Am J Transl Res 2022;14(12):8457-8472
www.ajtr.org /ISSN:1943-8141/AJTR0143582

Original Article

Analysis of M6A associated lncRNAs in prognosis and immune response of NSCLC patients

Jing Lin[1*], Xiao-Ling Gu[1*], Chu-Ling Li[1], Zi-Mu Wang[1], Zhao-Feng Wang[1], Ran-Pu Wu[1], Yong Song[1], Ying Wu[2,3], Hong-Bing Liu[1]

[1]Department of Respiratory Medicine, Jinling Hospital, Medical School of Nanjing University, Nanjing, China; [2]Affiliated Hospital of Nanjing University of Chinese Medicine, Jiangsu Province Hospital of Chinese Medicine, Nanjing, China; [3]First College of Clinical Medicine, Nanjing University of Chinese Medicine, Nanjing, China. [*]Equal contributors.

Received April 17, 2022; Accepted August 13, 2022; Epub December 15, 2022; Published December 30, 2022

Abstract: Distinguishing between N6-methyladenosine (m6A)-associated long noncoding RNAs (lncRNAs) is crucial in non-small-cell lung cancer (NSCLC) patients. In this research, the prognosis and immunotherapeutic response of lncRNAs and m6A in NSCLC were examined. lncRNAs related to m6A were identified using co-expression analyses, and their prognostic impact on patients with NSCLC was assessed using univariate Cox regression analysis. Sixty-three m6A-associated lncRNAs were determined as prognostic lncRNAs, and on this basis, 25 m6A-associated lncRNAs were screened by least absolute shrinkage and selection operator (lasso) Cox regression. Multivariable Cox analysis obtained 14 m6A-associated lncRNAs for the construction of risk model. The NSCLC patients were grouped into different risk subgroups in accordance with the median of the risk fraction in each data, and we evaluated the differences of potential immunotherapeutic characteristics and drug sensitivity prediction between the two subgroups. By using this model to recombine patients, they can be effectively distinguished in terms of the immunotherapy response. Furthermore, candidate compounds for the differentiation of NSCLC subtypes were identified. The model based on 14 m6A-associated lncRNAs is a promising prognostic biomarker, which may help to predict the efficacy of immunotherapy in NSCLC patients and provide a theoretical basis for improving the outcome of patients.

Keywords: M6A, lncRNAs, prognosis, NSCLC, immunotherapy response

Introduction

Pulmonary carcinoma is a malignant tumor originating from the lung bronchial mucosal or glandular cells, and its incidence and mortality rates have increased rapidly. Pulmonary carcinoma is one of the largest malignant tumors threatening human life and health [1]. NSCLC is one of the histological types of pulmonary carcinoma, accounting for about 85% of all lung carcinoma cases. Most patients with NSCLC are diagnosed with local advanced stage or metastasis, and the overall prognosis is very poor. The outcome of early NSCLC is better than that of late-stage NSCLC and it can usually be treated using surgical resection or radiotherapy. In addition to imaging-based examination, other methods are used for the early detection of NSCLC [2]. At present, some studies have indicated that the identification and utilization of molecular biomarkers provides new ideas for the prediction of curative effect and prognosis of patients.

Modification of RNA methylation is responsible for approximately 60% of all RNA modifications. N6-methyladenosine (m6A) is a prominent decoration in noncoding RNAs (ncRNAs) and messenger RNAs (mRNAs). It is one of the several RNA modifications currently known. m6A modification is reversible and exists in nearly every type of RNA, It is closely related to many disease processes, such as oncogenesis. m6A methylation modifications include common participation of methyltransferase, demethylase, and methylated reading proteins [3]. The most recent study on m6A in cancer has indicated that it plays a definite role in various cancers. For instance, METTL3 is upregulated in lung adenocarcinoma (LUAD) and has an influence

on promoting the invasion of human pulmonary carcinoma cells [4]. YTHDF1 is amplified in NSCLC to promote cell proliferation [5]. FEZF1-AS1 affected by m6A modification adjusts the ITGA11/miR-516b-5p axis and eventually increases its expression in NSCLC [6].

This study examined the prognostic function of m6A-associated lncRNAs to aid in the exploration of biomarkers related to the prognosis and immune response of patients with NSCLC.

Materials and methods

Information acquisition and sorting of patients with NSCLC

We obtained the data on mutation, RNA sequence transcription group, and the related clinical information from previous NSCLC patients using the Cancer Genome Atlas (TCGA). As a retrospective study, to decrease the statistical deviation in this study, NSCLC patients without accurate survival data were eliminated from the study. We obtained 1,037 tumor tissue specimens and 108 normal tissue specimens from the TCGA database. We used Perl software to organize the transcriptome data and converted the IDs by using the mRNA matrix and its corresponding script. Based on past research, the expression matrices of 23 m6A genes were obtained from TCGA, which consisted of authors (VIRMA, RBM15, RBM-15B, METTL3, METTL14, METTL16, ZC3H13, and WTAP), readers (IGFBP1, IGFBP2, IGFBP3, YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, RBMX, HNRNPC, HNRNPA2B1, LRPPRC, and FMR1), and erasers (ALKBH5 and FTO).

Appraisal of m6A-associated lncRNAs

We used the Perl program to structure a human configuration file and the gene expression matrix, and obtained mRNA and lncRNAs gene expression profiles by running the relevant script program. Based on the m6A-related gene types “writer”, “reader”, and “eraser” and the relevant gene names, the M6A-associated gene expression profile were obtained by applying the limma software package in the R software. Co-expression analyses were performed to define the correlation between the M6A-associated gene expression and lncRNAs. We obtained the m6A-associated lncRNAs through Pearson’s correlational analysis ($P < 0.001$ and $|Pearson\ r| > 0.4$) to determine 1651 M6A-associated lncRNAs. In addition, the co-expression network diagram was drawn by the dplyr software package, ggplot2 software package, and ggalluval software package. The expression data of M6A-associated lncRNAs was combined with the clinical survival data through the limma software package. The relevance between prognosis M6A-associated lncRNAs and M6A-associated genes was determined by using the limma software package, reshape2 software package, tidyverse software package, and ggplot2 software package, and the correlation heat map was drawn.

Establishment of m6A lncRNAs-related prognostic model

The whole set from TCGA was randomly split into the training group (50%) and the test group (50%). We used the training set to build the M6A-associated lncRNA model and to apply the entire training and test sets to the established model. In this study, univariate analysis was performed to screen the prognosis of M6A-associated lncRNAs from 1651 M6A-associated lncRNAs in the TCGA data set ($P < 0.05$). We used the R software package glmnet for lasso Cox regression, we discovered that 63 M6A-associated lncRNAs were obviously related with Overall Survival (OS) in NSCLC patients of the TCGA data set. In addition, we used multivariate Cox regression to analyze 14 M6A-associated lncRNAs, and eventually 14 M6A-associated lncRNAs risk models were set up. Among them, 6 lncRNAs (AL606489.1, AC084117.1, ‘NR2F2-AS1’, AC022165.1, ‘ZNF8-ERVK3-1’, AC020915.2) were independent adverse prognostic factors for OS, and the rest (AC027117.1, AC008114.1, ‘MAPKAPK5-AS1’, AL161757.2, ‘PDC-AS1’, ‘LMO7DN-IT1’, AC090948.1, AC005479.2) were independent favorable prognostic factors for OS. We used the following formula to calculate the risk fraction:

$$\text{risk fraction} = \sum_{n=x}^{n} \text{coef}(x)\,\text{expr}(x)$$

Where, coef (x) indicates the regression coefficient of lncRNAs associated with survival, and expr (x) indicates the expression of lnc-

RNAs. On the basis of the median value of risk fraction, they were split into different risk subgroups.

Statistical analysis

We obtained the survival curve through the survminer package and survival package in order to compare the difference of OS between the low-risk and high-risk groups. The corresponding subject receiver operating characteristic (ROC) curve was obtained through timeROC software package, and the concordance index (C-index) curve was obtained through the Survival software package, rms software, and pec software package to assess the correctness of our model in forecasting patient survival. We also used the pheatmap software package to obtain the risk curve. Multivariate and univariate independent prognostic analyses were implemented through the survival software package to assess whether our model is independent of other clinical prognostic factors that may affect the outcome of patients. We performed the validation of the clinical groups to examine whether our model was applicable to distinct clinical groups. Clinical and risk relevance analyses were performed to differentiate between the associated low-risk and high-risk M6A-associated lncRNAs as well as to clarify the relevance between clinical features and our prognostic risk model. Principal Component Analysis (PCA) was employed to identify and visualize the m6A gene, the total gene expression profile, M6A-associated lncRNA, and the model lncRNA data. For the sake of probing the potential mechanism of distinct prognosis between the different risk subgroups, the "limma" R package was used to define the discrepancy expressed genes (DEGs) among the subgroups (FDR < 0.05 and |log2fc| > 1). Further functional enrichment analyses of DEGs were unfolded to evaluate the differences in the biological process (BP), cellular components (CC), and molecular function (MF) between the low-risk and high-risk subgroups.

Exploration of the immunotherapy model

R software package was applied to assess and calculate the mutation data. The differences between high mutation load and patient survival and between low mutation load and patient survival were evaluated. We employed the tide algorithm to predict the efficacy of immunotherapy.

Analysis of potential compounds of target M6A-associated lncRNAs model

To obtain the getPotential pharmaceutical compound for the therapy of NSCLC clinically, we calculated the half maximal inhibitory concentration (IC50) of NSCLC database compounds in the TCGA project, which was acquired from the Genomics of Drug Sensitivity in Cancer (GDSC) website. The R software package pRRophetic was used to forecast the IC50 of compounds for NSCLC patients.

Construction of nomogram

The establishment of the column diagram demonstrated the predictive ability of OS for 1, 3, and 5 years. Line diagrams included several predictors (such risk fraction, T-stage, N-stage, TNM stage, gender, and age). The calibration curve was applied to illustrate the difference between the survival rate predicted by the model and the actual results.

Results

Assessment of m6A-associated lncRNAs in NSCLC patients

Figure 1 shows the flow diagram of risk model construction and subsequent analysis of m6A-associated lncRNAs in NSCLC patients. The expression matrices of 23 m6A genes and 14056 lncRNAs were obtained from The Cancer Genome Atlas (TCGA) data pool. According to Pearson coefficient ($P < 0.001$ and |Pearson r| > 0.4), the m6As significantly related to lncRNAs were acquired, and the Sankey diagram of m6A-lncRNA was drawn (**Figure 2A**). Altogether, 1651 m6A-associated lncRNAs were assessed. **Figure 2B** indicates the correlation between 14 lncRNAs and m6A-associated genes in TCGA data pool.

Structure of the risk model of m6A-associated lncRNAs in NSCLC patients

We used univariate Cox regression analysis to select m6A-associated prognostic lncRNAs from 1,651 m6A associated lncRNAs. Altoge-

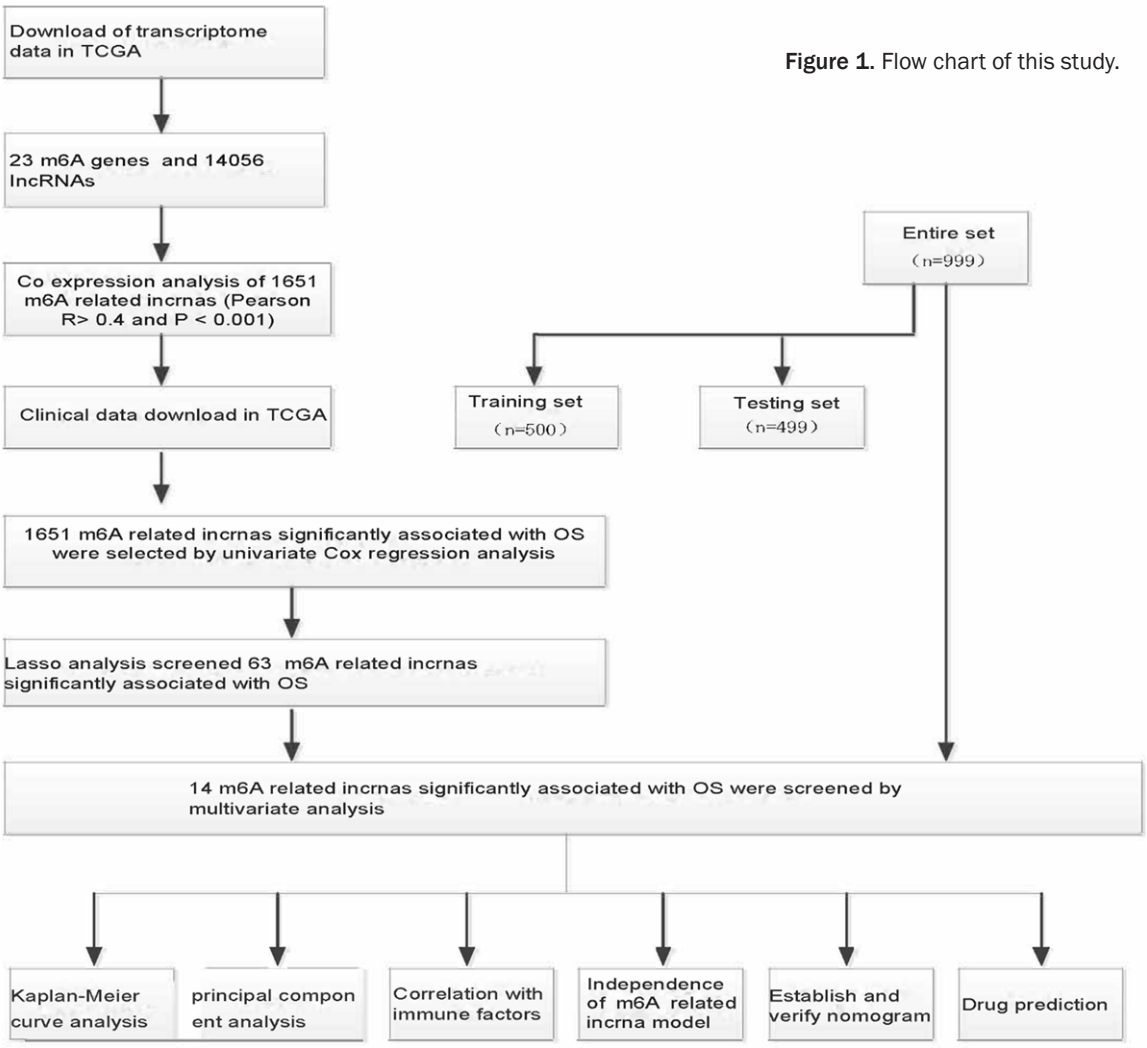


Figure 1. Flow chart of this study.

ther, 63 m6A associated lncRNAs were found to be remarkably associated with OS ($P < 0.05$) (**Figure 3A**). Lasso Cox method was analyzed for the 63 m6A-associated prognostic lncRNAs in the TCGA queue, and 25 m6A-associated lncRNAs were obtained (**Figure 3B** and **3C**). Next, multivariate Cox analysis was applied to obtain 14 m6A-associated lncRNAs for constructing risk models in NSCLC patients (**Table 1**). On the basis of the median value of the prognostic risk level, the NSCLC patients were divided into different risk subgroups. **Figure 4A** depicts the risk-level distribution among different risk subgroups in the training group, and **Figure 4B** presents the survival time and survival status of the patients in the two distinct risk subgroups. **Figure 4C** depicts the expression standards and risks of the 14 m6A-associated lncRNAs per patient. The Kaplan-Meier survival curve showed that the lower the risk score, the better the survival prognosis ($P < 0.001$) (**Figure 4D**).

To check the prognostic ability of the model, a unified method was applied to compute the risk fraction of every patient in the test set as well as the whole test set. **Figure 5** describes the scatter of the risk levels, survival status, and the description of m6A-associated lncRNAs in the test set (**Figure 5A-C**) and the whole test set (**Figure 5E-G**). Kaplan-Meier survival analysis of the test set and the whole test set indicated that there was no difference in the results of the TCGA training set. Patients with a lower risk fraction had better OS than those with a higher risk fraction (**Figure 5D** and **5H**).

The OS differences between different risk subgroups in TCGA group were analyzed with regard to the general clinicopathological char-

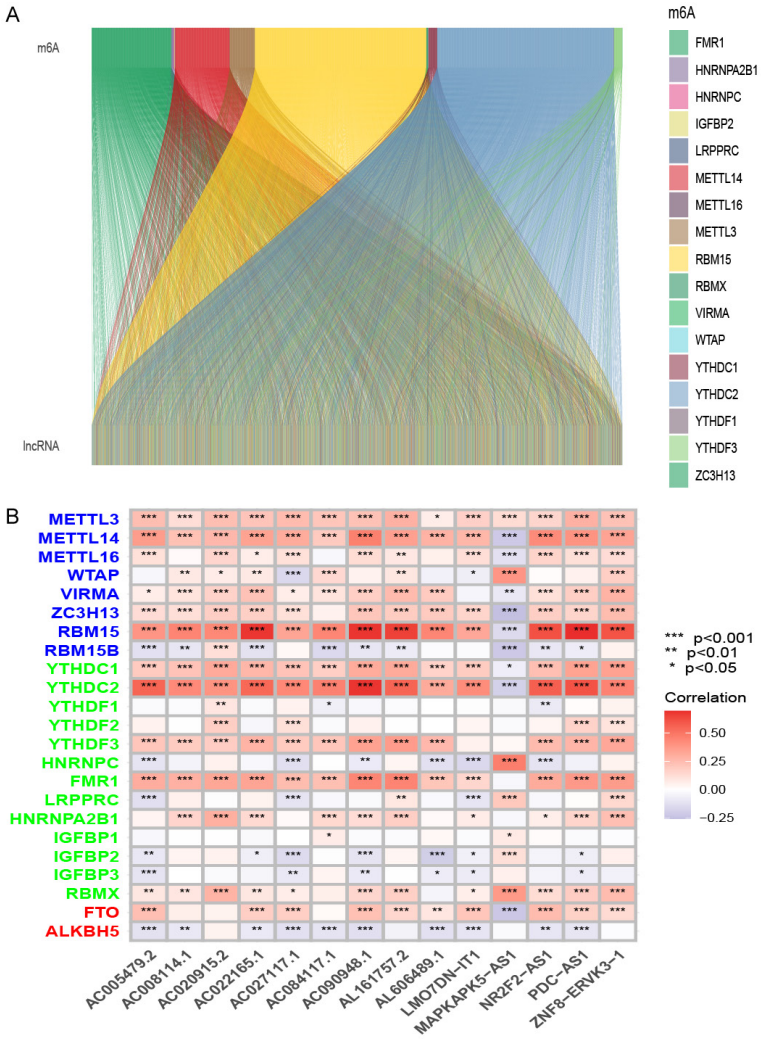


Figure 2. A. Co expression network of 17 m6A genes and m6A related lncRNAs. B. Heatmap of the correlation between 23 m6A genes and 14 prognostic m6A related lncRNAs.

acteristics. Compared with the high-risk group, the subgroup with lower risk had better OS upon dividing the subgroups based on age, sex, and tumor stage (**Figure 6**).

PCA verified the grouping capability of the m6A-associated lncRNA model

PCA was used to test the differences between different risk subgroups according to the risk model of the total gene expression profile, 23 m6A genes, 14 m6A-associated lncRNAs, and risk model of 14 m6A-associated lncRNAs (**Figure 7A-D**). **Figure 7A-C** show that the distribution of the high-risk and low-risk groups was comparatively dispersed. But, differences in the distribution between different risk subgroups were evident from the results of our model (**Figure 7D**). These results indicate that the prediction markers between different groups are clear.

Analysis of immunotherapy and tumor mutation load in different risk groups using the m6A-associated lncRNA model

Using the m6A-associated lncRNA model from 1037 NSCLC samples, the expression of several immunocytes in NSCLC were analyzed. There was a dramatic difference in the expression of immune indexes between different subgroups (**Figure 8A**). Gene Ontology (GO) enrichment analysis was used to study the potential molecular mechanism of the m6A model, which described the possible molecular functions of gene products, the cellular environment, and the biological processes involved (**Figure 8B**). Next, the relationship between m6A-associated lncRNA model and immunotherapy was studied. Tumor immune dysfunction and rejection (TIDE) score was applied to evaluate tumor immune escape. A higher tide score means a higher possibility of immune escape and a lower success rate of immunotherapy. In our research, the high-risk group has a lower score and is more likely to respond to immunotherapy (**Figure 8C**). Maftools package was applied to analyze and summarize the mutation data. The mutations were distinguished according to the predictors of variation effect. **Figure 8D** and **8E** illustrate the top 20 genes with mutation frequency between the low-risk and high-risk groups of LUAD. **Figure 8F** and **8G** show the top 20 genes with mutation frequency between lung squamous cell carcinoma (LUSC) low-risk and high-risk subgroups. The tumour mutational burden (TMB) score was then calculated according to TGCA somatic mutation data. The tumor mutation load in the high-risk group of patients with LUAD was higher than that in the low-risk group,

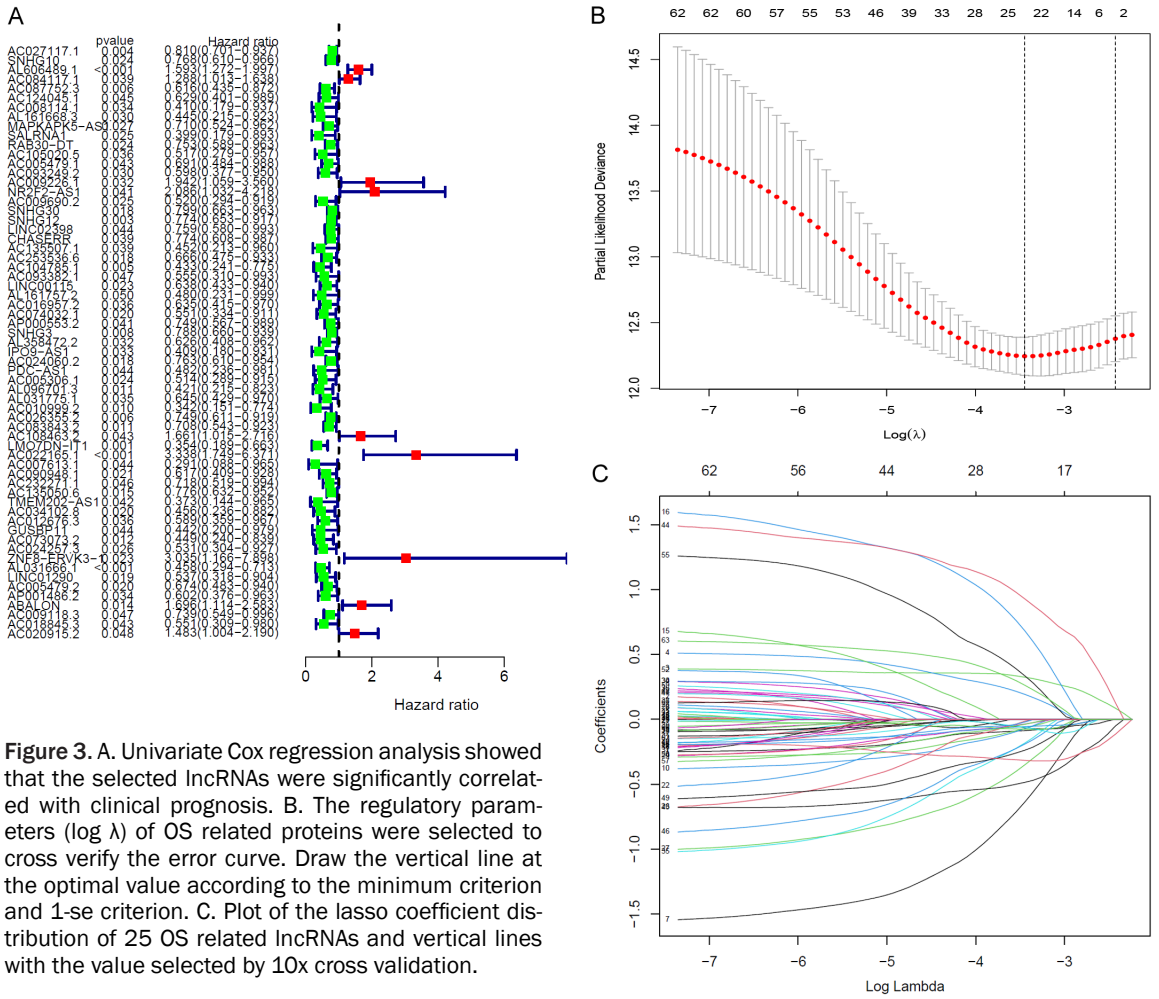


Figure 3. A. Univariate Cox regression analysis showed that the selected lncRNAs were significantly correlated with clinical prognosis. B. The regulatory parameters (log λ) of OS related proteins were selected to cross verify the error curve. Draw the vertical line at the optimal value according to the minimum criterion and 1-se criterion. C. Plot of the lasso coefficient distribution of 25 OS related lncRNAs and vertical lines with the value selected by 10x cross validation.

Table 1. Multivariate Cox regression analysis showed 14 independent prognostic lncRNAs

Id	coef
AC027117.1	-0.194147157644064
AL606489.1	0.455587975835069
AC084117.1	0.441591501889677
AC008114.1	-1.64879373061543
'MAPKAPK5-AS1'	-0.324057213978301
'NR2F2-AS1'	1.31330855789548
AL161757.2	-1.22285742672866
'PDC-AS1'	-0.835014653321143
'LMO7DN-IT1'	-0.655246626580307
AC022165.1	1.59089060719201
AC090948.1	-0.681906232917018
'ZNF8-ERVK3-1'	1.02052572330987
AC005479.2	-0.324707380405547
AC020915.2	0.575262289352546

which revealed that the TMB of the different risk groups varied (**Figure 8H**). In patients with LUSC, significant differences were not observed between the TMB scores of different groups (**Figure 8I**). TMB is the number of somatic mutations per megabase pair in a specific genomic region. The LUAD patients were split into low mutation load group and high mutation load group. The survival curve of the patients in the low mutation load group was similar to that of the patients in the high mutation load group, which signified that there was no obvious difference in survival between the two subgroups (**Figure 8J**). The OS of patients with a high mutation load and low mutation load in the high-risk group was worse than that of the patients with a high mutation load and low mutation load in the low-risk group. The survival curve of patients with a high mutation load in the high-risk group was analogous to that of the patients with a low mutation load, which implied that high and low mutation statuses cannot differentiate the survival rate of the high-risk group (**Figure 8K**). These dis-

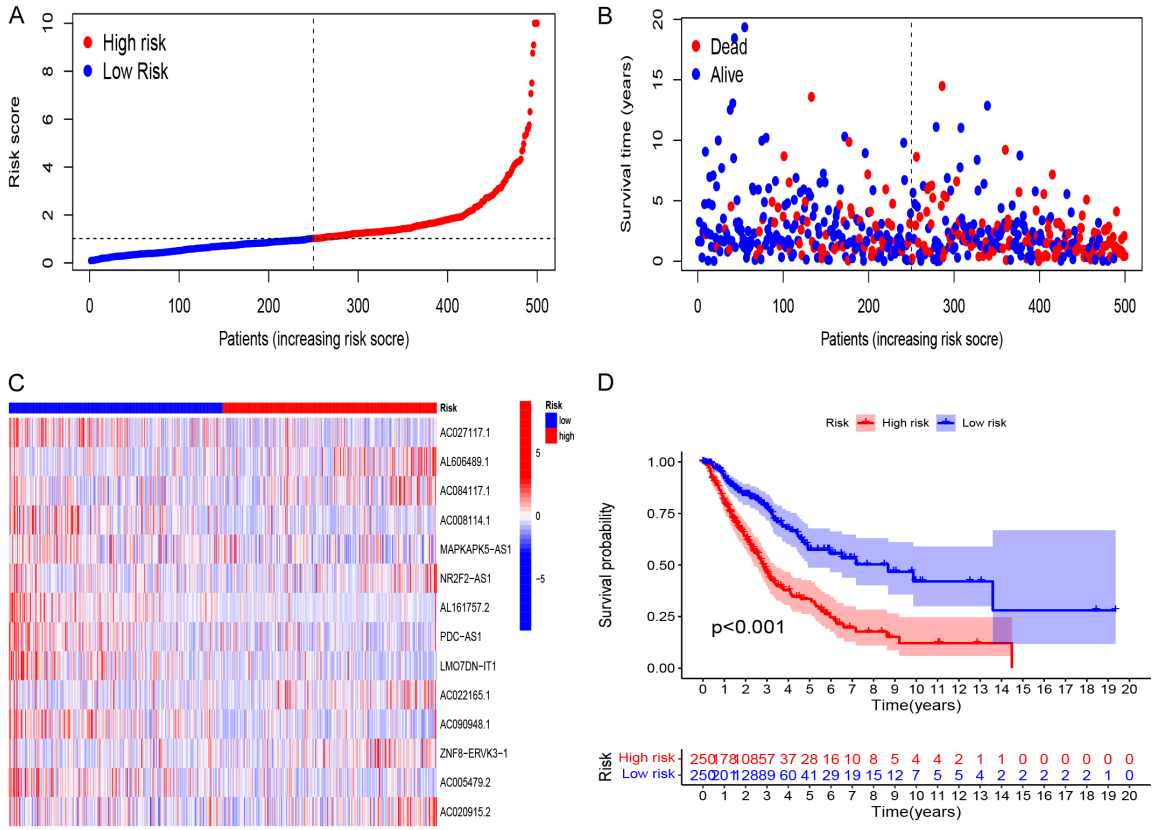


Figure 4. A. Distribution of risk scores of m6A related lncRNA model. B. Different patterns of survival status and survival time between high-risk group and low-risk group. C. Cluster analysis Heatmap showed the expression criteria of 14 prognostic lncRNAs in each patient. D. Kaplan Meier survival curve of patients with OS in high-risk and low-risk groups.

coveries indicate that m6A-associated lncRNA model may be more meaningful for disease prediction than high and low mutation load. The LUSC patients were split into high mutation load group and low mutation load group. The survival curve of the patients in the high mutation load group was higher than that of the patients in the low mutation load group (**Figure 8L**). The OS of the patients with a low mutation load in the high-risk group was worse than that of the patients with a high mutation load in the high-risk group, the survival outcome of the patients with a low mutation load in the low-risk group was worse than that of patients with a high mutation load in the low-risk group, the OS of the high-risk group in the low mutation load group was worse than that of the low-risk group, and the OS of the high-risk group in the high mutation load group was worse than that of the low-risk group (**Figure 8M**). These observations indicate that the m6A-associated lncRNA model has prognostic significance both in high and low mutation loads.

Screening of potential drugs for the m6A-related lncRNA model

To screen potential drugs through lncRNA model in NSCLC patients, the half maximum inhibitory concentration (IC50) was applied based on every sample in the GDSC data pool to estimate the treatment response. Ninety-five compounds were screened to determine whether there were obvious differences in the IC50 estimates between the two subgroups, and sensitivity of the high-risk group to all these potential drugs was higher (Supplementary Figure 1). In the high-risk group, sensitivity of 70 compounds to these potential drugs was higher. We selected 4 more sensitive compounds for analysis. Saracatinib (AZD.0530) is an SRC inhibitor, which can reduce the expression of PD-L1 in NSCLC cells, thereby enhancing the anti-tumor effect in NSCLC [7]. Bicalutamide is mainly used to treat prostate cancer [8]. GSK269962A is a rock inhibitor. The double effects of EGFR and rock on triple negative

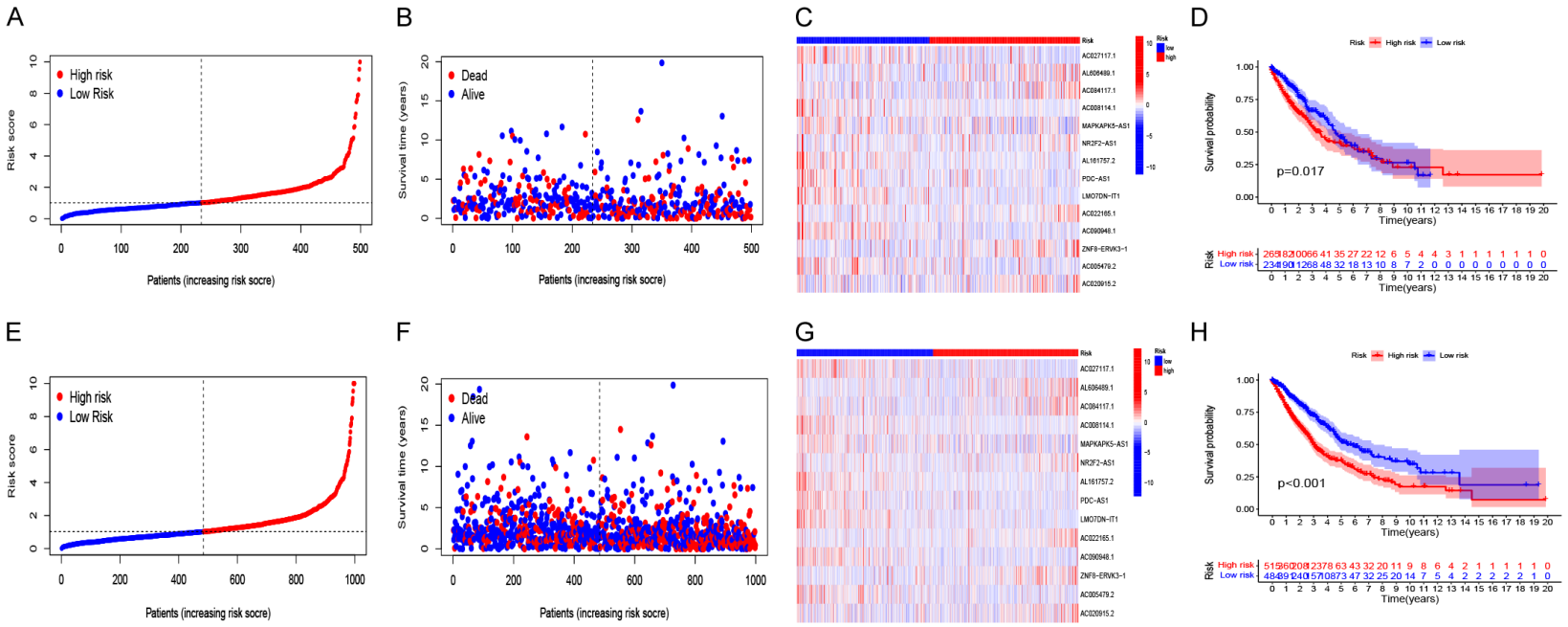


Figure 5. A. Distribution of m6A-related lncRNA model-based risk score for the testing set. B. Patterns of the survival time and survival status between the high-risk and low-risk groups for the testing set. C. Cluster analysis Heatmap showed the display level of 14 prognostic lncRNAs in each patient in the test set. D. Kaplan Meier survival curve of OS in high-risk group and low-risk group. E. Distribution of the m6A-related lncRNA model-based risk score for the entire set. F. Patterns of the survival time and survival status between the high-risk and low-risk groups for the entire set. G. Cluster analysis Heatmap showed the expression level of 14 prognostic lncRNAs in each patient in the whole set. H. Kaplan Meier survival curve of OS in low-risk group and high-risk group.

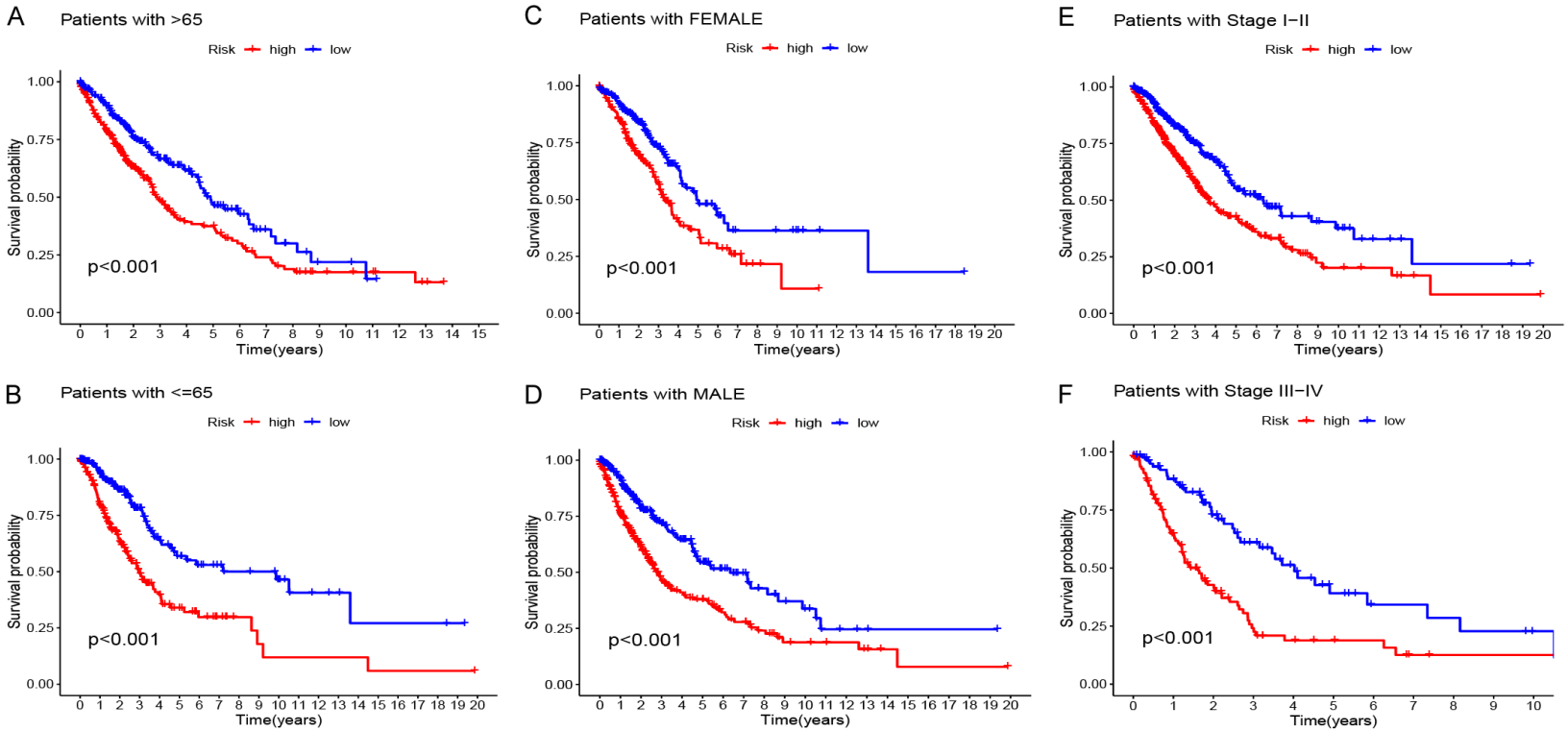


Figure 6. Kaplan Meier curve of OS difference by gender, age and tumor stage between high-risk group and low-risk group in TCGA.

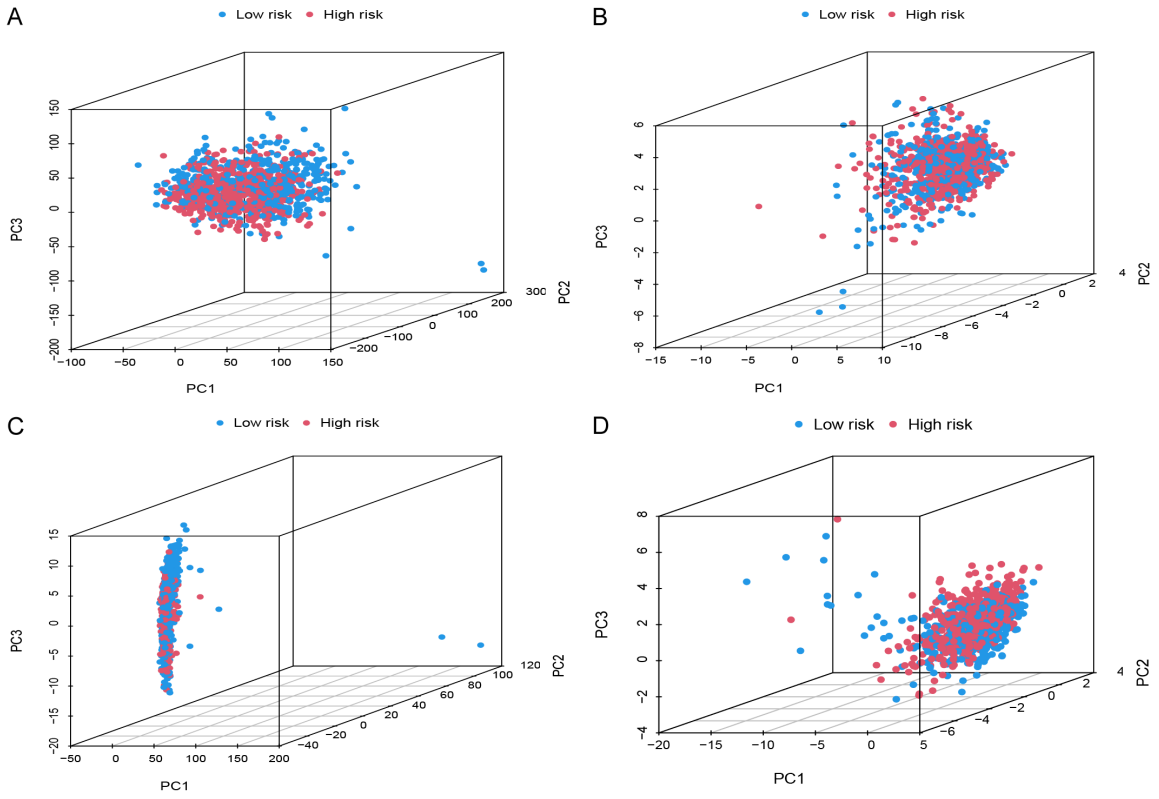


Figure 7. (A) Whole gene expression profiles, (B) 23 m6A genes, (C) 14 m6A related lncRNAs, and (D) Risk models based on 14 m6A related lncRNAs expression profiles.

breast cancer cells can enhance their autophagy and lead to cancer cell death [9]. These two are less studied in lung cancer. TAE684 (NVP-TAE684) is a selective ALK inhibitor. In lung cancer cell lines containing wild-type h694r or e1384k mutant ALK, TAE684 effectively inhibited the proliferation and induced apoptosis of cells expressing h694r or e1384k mutant ALK [10].

Evaluation of the m6A-associated lncRNA prognostic model in NSCLC

Single factor and multifactor Cox analysis was used to estimate whether the prognostic model of the 14 m6A-associated lncRNAs was an independent prognostic factor for NSCLC. In single factor Cox analysis, the hazard ratio (HR) and 95% confidence interval (CI) of the risk score were 1.251 and 1.198-1.307, respectively ($P < 0.001$) (**Figure 9A**). In multifactor Cox analysis, HR was 1.245 and 95% CI was 1.191-1.302 ($P < 0.001$) (**Figure 9B**), which demonstrated that the risk model was not associated with clinical features such as age, tumor stage, or sex. As an independent prognostic indicator, our m6A prognostic model is probably helpful for clinical prognostic evaluation. The compliance index and area under the ROC curve (AUC) of the risk score were evaluated. The consistency index of the risk score is always higher than that of other clinical factors and gradually decreases over time, implying that the risk level had a better effect in predicting the prognosis of NSCLC (**Figure 9C**). The study further drew the ROC curve of risk score, age, gender and tumor stage to predict OS, and the AUC values were 0.686, 0.539, 0.551 and 0.629 respectively, indicating that the prognostic model of risk score is reliable for the prognosis evaluation of NSCLC (**Figure 9D**). The ROC curve of time has AUC values of 0.686 for 1-year survival, 0.669 for 3-year survival, and 0.642 for 5-year survival, suggesting that the model is relatively stable, which means that the m6A related lncRNAs model has a good ability in evaluating prognosis (**Figure 9E**).

Establishment and evaluation of column line diagram

A nomogram containing the clinical risk characteristics and risk were used to predict the incidence rate of OS in 1, 3, and 5 years. In com-

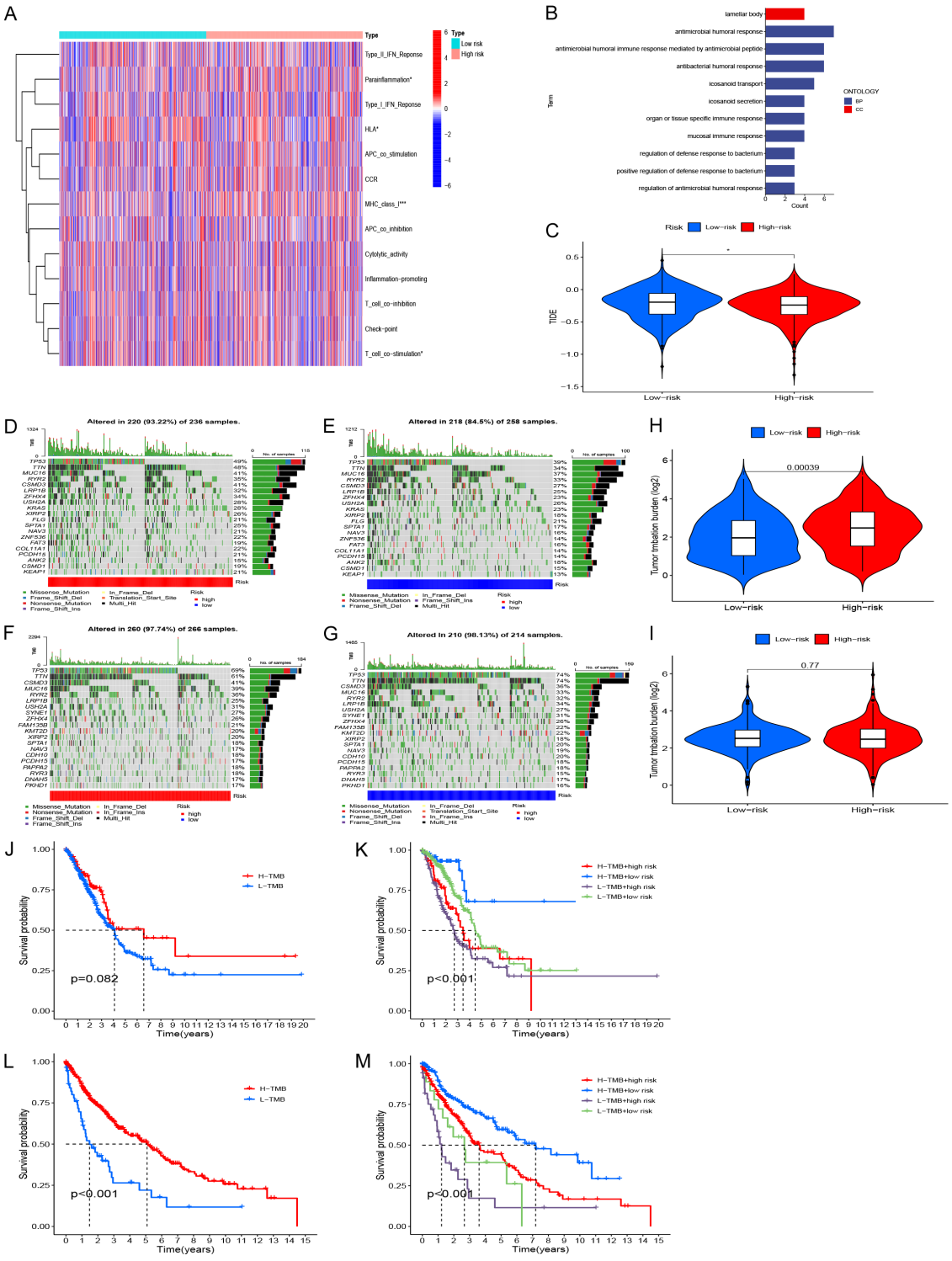


Figure 8. (A) The immune index of each patient indicates the standard. (B) GO enrichment analysis. (C) TIDE prediction differences between high-risk and low-risk patients. (D and E) Waterfall plots show mutation information of genes with high mutation frequency in LUAD high-risk group (D) and low-risk group (E). (F and G) Waterfall plots show the mutation information of high mutation frequency genes in LUSC high-risk group (F) and low-risk group (G). (H and I) TMB difference between high-risk and low-risk patients with LUAD (H). Difference in TMB between LUSC high-risk and low-risk patients (I). (J and L) Kaplan Meier curve analysis showing OS in LUAD patients according to mutation load status (J). Kaplan Meier curve analysis (L) showing OS in LUSC patients according to mutation load status. (K and M) Kaplan Meier curve analysis of LUAD patients classified according to mutation load status and m6A related lncRNA model (K). LUSC patients classified according to mutation load status and m6A related lncRNA model showed Kaplan Meier curve analysis of OS (M).

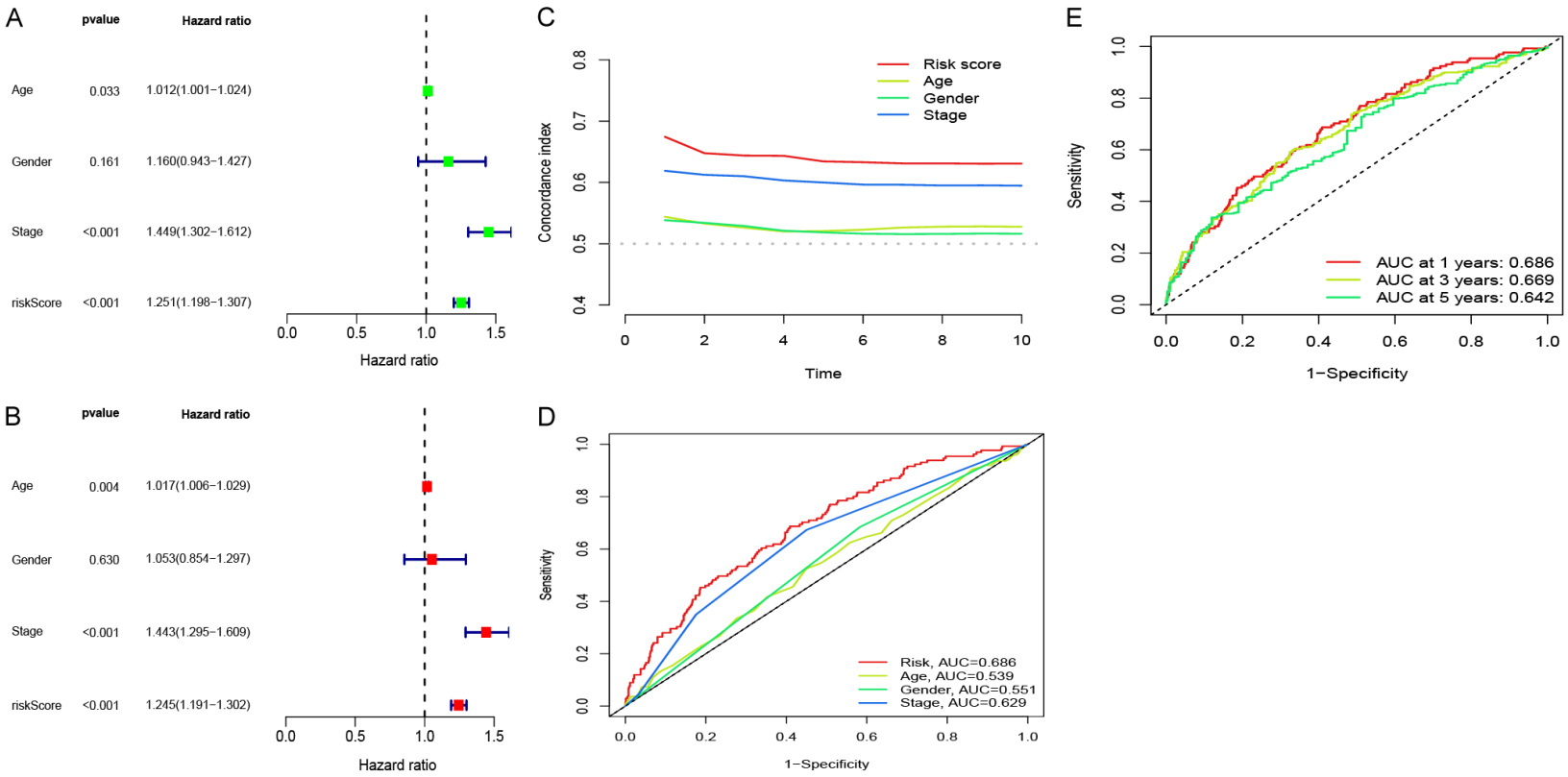


Figure 9. A. Univariate analysis of clinical characteristics and OS risk score. B. Multivariate analysis of clinical characteristics and OS risk score. C. Consistency index of risk score and clinical characteristics. D. ROC curve of clinical characteristics and risk score. E. ROC curve of time.

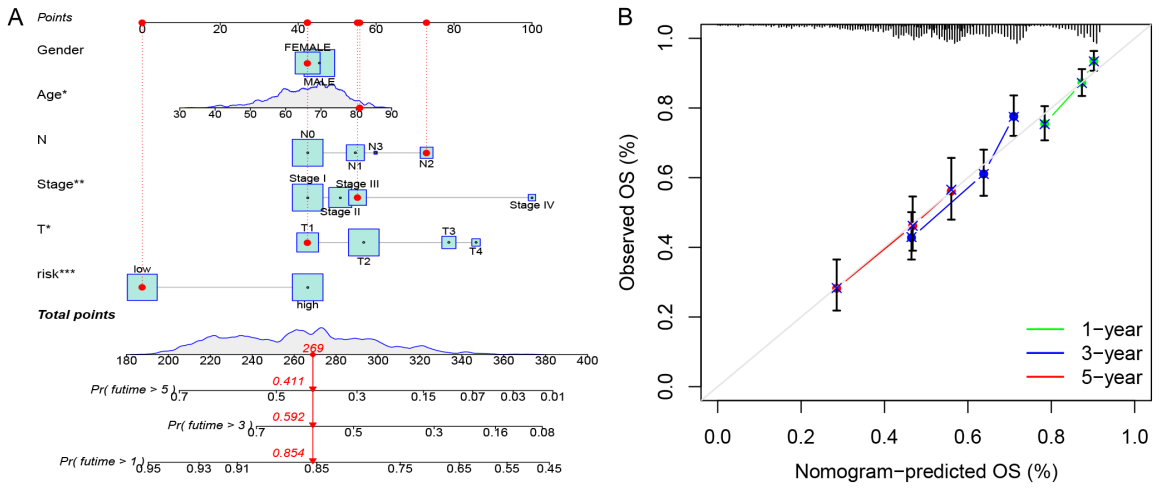


Figure 10. A. The nomogram predicts the probability of the 1-, 3-, and 5-year OS. B. The calibration plot of the nomogram predicts the probability of the 1-, 3-, and 5-year OS.

parison with the clinical features, the risk level of the prognostic model demonstrated outstanding predictive power in the nomogram (**Figure 10A**). The associated diagram showed that the calculated 1-, 3-, and 5-year OS ratios were consistent with the predicted ratios (**Figure 10B**).

Discussion

The most common type of pulmonary carcinoma is NSCLC. Several articles mainly identify the characteristics of ncRNAs to predict the survival rate and immunotherapy response of the NSCLC patients.

More than 100 chemical modifications have been identified in cellular RNA [11], including messenger RNA (mRNA) and non-coding RNA (ncRNA). m6A is the principal embellishment in the mRNA of many eukaryotic species [11]. Numerous studies have documented that m6A modification regulates the pathogenesis of cancer.

lncRNA is an important regulatory factor [12], and thousands of lncRNAs exist in the human genome [13]. They have tissue specificity, sequence conservation, and low expression [14]. Recent research has demonstrated that lncRNAs can regulate gene expression and affect many important physiological processes with a variety of functions, such as chromatin modification, enhancer, transcriptional regulation and so on [12, 16]. Mounting evidence has indicated that lncRNAs are associated with the development of many human diseases [15]. The expression profiles of lncRNAs are different in diverse physiological and pathological environments, which denotes that these biological molecules may reflect the disease state. Researchers have found that some lncRNAs are also involved in tuberculosis, interstitial lung disease, chronic obstructive pulmonary disease, and asthma [17]. Extracellular lncRNAs can also be used as diagnostic biomarkers of pulmonary carcinoma. For example, lncRNA growth arrest-special transcript 5 (GAS5) is important for cancer progression. In contrast to healthy controls, the expression of Exo-GAS5 was downregulated in patients with pulmonary carcinoma [18]. In addition, Exo-GAS5 expression was higher in NSCLC patients with small tumors and early Tumor Node Metastasis (TNM) classification. Thus, Exo-GAS5 can be used to distinguish patients with early NSCLC [19]. Another study showed that the expression of plasma exons ENSG00000245648 and SOX2-OT were increased in lung squamous cell carcinoma patients in comparison with the negative control group. This finding revealed that SOX2-OT level can reflect the tumor state to a certain extent [20]. In common, other research indicated that the expression of exosomal MALAT-1 was increased in the NSCLC patients. On the basis of their research, MALAT-1 included in the exons in serum can promote the invasion and metastasis of pulmonary carcinoma cells because of progress in the cell cycle and decrease in the apoptosis of pulmo-

nary carcinoma cells. Knockout of MALAT-1 in NSCLC cells can inhibit the invasion and metastasis of these carcinoma cells [21]. As described above, because exosomal lncRNAs are involved in the development of pulmonary carcinoma, they are helpful for the diagnosis of pulmonary carcinoma. In the future, inhibiting the generation and secretion of exosomes and downregulating the level of lncRNA in the exosomes may be used as a novel method for the treatment of pulmonary carcinoma. Researchers have shown that exogenous lncRNAs are also related to the drug resistance of some cancers, such as pulmonary carcinoma. Compared with normal NSCLC cells, the lncRNA RP11-838N2.4 was increased in cells with erlotinib resistance; and knocking out the lncRNA RP11-838N2.4 can remove this influence [22]. The H19 expression increased in Gefitinib-resistant drug resistance cells. Extracellular H19 promoted the resistance of the NSCLC cells to gefitinib by packaging them into exons [23, 24]. The pivotal *in vitro* role of lncRNAs in the acquisition of drug resistance of pulmonary carcinoma cells might provide a method for lowering the failure of chemotherapy for pulmonary carcinoma.

m6A-modified lncRNAs have been researched further. For example, Xue et al. found that the lncRNA ABHD11-AS1 was increased in NSCLC tissue samples and cells, and ectopic excessive expression would lead to poor prognosis in NSCLC patients. MeRIP-seq revealed that m6A methyltransferases, such as METTL3, had an m6A modification site, which enhanced the expression of ABHD11-AS1. The results emphasize that ABHD11-AS1 is highly expressed in NSCLC tissues and cancer cells. Knocking down the expression of ABHD11-AS1 can inhibit cell proliferation and migration *in vitro* [25, 26].

In our research, 1651 m6A-associated lncRNAs were identified to determine the prognostic value of m6A-associated lncRNAs. The TCGA data pool authenticated the prognostic value of 25 m6A-associated lncRNAs, of which 14 were used to construct m6A-associated lncRNA models to predict OS in the NSCLC patients. In view of the medium risk fraction, the NSCLC patients were divided into different risk subgroups. The clinical results of the high-risk group were obviously worse. ROC curve indicated that previous clinical characteristics were worse than the model in predicting the survival rate of the NSCLC patients. A nomogram was also built to show the uniformity between OS and the predicted rates at 1, 3, and 5 years. This prognostic risk model is quite accurate in predicting OS in NSCLC patients, and the prediction model played an important role in identifying new biomarkers in subsequent studies.

In recent times, immunotherapy has become the focus of cancer therapy. The only predictive biomarker used at present for patient selection is PD-L1, but it has some limitations. Hence, developing biomarkers, such as TMB, to help in patient selection remains the focus of ongoing research [27]. The survival curve of patients with a high mutation load was found to be higher than that of patients with a low mutation load, and the prognosis was better. A recent study compared nivolumab combined with chemotherapy for the first-line treatment of advanced NSCLC. For patients with high TMB, those treated with nivolumab had a higher effective rate than those treated with chemotherapy. Notably, patients with PD-L1 and high TMB had the best prognosis [28]. Our results are consistent with this observation. In our research, the TIDE score of high-risk group is lower than that of low-risk group, and the results are statistically significant, which means that patients in high-risk group are more sensitive to immunotherapy. In addition, our research has provided new insights into the regulatory mechanism of m6A-associated lncRNAs in NSCLC.

Conclusion

In conclusion, we analyzed the clinical data and expression profile of NSCLC specimens in TCGA database. The prognosis model based on the relevant genes of Bioscience analysis can effectively predict the efficacy and prognosis of immunotherapy in NSCLC patients. The potential drugs for patients with NSCLC were selected using the model, thereby yielding new therapeutic targets and prognostic biomarkers for follow-up studies.

Acknowledgements

This work is supported by the National Natural Science Foundation (81570078) and the Key Research and Development Program of Jiangsu Province (grant number BE2019719).

Disclosure of conflict of interest

None.

Address correspondence to: Hong-Bing Liu, Department of Respiratory Medicine, Jinling Hospital, Medical School of Nanjing University, Nanjing, China. Tel: +86-13852293363; E-mail: netlhb@126.com; Ying Wu, Affiliated Hospital of Nanjing University of Chinese Medicine, Jiangsu Province Hospital of Chinese Medicine, Nanjing, China; First College of Clinical Medicine, Nanjing University of Chinese Medicine, Nanjing, China. E-mail: waitforwy@126.com

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394-424.

[2] Balata H, Fong KM, Hendriks LE, Lam S, Ostroff JS, Peled N, Wu N and Aggarwal C. Prevention and early detection for NSCLC: advances in thoracic oncology 2018. J Thorac Oncol 2019; 14: 1513-1527.

[3] Huang H, Weng H and Chen J. M6A modification in coding and non-coding RNAs: roles and therapeutic implications in cancer. Cancer Cell 2020; 37: 270-288.

[4] Melstrom L and Chen J. RNA N6-methyladenosine modification in solid tumors: new therapeutic frontiers. Cancer Gene Ther 2020; 27: 625-633.

[5] Shi Y, Fan S, Wu M, Zuo Z, Li X, Jiang L, Shen Q, Xu P, Zeng L, Zhou Y, Huang Y, Yang Z, Zhou J, Gao J, Zhou H, Xu S, Ji H, Shi P, Wu DD, Yang C and Chen Y. YTHDF1 links hypoxia adaptation and non-small cell lung cancer progression. Nat Commun 2019; 10: 4892.

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

[7] Lu FT, Luo F, Qiu MZ, Cao JX, Luo QY, Yang DJ and Zhao HY. SRC inhibitor saracatinib enhances efficacy of PD-1/PD-L1 immune checkpoint blockade in non-small cell lung cancer. Cancer Res 2022; 82: 6094.

[8] Kolvenbag GJ, Blackledge GR and Gotting-Smith K. Bicalutamide (casodex) in the treatment of prostate cancer: history of clinical development. Prostate 1998; 34: 61-72.

[9] Rontogianni S, Iskit S, van Doorn S, Peeper DS and Altelaar M. Combined EGFR and ROCK inhibition in TNBC leads to cell death via impaired autophagic flux. Mol Cell Proteomics 2020; 19: 261-277.

[10] Wang YW, Tu PH, Lin KT, Lin SC, Ko JY and Jou YS. Identification of oncogenic point mutations and hyperphosphorylation of anaplastic lymphoma kinase in lung cancer. Neoplasia 2011; 13: 704-15.

[11] Roundtree IA, Evans ME, Pan T and He C. Dynamic RNA modifications in gene expression regulation. Cell 2017; 169: 1187-1200.

[12] Xu J, Bai J, Zhang X, Lv Y, Gong Y, Liu L, Zhao H, Yu F, Ping Y, Zhang G, Lan Y, Xiao Y and Li X. A comprehensive overview of lncRNA annotation resources. Brief Bioinform 2017; 18: 236-249.

[13] Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J and Guigó R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012; 22: 1775-1789.

[14] Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A and Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011; 25: 1915-1927.

[15] Ponting CP, Oliver PL and Reik W. Evolution and functions of long noncoding RNAs. Cell 2009; 136: 629-641.

[16] Heo JB, Lee YS and Sung S. Epigenetic regulation by long noncoding RNAs in plants. Chromosome Res 2013; 21: 685-693.

[17] Sun T, Kalionis B, Lv G, Xia S and Gao W. Role of exosomal noncoding RNAs in lung carcinogenesis. Biomed Res Int 2015; 2015: 125807.

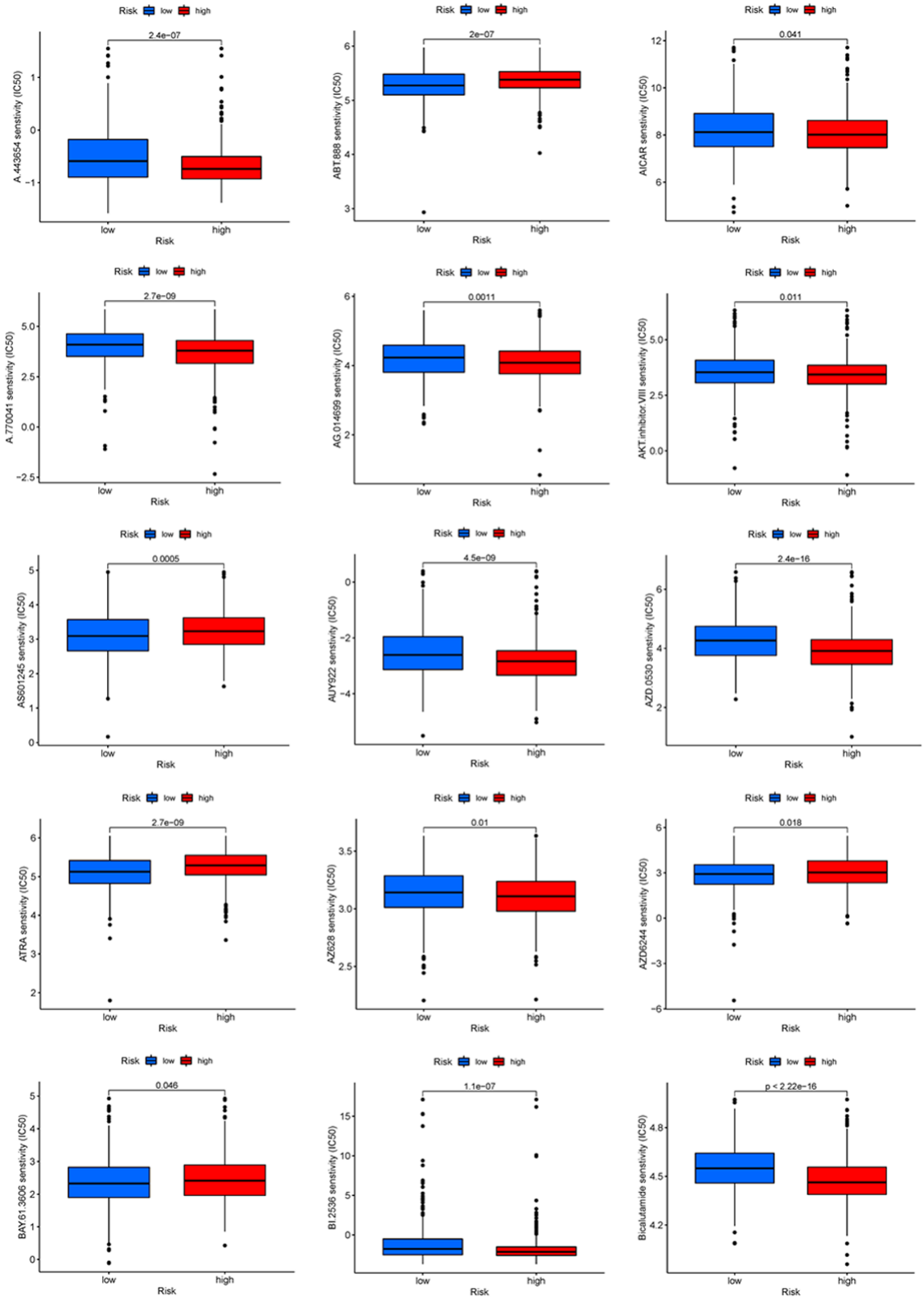
[18] Li Y, Yin Z, Fan J, Zhang S and Yang W. The roles of exosomal miRNAs and lncRNAs in lung diseases. Signal Transduct Target Ther 2019; 4: 47.

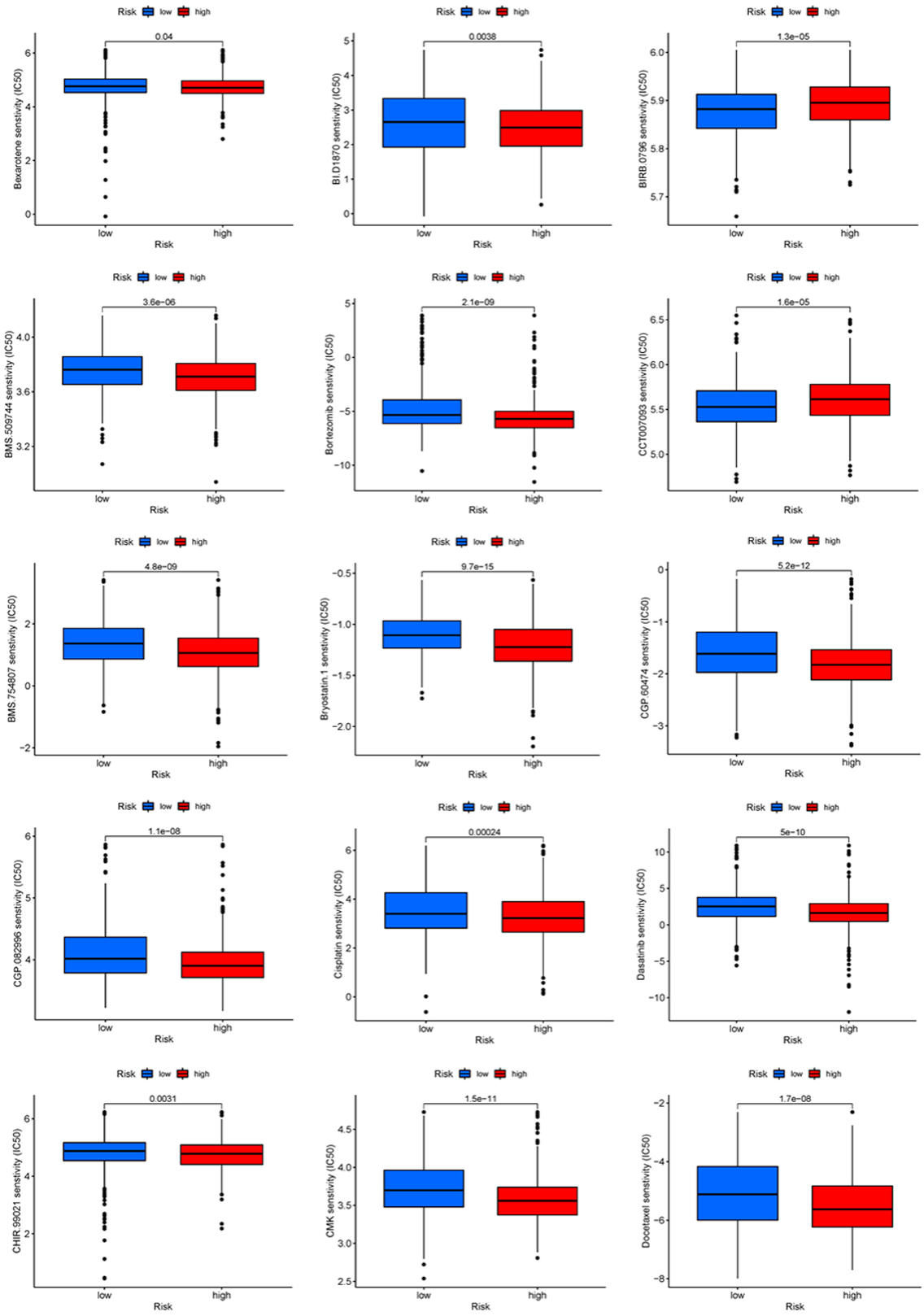
[19] Li C, Lv Y, Shao C, Chen C, Zhang T, Wei Y, Fan H, Lv T, Liu H and Song Y. Tumor-derived exosomal lncRNA GAS5 as a biomarker for early-stage non-small-cell lung cancer diagnosis. J Cell Physiol 2019; 234: 20721-20727.

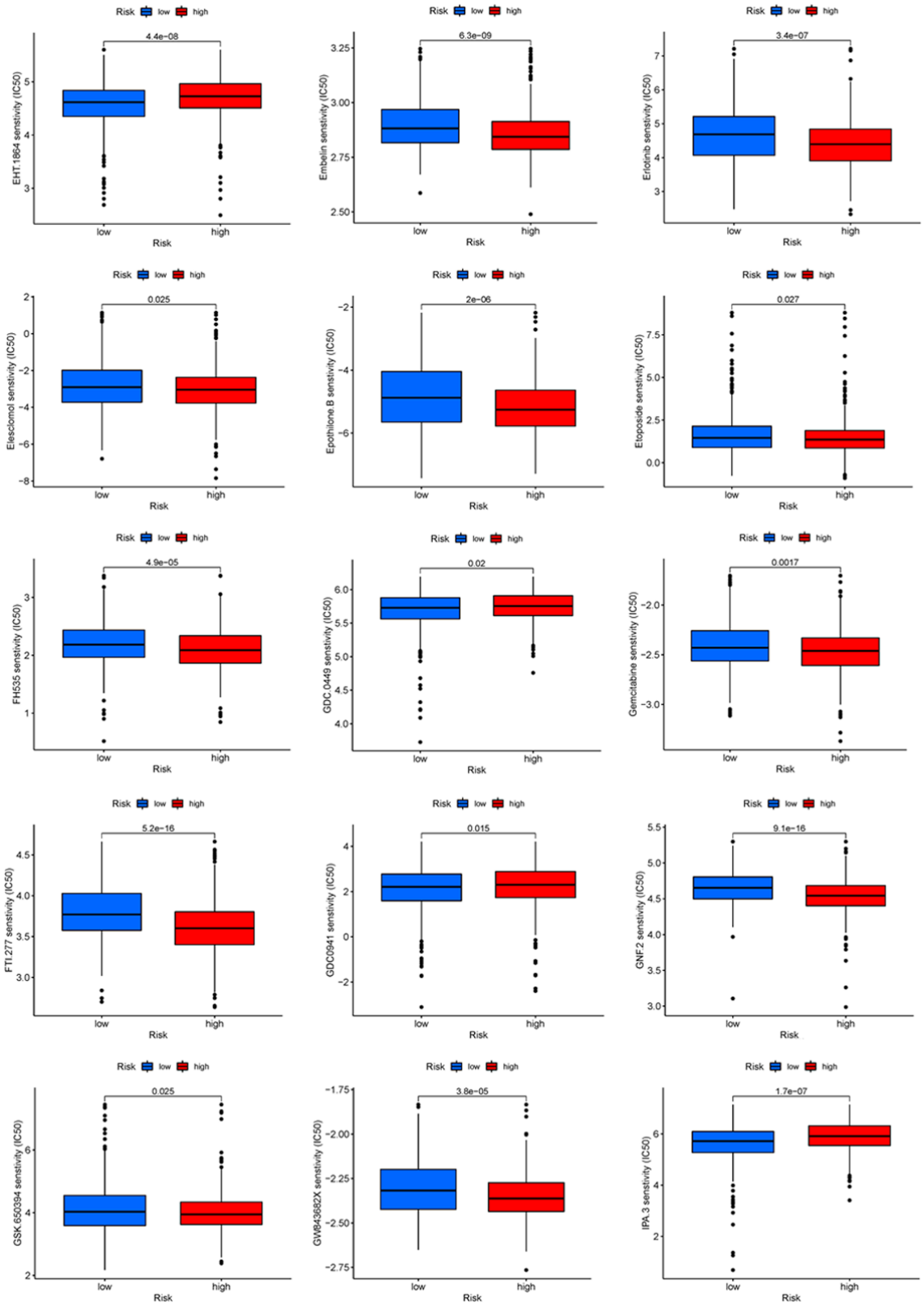
[20] Teng Y, Kang H and Chu Y. Identification of an exosomal long noncoding RNA SOX2-OT in plasma as a promising biomarker for lung squamous cell carcinoma. Genet Test Mol Biomarkers 2019; 23: 235-240.

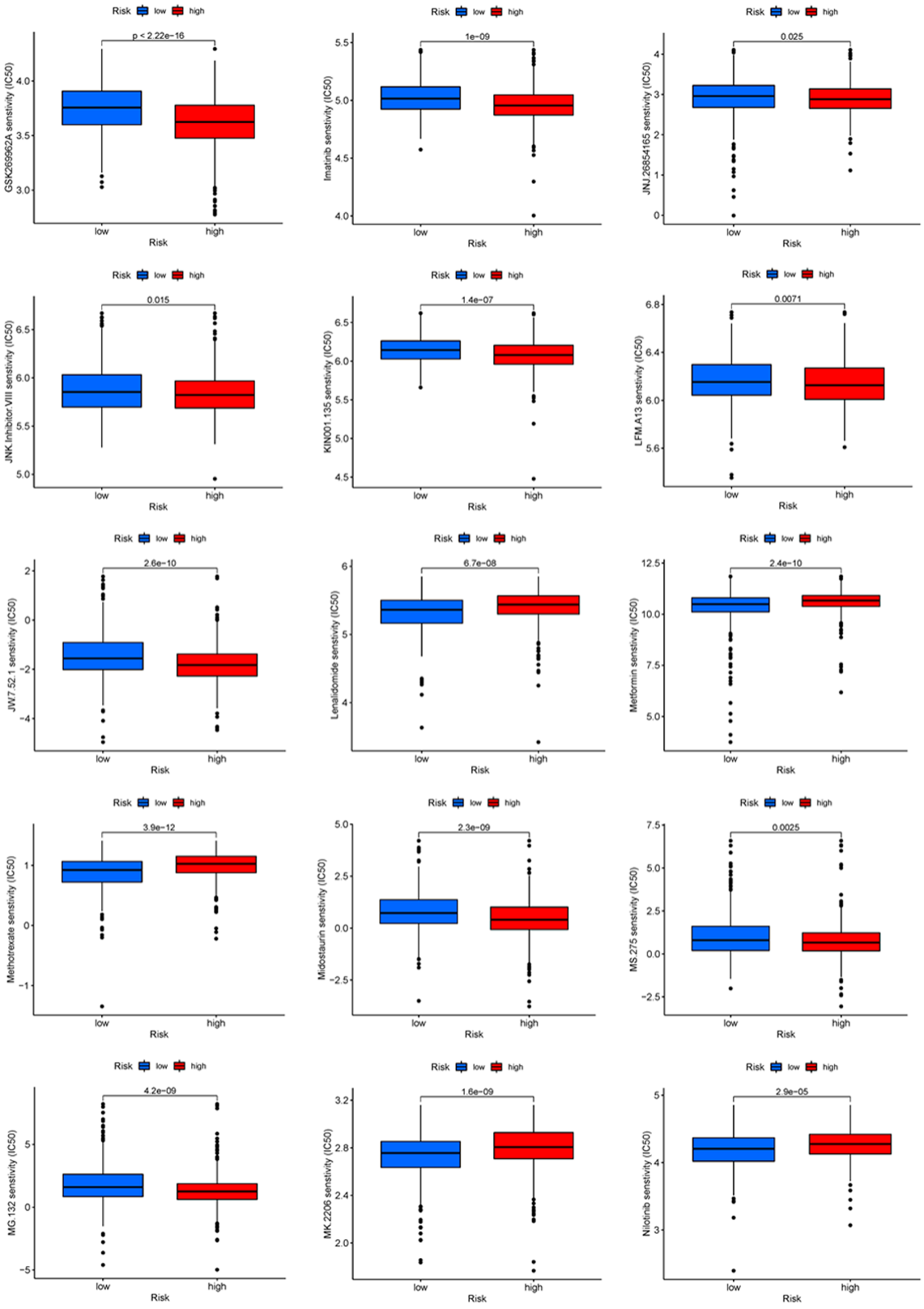
[21] Zhang R, Xia Y, Wang Z, Zheng J, Chen Y, Li X, Wang Y and Ming H. Serum long non coding RNA MALAT-1 protected by exosomes is up-regulated and promotes cell proliferation and migration in non-small cell lung cancer. Bio-

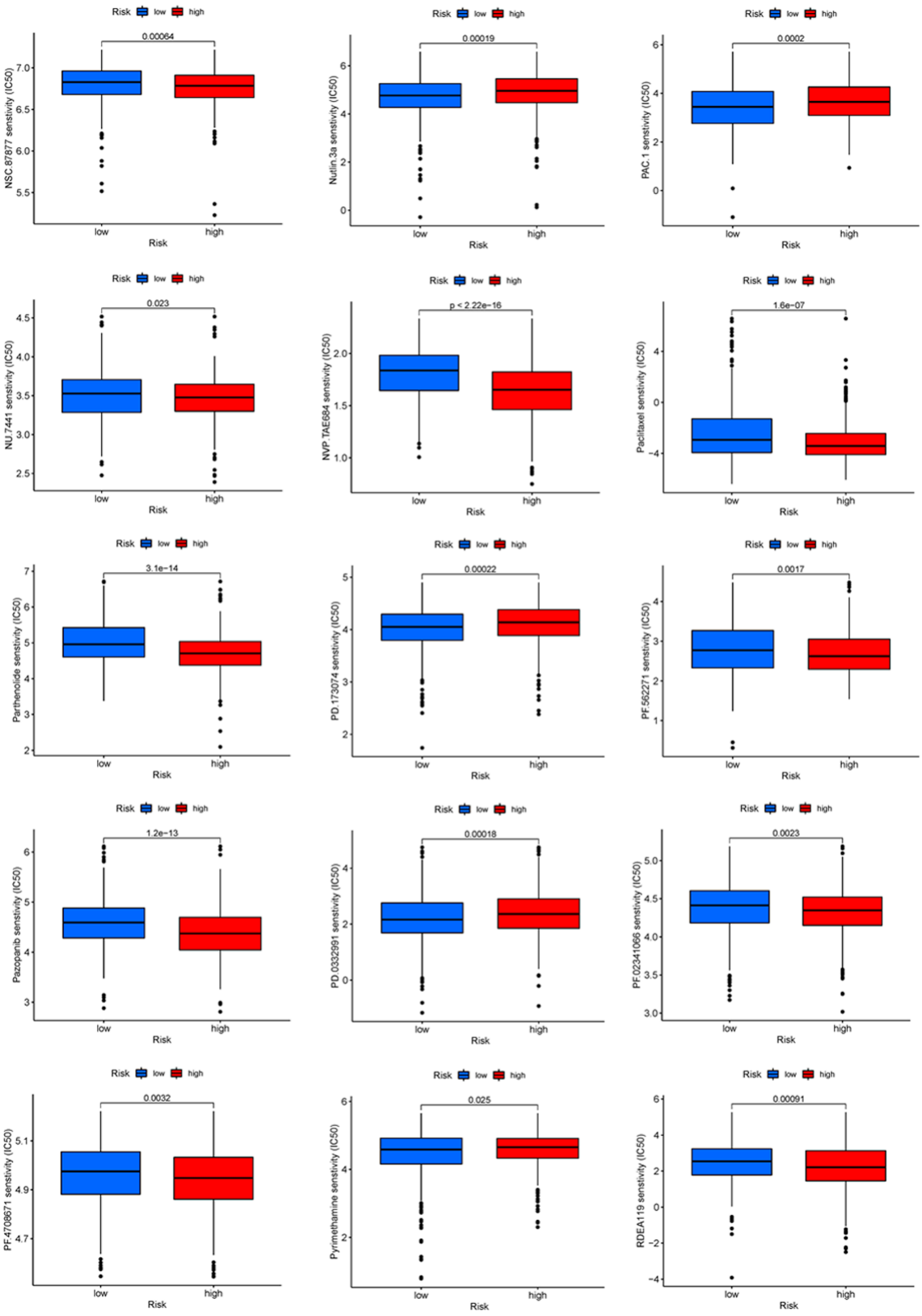
chem Biophys Res Commun 2017; 490: 406-414.
[22] Zhang W, Cai X, Yu J, Lu X, Qian Q and Qian W. Exosome-mediated transfer of lncRNA RP11-838N2.4 promotes erlotinib resistance in non-small cell lung cancer. Int J Oncol 2018; 53: 527-538.
[23] Lei Y, Guo W, Chen B, Chen L, Gong J and Li W. Tumor-released lncRNA H19 promotes gefitinib resistance via packaging into exosomes in non-small cell lung cancer. Oncol Rep 2018; 40: 3438-3446.
[24] Poulet C, Njock MS, Moermans C, Louis E, Louis R, Malaise M and Guiot J. Exosomal long non-coding RNAs in lung diseases. Int J Mol Sci 2020; 21: 3580.
[25] Xue L, Li J, Lin Y, Liu D, Yang Q, Jian J and Peng J. m6A transferase METTL3-induced lncRNA ABHD11-AS1 promotes the Warburg effect of non-small-cell lung cancer. J Cell Physiol 2021; 236: 2649-2658.
[26] Lan Y, Liu B and Guo H. The role of M6A modification in the regulation of tumor-related lncRNAs. Mol Ther Nucleic Acids 2021; 24: 768-779.
[27] Steuer CE and Ramalingam SS. Tumor mutation burden: leading immunotherapy to the era of precision medicine? J Clin Oncol 2018; 36: 631-632.
[28] Carbone DP, Reck M, Paz-Ares L, Creelan B, Horn L, Steins M, Felip E, van den Heuvel MM, Ciuleanu TE, Badin F, Ready N, Hiltermann TJN, Nair S, Juergens R, Peters S, Minenza E, Wrangle JM, Rodriguez-Abreu D, Borghaei H, Blumenschein GR Jr, Villaruz LC, Havel L, Krejci J, Corral Jaime J, Chang H, Geese WJ, Bhagavatheeswaran P, Chen AC and Socinski MA; CheckMate 026 Investigators. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. N Engl J Med 2017; 376: 2415-2426.

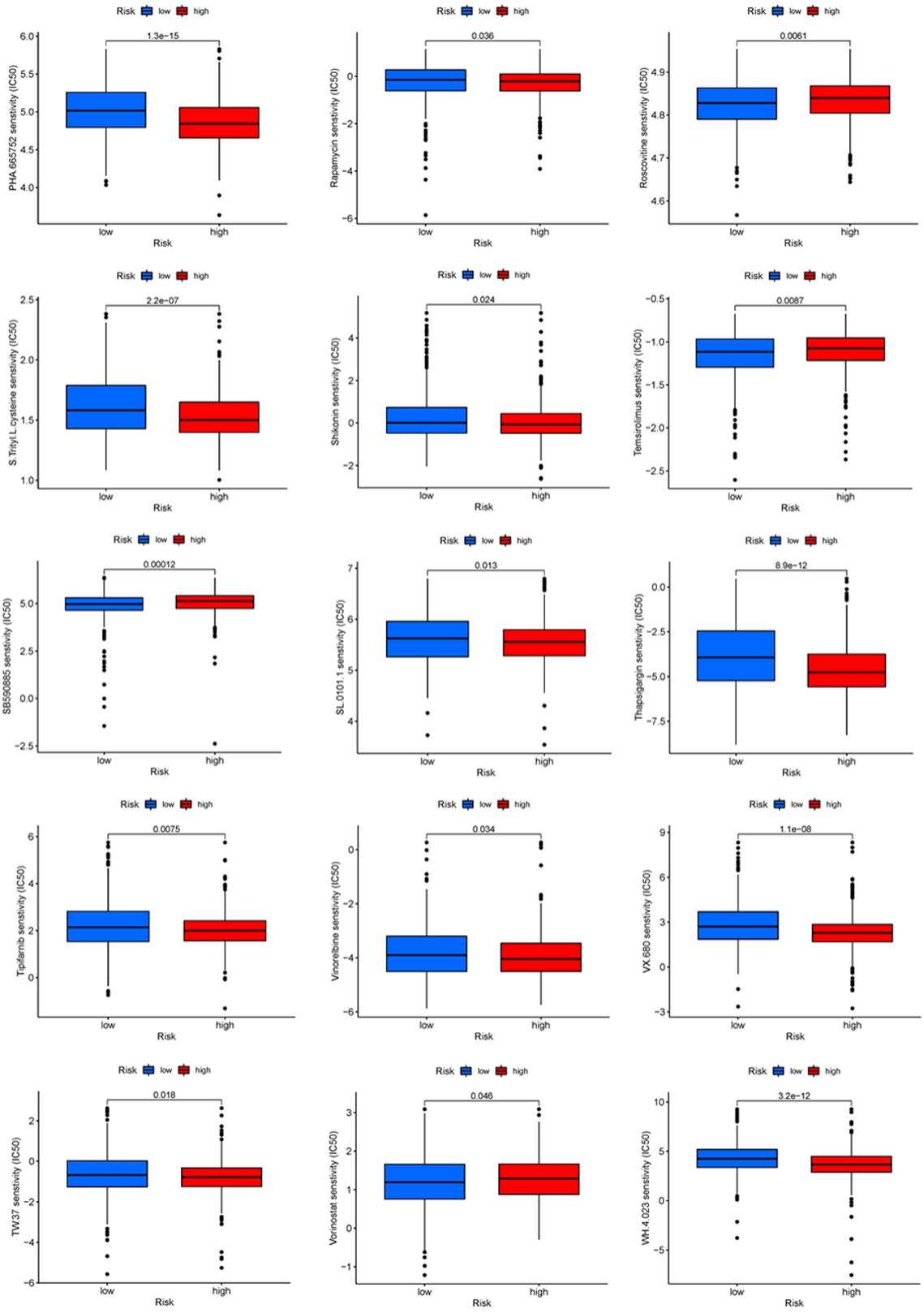

Risk low high
2.4e−07
A.443654 senstivity (IC50)
low high
Risk
Risk low high
2e−07
ABT.888 senstivity (IC50)
low high
Risk
Risk low high
0.041
AICAR senstivity (IC50)
low high
Risk
Risk low high
2.7e−09
A.770041 senstivity (IC50)
low high
Risk
Risk low high
0.0011
AG.014699 senstivity (IC50)
low high
Risk
Risk low high
0.011
AKT inhibitor VIII senstivity (IC50)
low high
Risk
Risk low high
0.0005
AS601245 senstivity (IC50)
low high
Risk
Risk low high
4.5e−09
AUY922 senstivity (IC50)
low high
Risk
Risk low high
2.4e−16
AZD.0530 senstivity (IC50)
low high
Risk
Risk low high
2.7e−09
ATRA senstivity (IC50)
low high
Risk
Risk low high
0.01
AZ628 senstivity (IC50)
low high
Risk
Risk low high
0.018
AZD6244 senstivity (IC50)
low high
Risk
Risk low high
0.046
BAY61.3606 senstivity (IC50)
low high
Risk
Risk low high
1.1e−07
BI.2536 senstivity (IC50)
low high
Risk
Risk low high
p < 2.22e−16
Bicalutamide senstivity (IC50)
low high
Risk

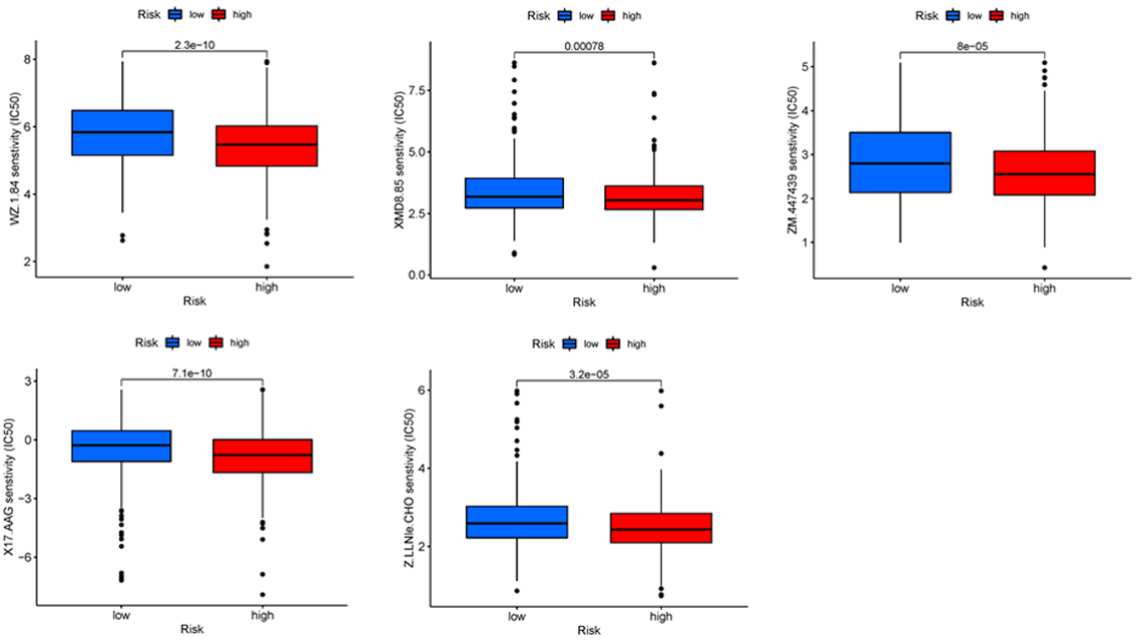

Risk low high
0.04
Bexarotene senstivity (IC50)
low high
Risk
Risk low high
0.0038
BI D1870 senstivity (IC50)
low high
Risk
Risk low high
1.3e−05
BIRB.0796 senstivity (IC50)
low high
Risk
Risk low high
3.6e−06
BMS.509744 senstivity (IC50)
low high
Risk
Risk low high
2.1e−09
Bortezomib senstivity (IC50)
low high
Risk
Risk low high
1.6e−05
CCT007093 senstivity (IC50)
low high
Risk
Risk low high
4.8e−09
BMS.754807 senstivity (IC50)
low high
Risk
Risk low high
9.7e−15
Bryostatin.1 senstivity (IC50)
low high
Risk
Risk low high
5.2e−12
CGP.60474 senstivity (IC50)
low high
Risk
Risk low high
1.1e−08
CGP.082996 senstivity (IC50)
low high
Risk
Risk low high
0.00024
Cisplatin senstivity (IC50)
low high
Risk
Risk low high
5e−10
Dasatinib senstivity (IC50)
low high
Risk
Risk low high
0.0031
CHIR.99021 senstivity (IC50)
low high
Risk
Risk low high
1.5e−11
CMK senstivity (IC50)
low high
Risk
Risk low high
1.7e−08
Docetaxel senstivity (IC50)
low high
Risk


Risk low high
4.4e−08
EHT.1864 senstivity (IC50)
low high
Risk
Risk low high
6.3e−09
Embelin senstivity (IC50)
low high
Risk
Risk low high
3.4e−07
Erlotinib senstivity (IC50)
low high
Risk
Risk low high
0.025
Elesclomol senstivity (IC50)
low high
Risk
Risk low high
2e−06
Epothilone B senstivity (IC50)
low high
Risk
Risk low high
0.027
Etoposide senstivity (IC50)
low high
Risk
Risk low high
4.9e−05
FH535 senstivity (IC50)
low high
Risk
Risk low high
0.02
GDC.0449 senstivity (IC50)
low high
Risk
Risk low high
0.0017
Gemcitabine senstivity (IC50)
low high
Risk
Risk low high
5.2e−16
FTI.277 senstivity (IC50)
low high
Risk
Risk low high
0.015
GDC0941 senstivity (IC50)
low high
Risk
Risk low high
9.1e−16
GNF.2 senstivity (IC50)
low high
Risk
Risk low high
0.025
GSK.650394 senstivity (IC50)
low high
Risk
Risk low high
3.8e−05
GW843682X senstivity (IC50)
low high
Risk
Risk low high
1.7e−07
IPA.3 senstivity (IC50)
low high
Risk


Risk low high
p < 2.22e−16
GSK269962A senstivity (IC50)
1e−09
Imatinib senstivity (IC50)
0.025
JNJ.26854165 senstivity (IC50)
0.015
JNK.Inhibitor.VIII senstivity (IC50)
1.4e−07
KIN001.135 senstivity (IC50)
0.0071
LFM.A13 senstivity (IC50)
2.6e−10
JW7.52.1 senstivity (IC50)
6.7e−08
Lenalidomide senstivity (IC50)
2.4e−10
Metformin senstivity (IC50)
3.9e−12
Methotrexate senstivity (IC50)
2.3e−09
Midostaurin senstivity (IC50)
0.0025
MS.275 senstivity (IC50)
4.2e−09
MG.132 senstivity (IC50)
1.6e−09
MK.2206 senstivity (IC50)
2.9e−05
Nilotinib senstivity (IC50)
low
high
Risk


Risk low high
0.00064
NSC.87877 senstivity (IC50)
low
high
Risk
Risk low high
0.00019
Nutlin.3a senstivity (IC50)
low
high
Risk
Risk low high
0.0002
PAC.1 senstivity (IC50)
low
high
Risk
Risk low high
0.023
NU.7441 senstivity (IC50)
low
high
Risk
Risk low high
p < 2.22e−16
NVP.TAE684 senstivity (IC50)
low
high
Risk
Risk low high
1.6e−07
Paclitaxel senstivity (IC50)
low
high
Risk
Risk low high
3.1e−14
Parthenolide senstivity (IC50)
low
high
Risk
Risk low high
0.00022
PD.173074 senstivity (IC50)
low
high
Risk
Risk low high
0.0017
PF.562271 senstivity (IC50)
low
high
Risk
Risk low high
1.2e−13
Pazopanib senstivity (IC50)
low
high
Risk
Risk low high
0.00018
PD.0332991 senstivity (IC50)
low
high
Risk
Risk low high
0.0023
PF.02341066 senstivity (IC50)
low
high
Risk
Risk low high
0.0032
PF.4708671 senstivity (IC50)
low
high
Risk
Risk low high
0.025
Pyrimethamine senstivity (IC50)
low
high
Risk
Risk low high
0.00091
RDEA119 senstivity (IC50)
low
high
Risk


Risk low high
1.3e−15
PHA 665752 senstivity (IC50)
0.036
Rapamycin senstivity (IC50)
0.0061
Roscovitine senstivity (IC50)
2.2e−07
S.Trityl.L.cysteine senstivity (IC50)
0.024
Shikonin senstivity (IC50)
0.0087
Temsirolimus senstivity (IC50)
0.00012
SB590885 senstivity (IC50)
0.013
SL.0101.1 senstivity (IC50)
8.9e−12
Thapsigargin senstivity (IC50)
0.0075
Tipifarnib senstivity (IC50)
0.034
Vinorelbine senstivity (IC50)
1.1e−08
VX.680 senstivity (IC50)
0.018
TW.37 senstivity (IC50)
0.046
Vorinostat senstivity (IC50)
3.2e−12
WH.4.023 senstivity (IC50)
low
high
Risk



Supplementary Figure 1. Ninety-five compounds.