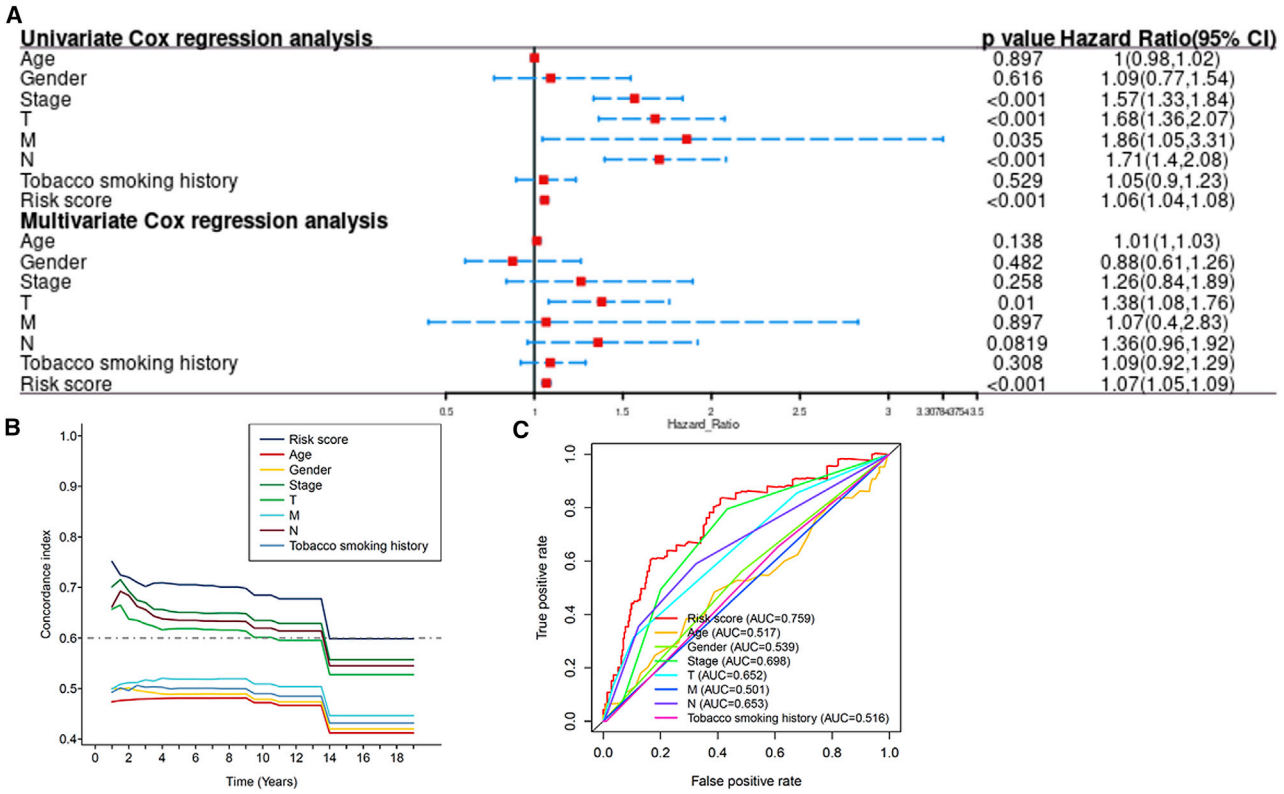

Figure 9. Assessment of the prognostic risk model of the m^{6}A-related lncRNAs and clinical features in LUAD in the TCGA entire set

(A) Univariate and multivariate analyses of the clinical characteristics and risk score with the OS. (B) Concordance indexes of the risk score and clinical characteristics. (C) ROC curves of the clinical characteristics and risk score.

of high-level hospitals in Guangdong Province–The First Affiliated Hospital of Shantou University Medical College Supporting Funding (2019-70); the Guangdong Basic and Applied Basic Research Foundation (2021A1515010137, 2020A1515011519); and the Medical Science and Technology Research Foundation of Guangdong Province (A2021409, A2020430).

AUTHOR CONTRIBUTIONS

F.X., X.H., and Y.L. designed the study, analyzed data, and wrote the manuscript. F.X., L.L., and Y.C. provided funding acquisition. L.L. and Y.C. supervised the research, analyzed data, and wrote the manuscript. All authors read and approved the final submitted manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Cao, M., Li, H., Sun, D., and Chen, W. (2020). Cancer burden of major cancers in China: A need for sustainable actions. Cancer Commun. (Lond.) *40*, 205–210.

2. Ferlay, J., Colombet, M., Soerjomataram, I., Dyba, T., Randi, G., Bettio, M., Gavin, A., Visser, O., and Bray, F. (2018). Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. Eur. J. Cancer *103*, 356–387.

3. Xu, F., He, L., Zhan, X., Chen, J., Xu, H., Huang, X., Li, Y., Zheng, X., Lin, L., and Chen, Y. (2020). DNA methylation-based lung adenocarcinoma subtypes can predict prognosis, recurrence, and immunotherapeutic implications. Aging (Albany NY) *12*, 25275–25293.

4. Zhang, C., Zhang, J., Xu, F.P., Wang, Y.G., Xie, Z., Su, J., Dong, S., Nie, Q., Shao, Y., Zhou, Q., et al. (2019). Genomic Landscape and Immune Microenvironment Features of Preinvasive and Early Invasive Lung Adenocarcinoma. J. Thorac. Oncol. *14*, 1912–1923.

5. Jurisic, V., Vukovic, V., Obradovic, J., Gulyaeva, L.F., Kushlinskii, N.E., and Djordjević, N. (2020). *EGFR* Polymorphism and Survival of NSCLC Patients Treated with TKIs: A Systematic Review and Meta-Analysis. J. Oncol. *2020*, 1973241.

6. Jiao, G., and Wang, B. (2016). NK Cell Subtypes as Regulators of Autoimmune Liver Disease. Gastroenterol. Res. Pract. *2016*, 6903496.

7. Sun, T., Wu, R., and Ming, L. (2019). The role of m^{6}A RNA methylation in cancer. Biomed. Pharmacother. *112*, 108613.

8. Liu, Z.X., Li, L.M., Sun, H.L., and Liu, S.M. (2018). Link Between m^{6}A Modification and Cancers. Front. Bioeng. Biotechnol. *6*, 89.

9. Tu, Z., Wu, L., Wang, P., Hu, Q., Tao, C., Li, K., Huang, K., and Zhu, X. (2020). N^6-Methylandenosine-Related lncRNAs Are Potential Biomarkers for Predicting the Overall Survival of Lower-Grade Glioma Patients. Front. Cell Dev. Biol. *8*, 642.

10. Zhou, K.I., Parisien, M., Dai, Q., Liu, N., Diatchenko, L., Sachleben, J.R., and Pan, T. (2016). N(6)-Methyladenosine Modification in a Long Noncoding RNA Hairpin Predisposes Its Conformation to Protein Binding. J. Mol. Biol. *428* (5 Pt A), 822–833.

11. Song, H., Li, H., Ding, X., Li, M., Shen, H., Li, Y., Zhang, X., and Xing, L. (2020). Long non-coding RNA FEZF1-AS1 facilitates non-small cell lung cancer progression via the ITGA11/miR-516b-5p axis. Int. J. Oncol. *57*, 1333–1347.

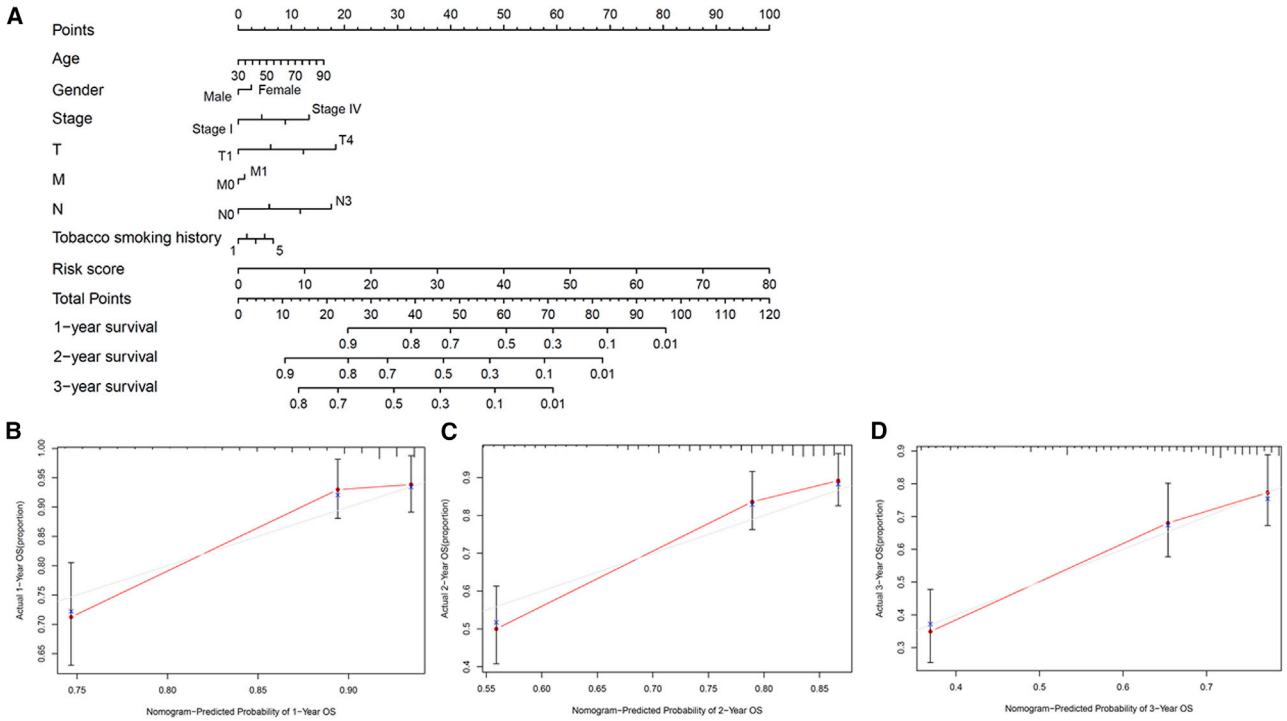

Figure 10. Construction and evaluation of a prognostic nomogram

(A) The nomogram predicts the probability of the 1-, 2-, and 3-year OS. (B–D) The calibration plot of the nomogram predicts the probability of the 1-, 2-, and 3-year OS.

12. Xue, L., Li, J., Lin, Y., Liu, D., Yang, Q., Jian, J., and Peng, J. (2021). m^6 A transferase METTL3-induced lncRNA ABHD11-AS1 promotes the Warburg effect of non-small-cell lung cancer. J. Cell. Physiol. *236*, 2649–2658.

13. Zhang, Y., Liu, X., Liu, L., Li, J., Hu, Q., and Sun, R. (2020). Expression and Prognostic Significance of m6A-Related Genes in Lung Adenocarcinoma. Med. Sci. Monit. *26*, e919644.

14. Li, Y., Gu, J., Xu, F.K., Zhu, Q.L., Chen, Y.W., Ge, D., and Lu, C.L. (2020). Molecular characterization, biological function, tumor microenvironment association and clinical significance of m6A regulators in lung adenocarcinoma. Brief Bioinform. Published online October 1, 2020. https://doi.org/10.1093/bib/bbaa225.

15. Bade, B.C., and Dela Cruz, C.S. (2020). Lung Cancer 2020: Epidemiology, Etiology, and Prevention. Clin. Chest Med. *41*, 1–24.

16. Zhao, X., Liu, X., and Cui, L. (2020). Development of a five-protein signature for predicting the prognosis of head and neck squamous cell carcinoma. Aging (Albany NY) *12*, 19740–19755.

17. Dong, H.X., Wang, R., Jin, X.Y., Zeng, J., and Pan, J. (2018). LncRNA DGCR5 promotes lung adenocarcinoma (LUAD) progression via inhibiting hsa-mir-22-3p. J. Cell. Physiol. *233*, 4126–4136.

18. Tian, Y., Yu, M., Sun, L., Liu, L., Wang, J., Hui, K., Nan, Q., Nie, X., Ren, Y., and Ren, X. (2020). Distinct Patterns of mRNA and lncRNA Expression Differences Between Lung Squamous Cell Carcinoma and Adenocarcinoma. J. Comput. Biol. *27*, 1067–1078.

19. Song, Q., Shang, J., Yang, Z., Zhang, L., Zhang, C., Chen, J., and Wu, X. (2019). Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma. J. Transl. Med. *17*, 70.

20. Zhang, H., Guo, L., and Chen, J. (2020). Rationale for Lung Adenocarcinoma Prevention and Drug Development Based on Molecular Biology During Carcinogenesis. OncoTargets Ther. *13*, 3085–3091.

21. Ghafouri-Fard, S., Shoorei, H., Branicki, W., and Taheri, M. (2020). Non-coding RNA profile in lung cancer. Exp. Mol. Pathol. *114*, 104411.

22. Loewen, G., Jayawickramarajah, J., Zhuo, Y., and Shan, B. (2014). Functions of lncRNA HOTAIR in lung cancer. J. Hematol. Oncol. *7*, 90.

23. Fu, L., Wang, H., Wei, D., Wang, B., Zhang, C., Zhu, T., Ma, Z., Li, Z., Wu, Y., and Yu, G. (2020). The value of CEP55 gene as a diagnostic biomarker and independent prognostic factor in LUAD and LUSC. PLoS ONE *15*, e0233283.

24. Fazi, F., and Fatica, A. (2019). Interplay Between N^6-Methyladenosine (m^6A) and Non-coding RNAs in Cell Development and Cancer. Front. Cell Dev. Biol. *7*, 116.

25. Luo, Z.B., Lai, G.E., Jiang, T., Cao, C.L., Peng, T., and Liu, F.E. (2020). A Competing Endogenous RNA Network Reveals Novel lncRNA, miRNA and mRNA Biomarkers With Diagnostic and Prognostic Value for Early Breast Cancer. Technol. Cancer Res. Treat. *19*, 1533033820983293.

26. Miao, W., Li, N., Gu, B., Yi, G., Su, Z., and Cheng, H. (2020). LncRNA DLGAP1-AS2 modulates glioma development by up-regulating YAP1 expression. J. Biochem. *167*, 411–418.

27. Li, J.P., Li, R., Liu, X., Huo, C., Liu, T.T., Yao, J., and Qu, Y.Q. (2020). A Seven Immune-Related lncRNAs Model to Increase the Predicted Value of Lung Adenocarcinoma. Front. Oncol. *10*, 560779.

28. Allgäuer, M., Budczies, J., Christopoulos, P., Endris, V., Lier, A., Rempel, E., Volckmar, A.L., Kirchner, M., von Winterfeld, M., Leichsenring, J., et al. (2018). Implementing tumor mutational burden (TMB) analysis in routine diagnostics-a primer for molecular pathologists and clinicians. Transl. Lung Cancer Res. *7*, 703–715.

29. Topalian, S.L., Taube, J.M., Anders, R.A., and Pardoll, D.M. (2016). Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. Nat. Rev. Cancer *16*, 275–287.

30. Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., Li, Z., Traugh, N., Bu, X., Li, B., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat. Med. *24*, 1550–1558.

31. Jurišić, V., Obradovic, J., Pavlović, S., and Djordjevic, N. (2018). Epidermal Growth Factor Receptor Gene in Non-Small-Cell Lung Cancer: The Importance of Promoter Polymorphism Investigation. Anal. Cell. Pathol. (Amst.) *2018*, 6192187.

32. Rafei, H., El-Bahesh, E., Finianos, A., Nassereddine, S., and Tabbara, I. (2017). Immune-based Therapies for Non-small Cell Lung Cancer. Anticancer Res. *37*, 377–387.

33. Xu, F., Zhan, X., Zheng, X., Xu, H., Li, Y., Huang, X., Lin, L., and Chen, Y. (2020). A signature of immune-related gene pairs predicts oncologic outcomes and response to immunotherapy in lung adenocarcinoma. Genomics *112*, 4675–4683.

34. Xu, F., Lin, H., He, P., He, L., Chen, J., Lin, L., and Chen, Y. (2020). A *TP53*-associated gene signature for prediction of prognosis and therapeutic responses in lung squamous cell carcinoma. OncoImmunology *9*, 1731943.

35. Hong, W., Liang, L., Gu, Y., Qi, Z., Qiu, H., Yang, X., Zeng, W., Ma, L., and Xie, J. (2020). Immune-Related lncRNA to Construct Novel Signature and Predict the Immune Landscape of Human Hepatocellular Carcinoma. Mol. Ther. Nucleic Acids *22*, 937–947.

36. Gong, K., Guo, G., Beckley, N., Zhang, Y., Yang, X., Sharma, M., and Habib, A.A. (2021). Tumor necrosis factor in lung cancer: Complex roles in biology and resistance to treatment. Neoplasia *23*, 189–196.

37. Wu, Z., Wang, M., Liu, Q., Liu, Y., Zhu, K., Chen, L., Guo, H., Li, Y., and Shi, B. (2020). Identification of gene expression profiles and immune cell infiltration signatures between low and high tumor mutation burden groups in bladder cancer. Int. J. Med. Sci. *17*, 89–96.

38. Li, X., Li, Y., Yu, X., and Jin, F. (2020). Identification and validation of stemness-related lncRNA prognostic signature for breast cancer. J. Transl. Med. *18*, 331.

OMTN, Volume *24*

Supplemental information

m^6A-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with LUAD

Feng Xu, Xiaoling Huang, Yangyi Li, Yongsong Chen, and Ling Lin

Figure S1

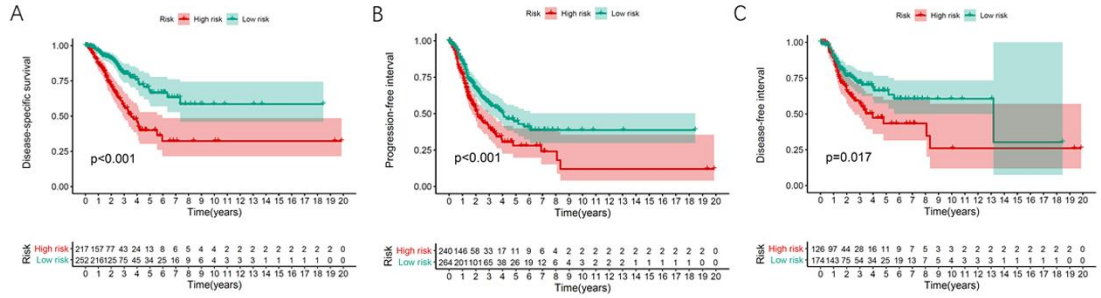

Figure S1. (A-C) Kaplan-Meier survival curves of the relative DSS, PFI, and DFI between the high- and low-risk groups.

Figure S2

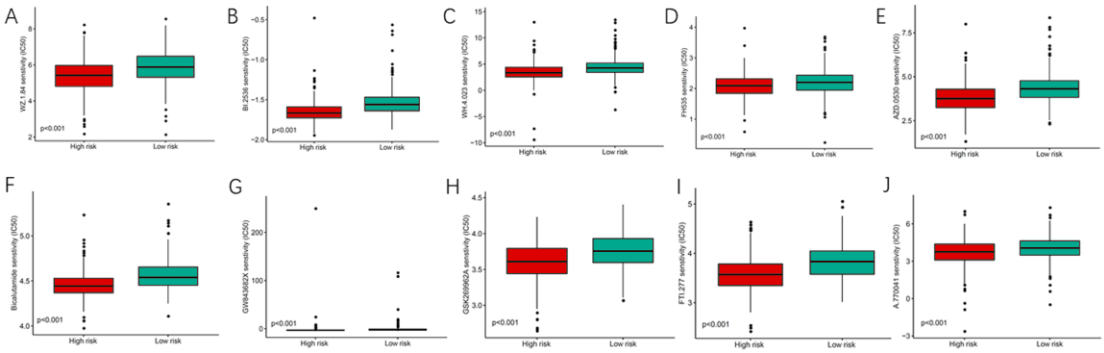

Figure S2. (A-J) Identification of novel candidate compounds targeting the m6A-related lncRNA model

Table S1. Summary of patients' characteristics

	Entire set		Training set		Testing set		
Characteristics	Number	Percentage	Number	Percentage	Number	Percentage	p-Value
Age							
<60 year	136	26.98	89	29.28%	47	23.5	0.2025
⩾60 year	358	71.03	210	69.08	148	74	
Not available	10	1.98	5	1.64	5	2.5	
Gender							
Female	270	53.57	165	54.28	105	52.5	0.7642
Male	234	46.43	139	45.72	95	47.5	
Stage							
Stage I-II	389	77.18	236	77.63	153	76.5	0.8231
Stage III-IV	107	21.23	63	20.72	44	22	
Not available	8	1.59	5	1.64	3	1.5	
T							
T1-2	437	86.71	269	88.49	168	84	0.1648
T3-4	64	12.7	33	10.86	31	15.5	
Not available	3	0.6	2	0.66	1	0.5	
M							
M0	335	66.47	203	66.78	132	66	0.8093
M1	25	4.96	14	4.61	11	5.5	
Not available	144	28.57	87	28.62	57	28.5	
N							
N0	325	64.48	197	64.8	128	64	0.9517
N1-3	167	33.13	100	32.89	67	33.5	
Not available	12	2.38	7	2.3	5	2.5	