Research Paper

Consensus clustering and novel risk score model construction based on m6A methylation regulators to evaluate the prognosis and tumor immune microenvironment of early-stage lung adenocarcinoma

Miao He[1,*], Yuxue Zhi[1,*], Chao Li[1], Changming Zhao[2], Guangquan Yang[1], Jing Lv[2], Hong You[1], Hai Huang[1], Xiaoyu Cao[1]

[1]Department of Radiation Oncology, People's Hospital of Deyang, Deyang 618000, Sichuan, P.R. China
[2]Department of Cardiovascular Surgery, People's Hospital of Deyang, Deyang 618000, Sichuan, P.R. China
*Equal contribution and co-first author

Correspondence to: Miao He; **email:** hmiao2021123@163.com, https://orcid.org/0000-0001-7337-9355
Keywords: early-stage lung adenocarcinoma, tumor immune microenvironment, m^{6}A methylation regulators, risk score, prognostic model
Received: November 15, 2023 **Accepted:** May 30, 2024 **Published:** July 5, 2024

ABSTRACT

Background: The aim of this study was to investigate the correlation between m^{6}A methylation regulators and cell infiltration characteristics in tumor immune microenvironment (TIME), so as to help understand the immune mechanism of early-stage lung adenocarcinoma (LUAD).
Methods: The expression and consensus cluster analyses of m^{6}A methylation regulators in early-stage LUAD were performed. The clinicopathological features, immune cell infiltration, survival and functional enrichment in different subtypes were analyzed. We also constructed a prognostic model. Clinical tissue samples were used to validate the expression of model genes through real-time polymerase chain reaction (RT-PCR). In addition, cell scratch assay and Transwell assay were also performed.
Results: Expression of m^{6}A methylation regulators was abnormal in early-stage LUAD. According to the consensus clustering of m^{6}A methylation regulators, patients with early-stage LUAD were divided into two subtypes. Two subtypes showed different infiltration levels of immune cell and survival time. A prognostic model consisting of HNRNPC, IGF2BP1 and IGF2BP3 could be used to predict the survival of early-stage LUAD. RT-PCR results showed that HNRNPC, IGF2BP1 and IGF2BP3 were significantly up-regulated in early-stage LUAD tissues. The results of cell scratch assay and Transwell assay showed that overexpression of HNRNPC promotes the migration and invasion of NCI-H1299 cells, while knockdown HNRNPC inhibits the migration and invasion of NCI-H1299 cells.
Conclusions: This work reveals that m^{6}A methylation regulators may be potential biomarkers for prognosis in patients with early-stage LUAD. Our prognostic model may be of great value in predicting the prognosis of early-stage LUAD.

INTRODUCTION

Non-small-cell lung cancer (NSCLC) accounts for 85% of cases of lung cancer [1]. Among NSCLC, lung adenocarcinoma (LUAD) is the most common [2]. One of the deadliest and most aggressive tumor kinds, LUAD has a less than 5-year overall survival [3]. Therefore, early diagnosis and treatment of LUAD are

particularly important. Low-dose computed tomography (CT) can be helpful in the diagnosis of early-stage LUAD [4]. Surgical resection and radiotherapy are commonly used in early-stage LUAD [5]. However, overall survival after treatment is not ideal [6]. Recently, immunotherapy has become a new type of treatment. This method provides new ideas for clinical treatment management. The changes of immune microenvironment of early-stage LUAD affect disease progression [7, 8]. Therefore, in order to accurately predict the prognosis and significantly optimize immunotherapy management of early-stage LUAD, further exploration of the regulatory mechanisms of tumor immune microenvironment (TIME) is needed.

The modification of RNA by m^6A is determined by the dynamic interaction between its methyltransferase (writer), binding protein (reader) and demethylase (erase) [9]. It has recently been discovered that m^6A methylation regulators are crucial for TIME. The ALKBH5 can improve the efficacy of immunotherapy by regulating lactate content and inhibiting immune cell accumulation in TIME [10]. Evaluation of m^6A modification patterns in patients with gastric cancer can improve our understanding of the characteristics of TIME infiltration and guide immunotherapy strategies more effectively [11]. The METTL3- or METTL14-deficient increases cytotoxic tumor-infiltrating CD8 T cells and alters TIME [12]. Therefore, exploring the role of m^6A methylation regulators in TIME may be helpful for the implementation of immunotherapy.

According to earlier research, m^6A methylation regulators are essential for controlling lung cancer. YTHDC2, a m^6A reader inhibits the development of LUAD by SLC7A11-dependent antioxidant function. The expression of m^6A methyltransferase METTL3 is increased in LUAD, and can regulate the growth of cancer cells [13]. The m^6A methylation regulators also regulate the development of LUAD by regulating the TIME [14, 15]. Nonetheless, the potential mechanism of m^6A methylation regulators in immune infiltration in the TIME of early-stage LUAD remains unclear.

Herein, all data of early-stage LUAD were downloaded from public databases. Then, the expression and consensus cluster analyses of m^6A methylation regulators in early-stage LUAD was performed. Two subtypes of early-stage LUAD were identified. Subsequently, we also constructed a prognostic model.

MATERIALS AND METHODS

Data sources

All data of LUAD were obtained from TCGA (https://tcga-data.nci.nih.gov/tcga/) and GEO (http://www.ncbi.nlm.nih.gov/geo) databases. Early-stage LUAD patients with TNM stage I-II and survival information were selected. UCSC Xena (https://gdc.xenahubs.net) was used to download RNA sequencing data and clinical information in TCGA. The original "CEL" file of the GSE31210 dataset was downloaded from GEO database. Finally, 398 early-stage LUAD samples and 59 paracancerous samples were obtained from TCGA database, and 226 early-stage LUAD samples and 20 paracancerous samples were obtained from GEO database (Supplementary Table 1).

Expression of m^6A methylation regulators

Herein, 22 m^6A methylation regulators were collected from the published literature [9, 16, 17]. These 22 m^6A methylation regulators include 2 erasers (ALKBH5 and FTO), 12 readers (IGF2BP1, IGF2BP2, IGF2BP3, LRPPRC, YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, FMR1, HNRNPA2B1 and HNRNPC) and 8 writers (KIAA1429, METTL3, RBM15, RBM15B, WTAP, ZC3H13, METTL14 and METTL16). Then, we investigated the differential expression of these molecules in early-stage LUAD and normal tissue in TCGA. We also analyzed somatic mutations of 22 m^6A methylation regulators in early-stage LUAD. Mutation data were obtained from TCGA.

Unsupervised clustering

The "ConsensusClusterPlus" software [18] (http://www.bioconductor.org/) was used to divide patients with early-stage LUAD into different clusters. Then, the expression of 22 m^6A regulators in different clusters was analyzed. Subsequently, we analyzed clinicopathological features and survival among early-stage LUAD patients in different clusters. Principal component analysis (PCA) was used to evaluate gene expression patterns between different clusters.

Immune cell infiltration in TIME

We obtained gene set from Charoentong's study that marked each TIME infiltration immune cell type [11, 19]. The enrichment score calculated by single-sample gene-set enrichment analysis (ssGSEA) was applied to represent the relative abundance of each TIME infiltration immune cell in each sample [20, 21]. The R "ESTIMATE" package was used to obtain the immune score, tumor purity and ESTIMATE score of patients [22]. The expression of PD-L1 in clusterA and cluster was tested. The R "GSVA" package was used for gene set variation analysis (GSVA) [23].

Prognostic model

The correlation between each m^{6}A regulator and the prognosis of patients with early-stage LUAD was examined using univariate analysis (P<0.05). Least absolute shrinkage and selection operator (LASSO)-Cox regression analysis was applied to determine the prognostic model. The risk score of each sample is calculated as follows: $Risk\ Score = \sum_{i=1}^{n}(exp_i * \beta_i)$. Patients were categorized into high and low risk groups based on the median risk score. Subsequently, Kaplan-Meier (KM) analysis was carried out. The prognostic model's accuracy was evaluated using receiver operator characteristic (ROC) analysis. The area under the curve (AUC) is an evaluation index of model performance. The AUC ranges from 0 to 1, where 0.6-0.7 indicates sufficient diagnostic accuracy [24]. Moreover, we also used univariate and multivariate Cox regression analysis to determine whether risk score was an independent prognostic factor.

Risk score and immune cells

To research the correlation between the risk score and immune cell infiltration, we examined the relationship between the risk score and the infiltration level of 6 immune cells based on TIMER (https://cistrome.shinyapps.io/timer/). We also used the TIMER to evaluate the impact of somatic cell copy number alternations (CNAs) of m^{6}A methylation regulators on the infiltration level of immune cell. The GISTIC 2.0 database was utilized in the TIMER.

Validation of model genes expression in clinical tissue samples

Totally, 13 patients were included in this study, and the specific clinical information is displayed in Supplementary Table 2. Early-stage LUAD tissue and adjacent normal tissue were obtained for real-time polymerase chain reaction (RT-PCR) validation. Total RNA of samples was extracted using TRIzol kit. Subsequently, the FastKing cDNA first strand synthesis kit (TIANGEN, KR116) was used for reverse transcription. Finally, SuperReal PreMix Plus (SYBR Green) (TIANGEN, FP205) was used for RT-PCR validation. ACTB was used as internal reference. The data were calculated by $2^{-\Delta\Delta Ct}$ method [25].

Validation of *in vitro* cell experiments

Cell line NCI-H1299 (Procell, CL-0165) was used for cell validation. The model gene HNRNPC was selected for knockdown and overexpression in NCI-H1299 cells. Overexpression lentivirus (vector: GV492) and target gene RNA interference lentivirus (vector: GV493) were purchased from Shanghai Jikai Gene Technology Co. The effect of HNRNPC on the migration of NCI-H1299 cells was analyzed by cell scratch assay. The cell density of knockdown control group, knockdown group, over-expression control group and overexpression group was adjusted to 2×10^5 cells/well and seeded in 6-well plate, with 3 replicates in each group, and incubated at 37° C, 5% CO_2. After the cells had grown to a monolayer, scratches were made with the micro pipette tip, and washed gently twice with phosphate buffer saline (PBS) to remove the scratched cells. After that, the culture was continued, and samples were taken and photographed according to the culture time of 0h, 24h and 48h. In addition, the effect of HNRNPC on the migration and invasion of NCI-H1299 cells was detected by Transwell assay. For Transwell assay, the cell density was adjusted to 4×10^4 cells/well and seeded in 24-well plates. Cells on the bottom surface of the Transwell chamber were fixed with precooled ethanol for 30 min. Crystal violet solution was stained for 30 min and observed under a microscope.

Statistical analysis

Wilcoxon test was applied to statistically analyze the differential expression of genes and the infiltration level of immune cell. The log-rank test was used to analyze significant differences in survival between different groups in Kaplan-Meier. In RT-PCR, the data were statistically analyzed by T-test. R software (version 3.6.3) was used for all statistics.

Availability of data and materials

We searched for LUAD public gene expression data and complete clinical annotations from GEO (http://www.ncbi.nlm.nih.gov/geo) and TCGA (https://tcga-data.nci.nih.gov/tcga/) databases. The accession numbers are GEO: GSE31210 (Platform: Affymetrix Human Genome U133 Plus 2.0 Array) and TCGA: LUAD (Platform: Illumina RNAseq), respectively.

RESULTS

Genetic alterations of m^{6}A methylation regulators in patients with early-stage LUAD

In total, 22 m^{6}A methylation regulators (2 erasers, 12 readers and 8 writers) were identified. Differential expression analysis revealed that m^{6}A methylation regulators expression was abnormal in early-stage LUAD (Figure 1A, 1B). Moreover, 22 m^{6}A regulators were basically positively correlated (Figure 1C). Among 381 samples, 85 samples had m^{6}A methylation regulators mutations, and the mutation rate was 22.31%

(Figure 1D). In early-stage LUAD samples, the ZC3H13 had the highest mutation frequency, followed by IGF2BP1, while WTAP did not show any mutation. These results imply that genetic and expression alteration landscapes of m^{6}A methylation regulators are highly heterogeneous between normal and early-stage LUAD samples.

Clustering based on m^{6}A methylation regulators

In the process of increasing the number of clusters K from 2 to 5, when K = 2, there is a higher intra-group correlation and a lower inter-group correlation (Supplementary Figure 1A–1E). Similar results were verified in GEO (Supplementary Figure 1F–1J). Thus, patients in TCGA were divided into clusterA (264 early-stage LUAD samples) and clusterB (134 early-stage LUAD samples) according to K=2. Moreover, verification in GEO also found similar results (clusterA contains 143 early-stage LUAD samples; clusterB contains 83 early-stage LUAD samples). The results of differential expression analysis showed that the expression levels of m^{6}A methylation regulators in clusterB were higher than that in clusterA (Figure 2A). Then, the clinicopathological features between clusterA and clusterB were compared. The results showed that clusterB mainly contained male patients with early-stage LUAD and was associated with higher age and smoking frequency (Figure 2B). KM analysis in TCGA (Figure 2C) and GEO (Figure 2D) indicated that the

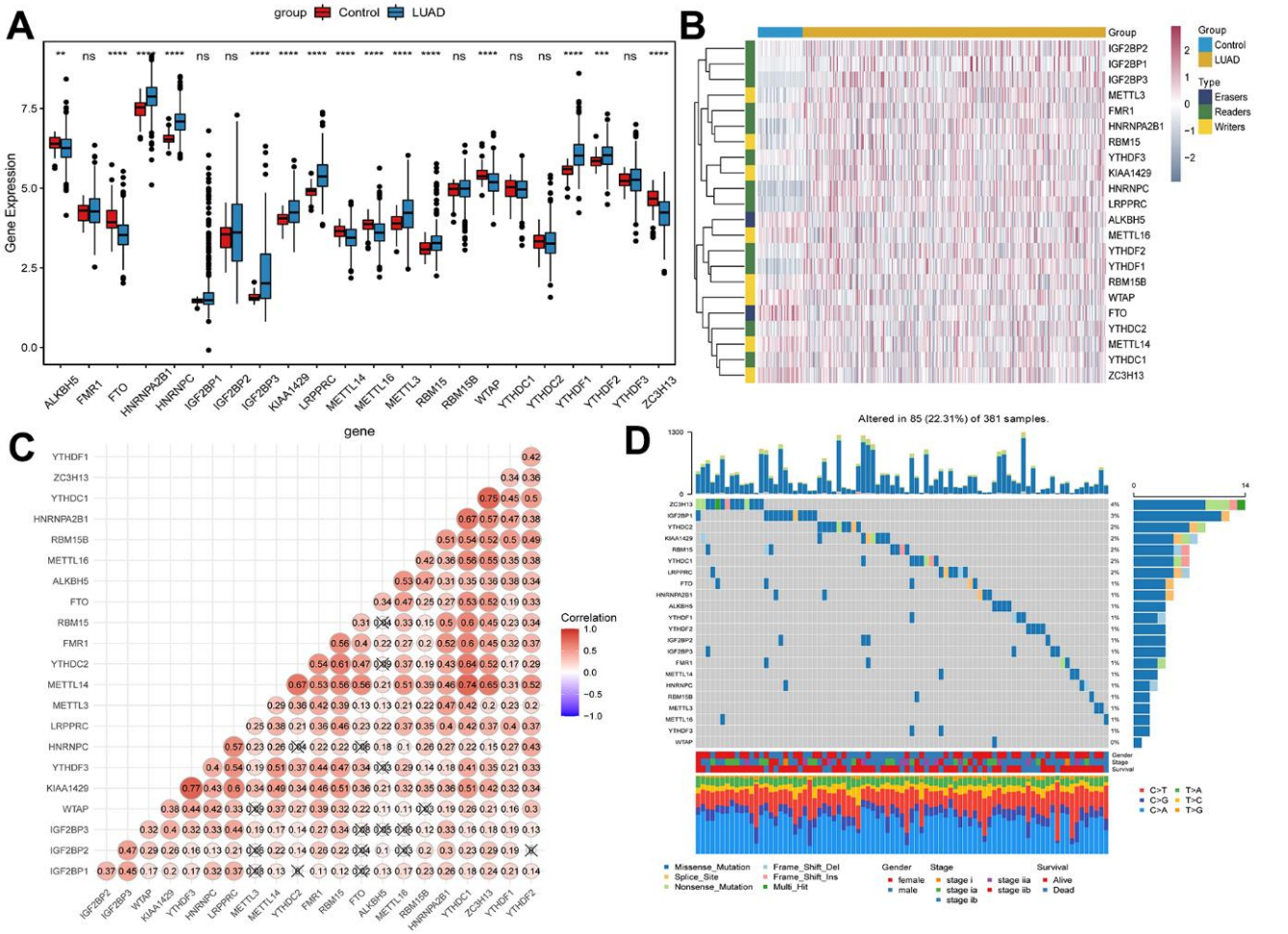


Figure 1. Expression of m^{6}A methylation regulators in early-stage LUAD. (**A**) Expression boxplot of m^{6}A methylation regulators in early-stage LUAD tissues and normal control tissues. *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001. (**B**) Heatmap of m^{6}A methylation regulators expression in early-stage LUAD tissues and normal control tissues. Complete-linkage method combined with Euclidean distance was used to construct clustering. (**C**) Correlation between m^{6}A methylation regulators. Red and blue represent positive and negative correlation, respectively. (**D**) Mutation frequency of m^{6}A methylation regulators in early-stage LUAD. The numbers and barplot on the right represent the mutation frequency of each m6A RNA methylation regulator and the proportion of each variant type, respectively.

survival time of patients in clusterB was significantly shorter than that in clusterA. We also used the PCA method to further analyze the gene expression patterns of clusterA and clusterB. Gene expression profiles between clusterA and clusterB were well distinguished (Figure 2E). Our data showed that the cluster subtypes defined by the expression of the m^{6}A regulators were closely related to the heterogeneity of early-stage LUAD patients.

Characteristics of immune cell infiltration in TIME

To investigate the degree of immune cell infiltration in TIME of clusterA and clusterB, we performed the ssGSEA analysis. The degree of immune cell (for example, Activated CD8 T cell and Immature B cell) infiltration in TIME was significantly reduced in clusterB (Figure 3A). Moreover, the immune and ESTIMATE scores were significantly lower in clusterB subtype than in clusterA subtype, and the tumor purity was significantly higher (Figure 3B–3D). We also detected the expression of PD-L1 in clusterA and clusterB. The expression of PD-L1 in clusterA was significantly higher than that in clusterB (Figure 3E). In addition, GSVA results showed that pathways (cell adhesion molecules, PPAR signaling pathway and B cell receptor signaling pathway) related to immunity and inhibition of tumor progression were significantly enriched in clusterA compared with clusterB (Figure 3F).

Construction of m^{6}A regulators prognostic model

Univariate Cox analysis showed that 3 m^{6}A regulators (HNRNPC, IGF2BP1 and IGF2BP3) were significantly associated with the prognosis of early-stage LUAD (Figure 4A). We performed LASSO-Cox regression analysis for the 3 characteristic genes. A prognostic model consisting of HNRNPC, IGF2BP1 and IGF2BP3 was constructed (Figure 4B, 4C). Risk Score = (0.1940*HNRNPC) + (0.1836*IGF2BP1) + (0.0805* IGF2BP3). Subsequently, the patients were categorized into high and low risk groups. The survival time of patients in high risk group was lower (Figure 4D, 4E).

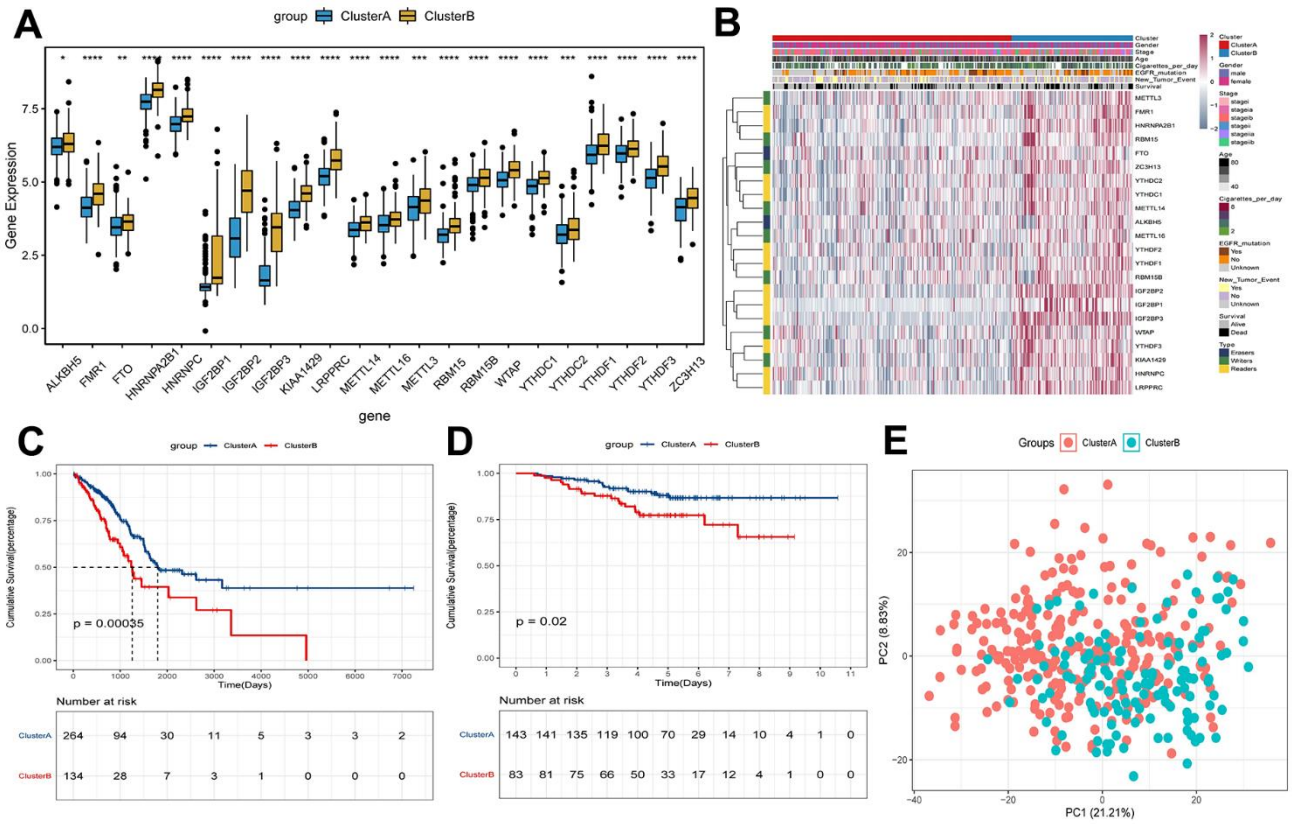


Figure 2. Survival and expression analysis of m^{6}A methylation regulators of early-stage LUAD in clusterA and clusterB subtypes. (**A**) Expression boxplot of expression of regulatory factors in clusterA and clusterB subtypes in TCGA. *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001. (**B**) Heatmap and clinicopathologic features of the clusterA and clusterB subtypes in TCGA. (**C**) Kaplan-Meier survival curves of clusterA and clusterB subtypes in TCGA; (**D**) Kaplan-Meier survival curves of clusterA and clusterB subtypes in GEO; (**E**) PCA was used to analyze gene expression patterns of clusterA and clusterB subtypes in TCGA.

The results of heat map indicated that HNRNPC, IGF2BP1 and IGF2BP3 were lowly expressed in the lower risk group (Figure 4F). Drugs related to HNRNPC, IGF2BP1 and IGF2BP3 were screened based on DGIdb database (https://dgidb.org/). The results showed that only one drug DABIGATRAN related to HNRNPC was screened (Supplementary Figure 2).

Additionally, patients in the high-risk group had a lower survival time, according to the KM analysis (Figure 5A). ROC curve analyses results showed that the AUC in 1, 3 and 5 years were all greater than 0.6 (Figure 5B). The results of KM analysis and ROC analysis in GEO (Figure 5C, 5D) and cBioPortal (https://www.cbioportal.org/) (Supplementary Figure 3A, 3B) databases are similar to those in TCGA. Decision curve analysis (DCA) was also performed based on TCGA. Prognostic indicator genes TP53 and NPM1 for LUAD and risk score were selected to perform DCA (Supplementary Figure 3C). The results of DCA showed that the predictive effect of risk score was better. Univariate analysis showed that T stage, N stage, TNM stage and risk score were correlated with overall survival of early-stage LUAD (Figure 5E). Moreover, multivariate analysis showed that TNM stage, risk score and overall survival were significantly correlated (Figure 5F). Additionally, risk score was significantly associated with survival, stage and gender (Figure 5G). In summary, the risk score may be an independent prognostic factor for early-stage LUAD.

Effect of risk score on immune cell infiltration

The risk score was negatively correlated with the infiltration level of activated B cell, activated CD8 T cell, macrophage, neutrophil and plasmacytoid dendritic

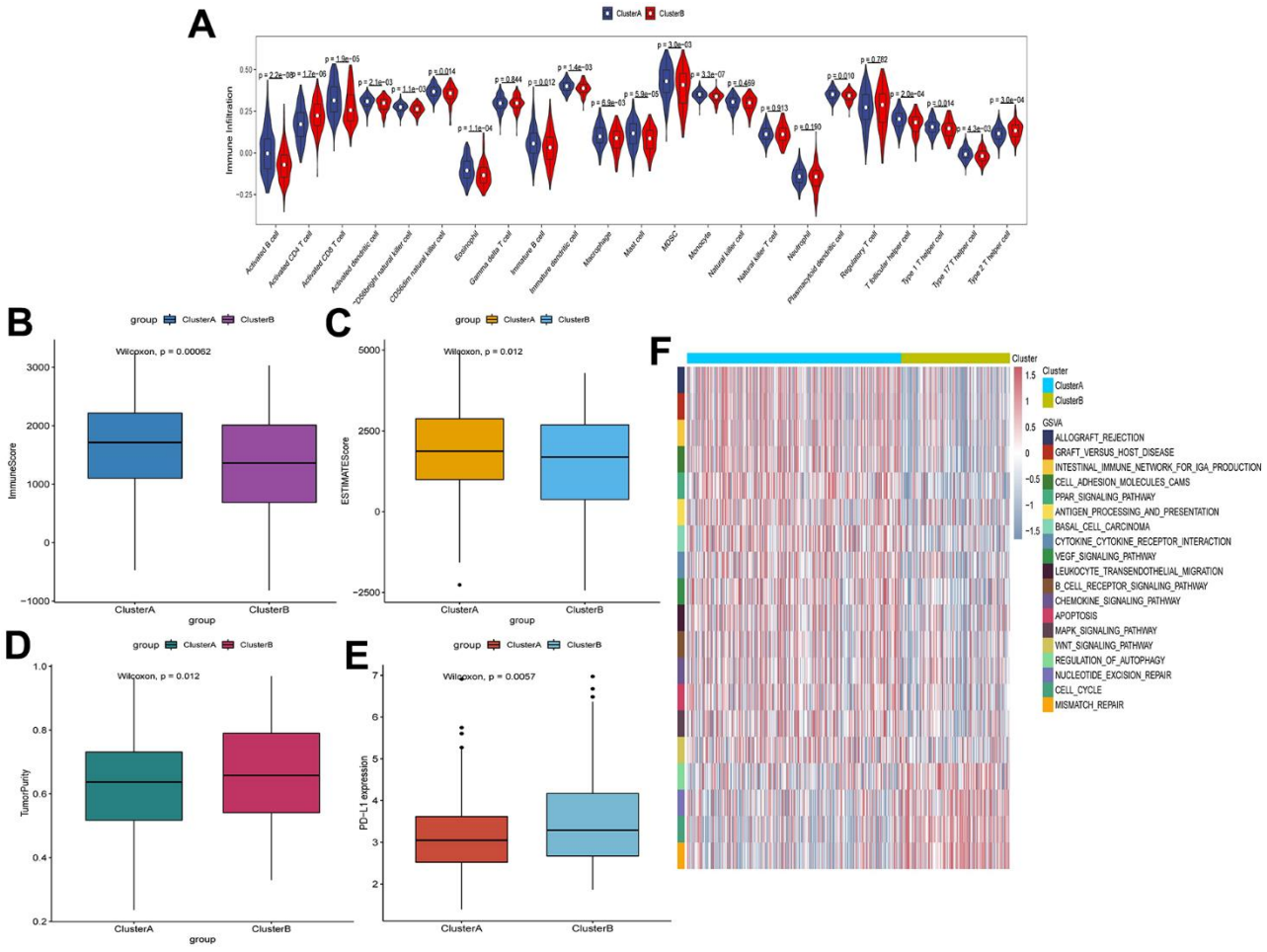


Figure 3. Characteristics of cell infiltration in TIME of clusterA and clusterB subtypes in TCGA. (**A**) The degree of immune cell infiltration in TIME of clusterA and clusterB subtypes; (**B**) Immunoscore in clusterA and clusterB subtypes; (**C**) ESTIMATEscore in clusterA and clusterB subtypes; (**D**) Tumor purity in clusterA and clusterB subtypes; (**E**) Expression level of PD-L1 in clusterA and clusterB subtypes; (**F**) GSVA enrichment analysis of clusterA and clusterB subtypes.

cell, and positively correlated with the infiltration level of activated CD4 T cell (Figure 6). This result confirmed that risk score based on HNRNPC, IGF2BP1 and IGF2BP3 was associated with the immune micro-environment of early-stage LUAD. Subsequently, we assessed the effect of CNAs of HNRNPC, IGF2BP1 and IGF2BP3 on the infiltration of immune cell. Somatic cell CNAs could affect the infiltration of immune cells (Figure 7).

Expression validation of HNRNPC, IGF2BP1 and IGF2BP3 in clinical tissue samples through RT-PCR

Primers of HNRNPC, IGF2BP1 and IGF2BP3 are shown in Supplementary Table 3. Expression levels of HNRNPC, IGF2BP1 and IGF2BP3 were significantly up-regulated in early-stage LUAD tissues compared with adjacent normal tissues (Figure 8). This result was consistent with bioinformatics analysis. This further indicates that HNRNPC, IGF2BP1 and IGF2BP3 may play important regulatory roles in the pathological mechanism of early-stage LUAD.

Effect of HNRNPC on cell migration and invasion

Cell scratch assay showed that overexpression of HNRNPC increased the migration ability of NCI-H1299 cells, while knockdown of HNRNPC decreased the migration ability (Figure 9 and Supplementary Figure 4A). Transwell assay also showed that overexpression of HNRNPC increased the migration ability of NCI-H1299 cells, while knockdown of HNRNPC decreased the migration ability (Figure 10 and Supplementary Figure 4B). In addition, Transwell assay also showed that overexpression of HNRNPC increased the invasion ability of NCI-H1299 cells, while knockdown of HNRNPC reduced the invasion ability (Figure 10 and Supplementary Figure 4C).

DISCUSSION

With the progress of science and technology, more and more evidences indicate that m^{6}A regulators play an important role in inflammation, immunity and anti-tumor [26, 27]. Exploring the role of m^{6}A RNA regulators in TIME has become a hot research topic. LUAD is a common pulmonary malignant disease with a poor prognosis [28]. Therefore, exploring the effect of m^{6}A regulators in the TIME of early-stage LUAD is beneficial for early management. In our study, we selected 22 m^{6}A regulators for research in early-stage LUAD. Compared with normal controls, these m^{6}A regulators showed obvious mutations and expression heterogeneity in early-stage LUAD.

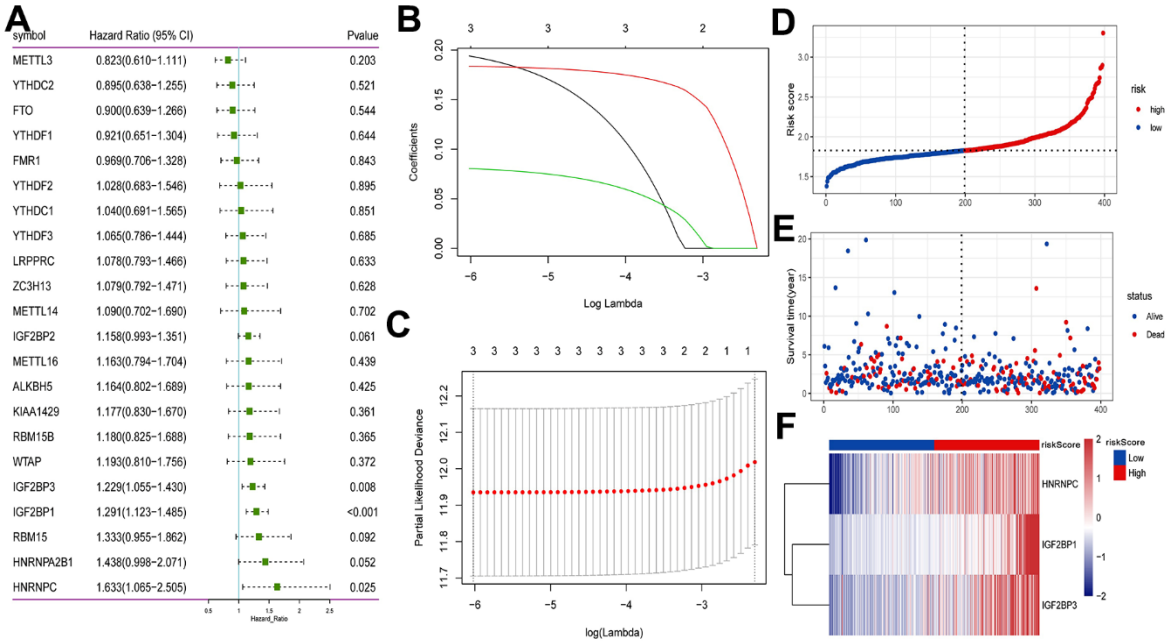


Figure 4. Construction of prognostic model based on m^{6}A methylation regulators in TCGA. (**A**) Univariate Cox analysis screened out the m^{6}A RNA methylation regulators that were significantly correlated with the prognosis of patients with early-stage LUAD. (**B**) The LASSO regression model was constructed; (**C**) the optimal value of the lambda penalty parameter was determined by 10-cross validation; (**D**) risk score distribution of early-stage LUAD. Red and blue represent high risk score and low risk score, respectively; (**E**) distribution of overall survival status of early-stage LUAD; (**F**) heat map of HNRNPC, IGF2BP1 and IGF2BP3. Red indicates above the reference channel. Blue indicates below the reference channel.

Expression levels of m^6A regulators in clusterB were higher than that in clusterA. KM analysis indicated that the survival time of patients in clusterB was significantly shorter. Notably, there are significant differences of TIME between clusterA and clusterB. In clusterB, infiltration degree of immune cells was significantly reduced. The higher the degree of infiltration of CD8 T cells and other immune cells in the TIME may indicate a better prognosis [29]. In addition, clusterB had lower immune and ESTIMATE scores, and higher tumor purity. Previous studies have indicated higher immune and ESTIMATE scores and lower tumor purity were associated with higher prognosis [30, 31]. We also studied the difference of PD-L1 in clusterA and clusterB. The expression of PD-L1 in clusterA was lower than that in clusterB. PD-1/PD-L1 can participate in regulating the immune escape of tumor cells [32]. The abnormally high expression of PD-L1 in TIME may be associated with activation of various carcinogenic signals [33]. These results indicate that the poor prognosis of clusterB may be related to immune cell infiltration and PD-L1 expression level.

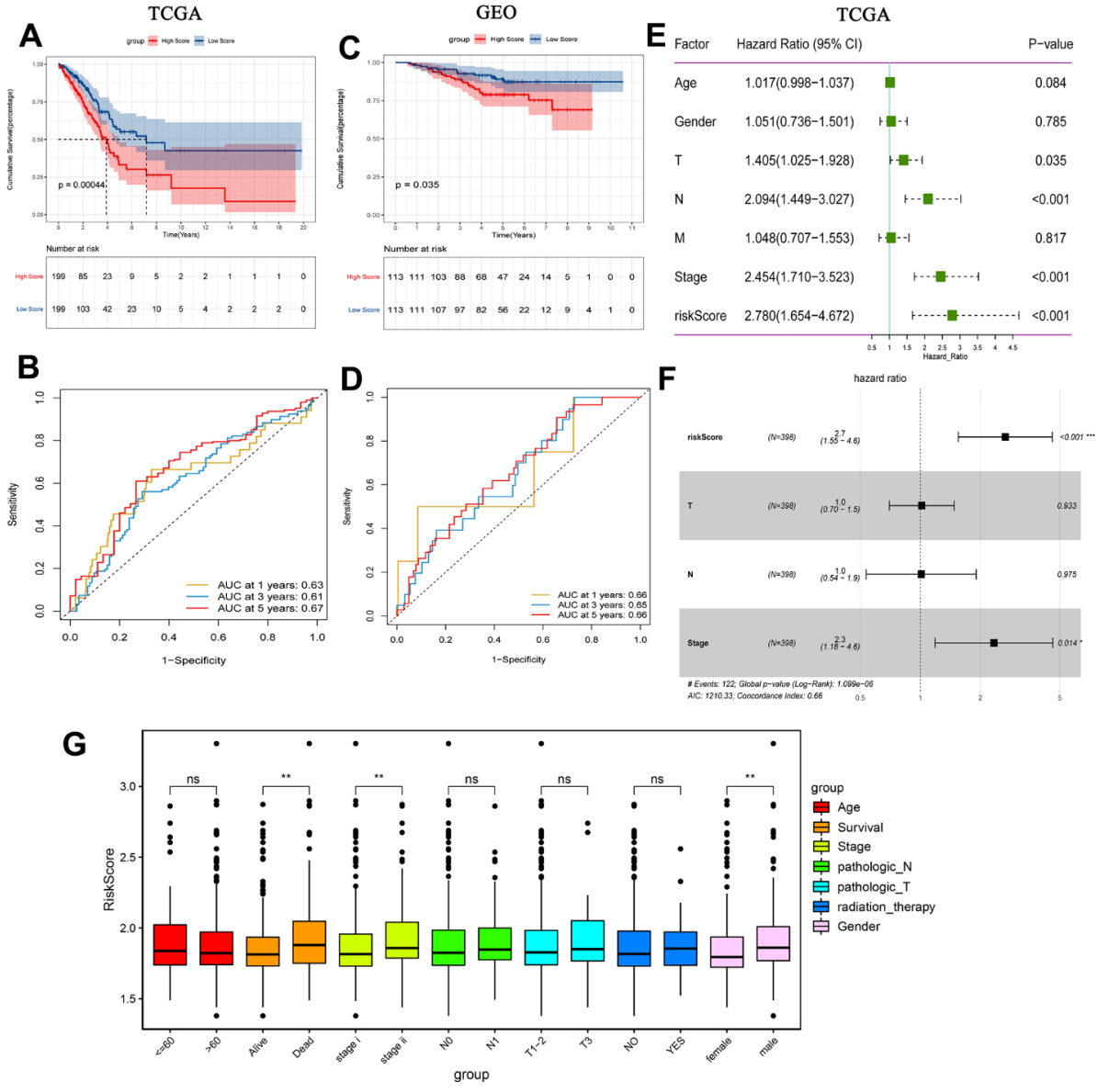


Figure 5. Prognostic analysis of risk score. (**A**) Kaplan-Meier survival analysis of early-stage LUAD in TCGA high and low risk score group; (**B**) Time-dependent ROC analysis measuring the predictive value of the risk score in TCGA; (**C**) Kaplan-Meier survival analysis of early-stage LUAD in GEO high and low risk score group; (**D**) Time-dependent ROC analysis measure the predictive value of the risk score in GEO; (**E**) Univariate Cox analysis confirmed that risk score was associated with overall survival in TCGA; (**F**) Multivariate Cox analysis found that risk score was an independent prognostic factor for early-stage LUAD in TCGA; (**G**) Correlation between risk score and clinical characteristics.

The GSVA results showed that pathways related to immunity and inhibition of tumor progression were significantly enriched in clusterA compared with clusterB. Low cell adhesion molecules expression predicts poor prognosis of LUAD [34]. B cell receptor signaling pathway is an important pathway through which B cells recognize antigen and initiate immune response [35]. The up-regulation of PPAR signaling pathway can promote apoptosis of colorectal cancer cells and inhibit tumor progression [36]. PPAR signaling pathway also plays a vital regulatory role in LUAD [37]. Thus, we hypothesized that m^6A regulators may play a regulatory role in the TIME of early-stage LUAD by regulating related signaling pathways.

A prognostic model was also constructed. The prognostic model consisted of 3 m^6A regulators (HNRNPC, IGF2BP1 and IGF2BP3). It has been found that silencing heterogeneous nuclear ribonucleoprotein C (HNRNPC) can inhibit the proliferation and metastasis of prostate cancer cells. Furthermore, HNRNPC expression is negatively correlated with the levels of most immune cell infiltration in prostate cancer [38]. HNRNPC expression is lower in the high risk group of patients with lung squamous cell carcinoma and is more sensitive to immunotherapy and chemotherapy [39]. HNRNPC is significantly dysregulated and affects immune pathways and immune cell infiltration of endometriosis [40]. HNRNPC is also highly expressed in pancreatic adenocarcinoma and esophageal squamous cell carcinoma (ESCC), and can regulate the TIME [41, 42]. Herein, HNRNPC was highly expressed in early-stage LUAD tissues. Knockdown of HNRNPC inhibited the migration and invasion of NCI-H1299. Drug prediction results showed that DABIGATRAN was related to HNRNPC. Previous studies have shown that DABIGATRAN has some therapeutic efficacy in breast cancer, glioblastoma and pancreatic cancer [43, 44]. A case report indicates that DABIGATRAN was also involved in the treatment of a case of LUAD-associated hypercoagulability leading to venous limb gangrene [45]. The identification of DABIGATRAN provides new perspectives on the treatment of early-stage LUAD. Insulin like growth factor 2 mRNA binding protein (IGF2BP) 1 and IGF2BP3 belong to the same protein family. IGF2BP1 can regulate the invasion, migration, growth and proliferation of hepatocellular carcinoma (HCC) cells by long non-coding RNA LIN28B-AS1 regulation [46]. Increased expression of IGF2BP1 in endometrial cancer (EC) can promote cell proliferation

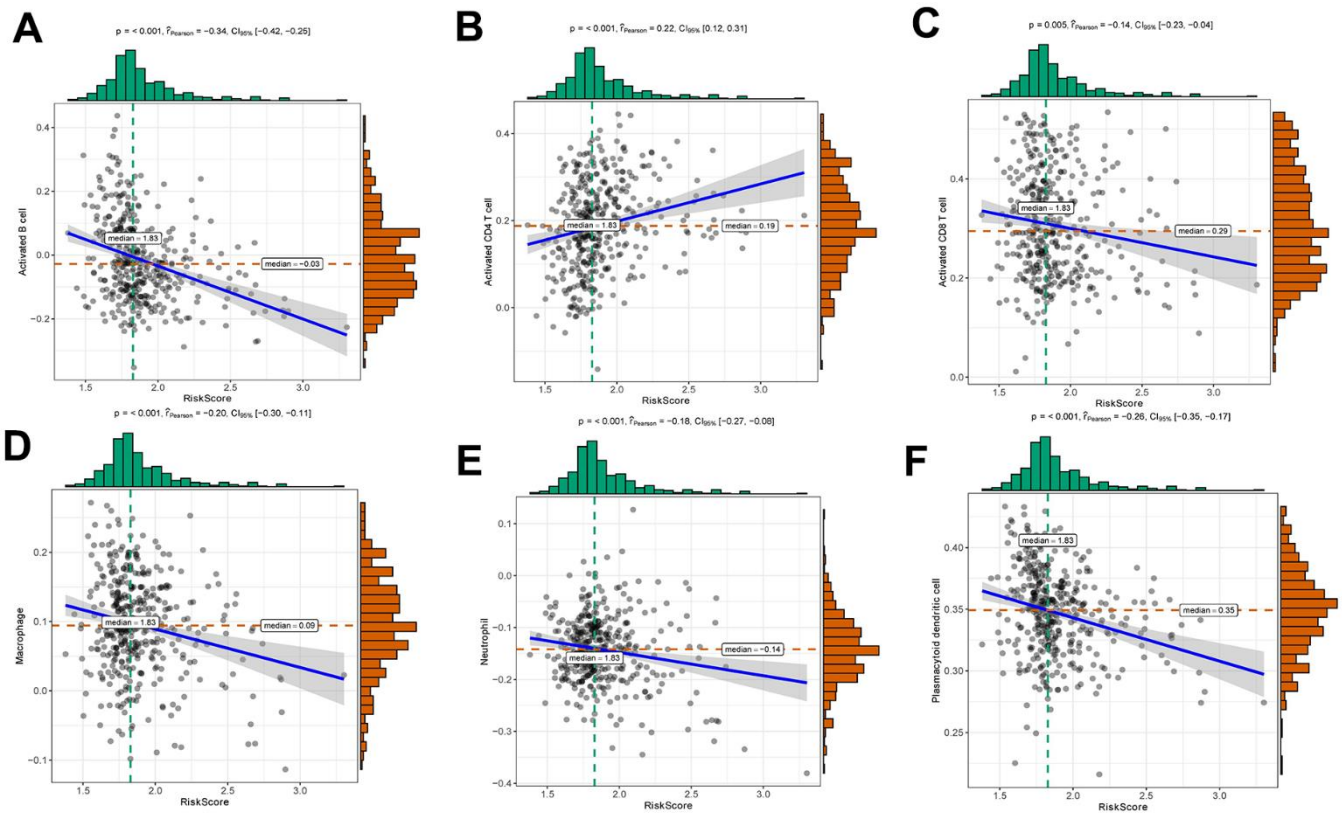


Figure 6. Relationships between risk score and activated B cell (**A**), activated CD4 T cell (**B**), activated CD8 T cell (**C**), macrophage (**D**), neutrophil (**E**) and plasmacytoid dendritic cell (**F**).

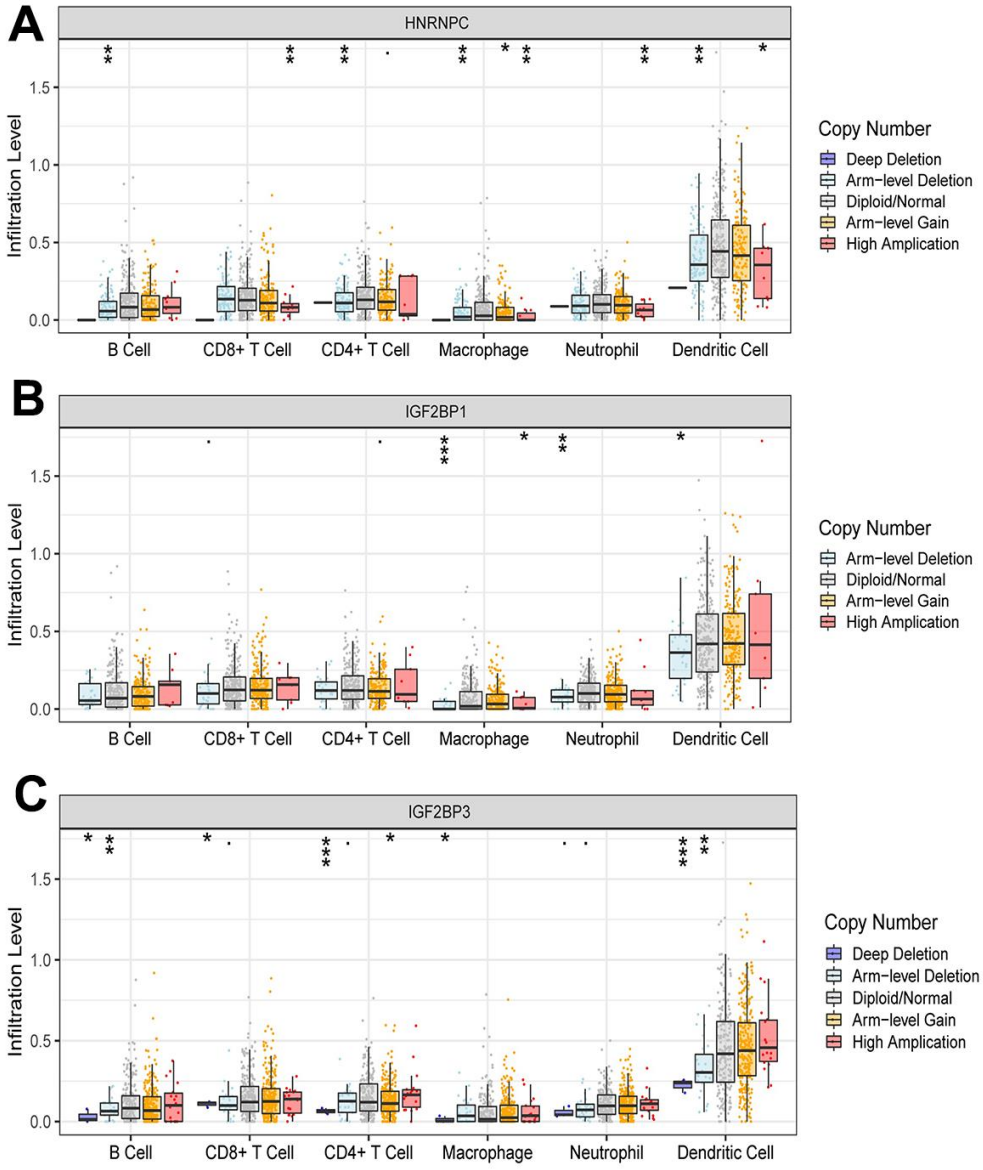


Figure 7. Effect of CNAs of HNRNPC (**A**), IGF2BP1 (**B**) and IGF2BP3 (**C**) on immune cell infiltration level. *, P < 0.05; **, P < 0.01; ***, P < 0.001.

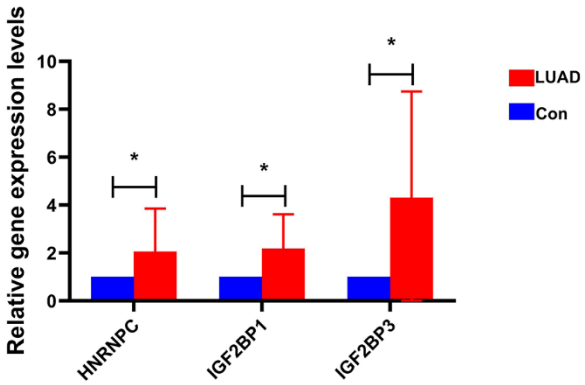


Figure 8. The expression of HNRNPC, IGF2BP1 and IGF2BP3 in early-stage LUAD was verified by RT-PCR. * represents P<0.05.

and regulate tumor progression [47]. In ovarian cancer, expression of IGF2BP1 is negatively correlated with immune cell infiltration [48]. Inhibition of IGF2BP1 also inhibits the proliferation and migration of NSCLC cells [49]. In addition, high expression of IGF2BP1 in LUAD is related to poorer prognosis [50]. Down-regulation of IGF2BP3 in colon cancer can inhibit DNA replication and angiogenesis [51]. IGF2BP3 can also promote the occurrence of bladder cancer [52]. IGF2BP3 promotes the proliferation and migration of melanoma cells and is associated with the infiltration level of immune cell in the TIME [53]. In addition, IGF2BP3 is also related to the expression of immunomodulators and immune infiltration in NSCLC [54]. These findings suggest that HNRNPC, IGF2BP1 and IGF2BP3 play key regulatory roles in the development of tumors. It also implies that the potential functions of HNRNPC, IGF2BP1 and IGF2BP3 in early-stage LUAD are worthy of further study.

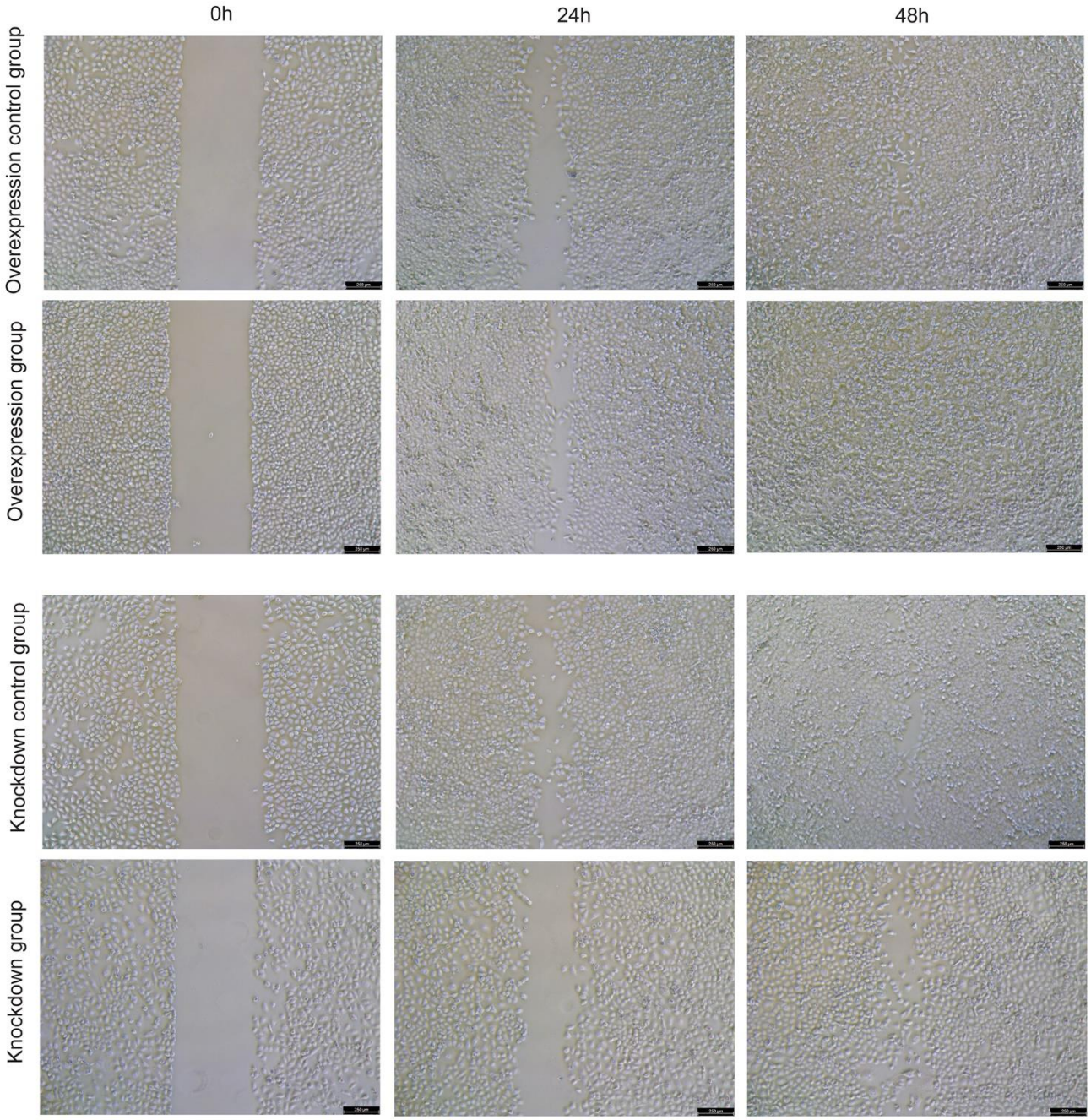


Figure 9. The effect of overexpression and knockdown of HNRNPC on the migration of NCI-H1299 cells was detected by cell scratch assay at 24 h and 48 h.

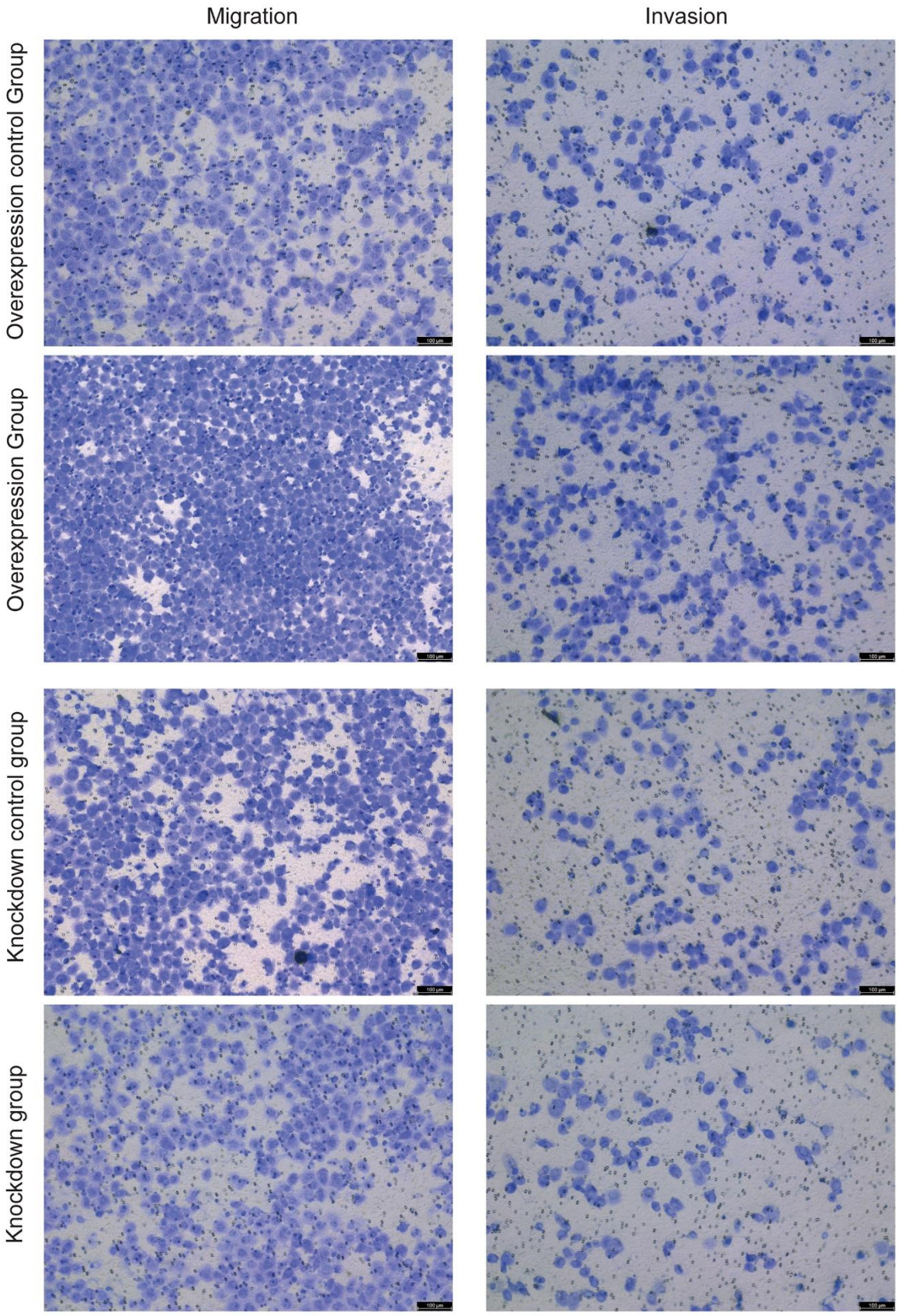


Figure 10. The effects of overexpression and knockdown of HNRNPC on the migration (left) and invasion (right) of NCI-H1299 cells were detected by cell scratch assay.

The relationship between immune cell infiltration and risk score was analyzed. Immune cell infiltration is associated with the prognosis of digestive system tumors and LUAD [55, 56]. Increased levels of CD8 T cell infiltration are associated with improved prognosis in NSCLC [57]. In this study, risk score was correlated with the level of immune cell infiltration. In addition, the somatic cell CNAs of m^6A methylation regulators were found to influence the infiltration of immune cells. Thus, we hypothesized that m^6A regulators may play a role in the prognosis of early-stage LUAD by regulating the TIME.

It is undeniable that our research has certain limitations. First of all, the data of this study were obtained through public databases, so a vast number of clinical sample data are needed for verification. Thus, a vast number of clinical sample data must be collected for further verification in the later stage. Second, our study only preliminarily revealed the correlation between m^6A methylation regulators and prognosis and TIME of early-stage LUAD, and its specific molecular mechanism needs to be further studied.

CONCLUSIONS

Expression analysis revealed that expression of m^6A methylation regulators was abnormal in early-stage LUAD. Patients with early-stage LUAD were divided into two subtypes (clusterA and clusterB) based on the consensus clustering. The two subtypes have significant differences in TIME, prognosis, PD-L1 expression level and m^6A regulators expression level. A prognostic model consisting of HNRNPC, IGF2BP1 and IGF2BP3 could be used to predict the prognosis of early-stage LUAD. Moreover, risk score was associated with infiltration level of immune cell in early-stage LUAD. Remarkably, overexpression of HNRNPC promotes the migration and invasion of NCI-H1299 cells, while knockdown HNRNPC inhibits the migration and invasion of NCI-H1299 cells.

AUTHOR CONTRIBUTIONS

Conception and design: M.H., C.L.; Administrative support: M.H.; Provision of study materials or patients: C.Z., J.L.; Collection and assembly of data: Y.Z.; Experiments: M.H., Y.Z.; Data analysis and interpretation: M.H., Y.Z., C.L., C.Z., G.Y., J.L., H.Y., H.H., X.C. All authors reviewed the manuscript.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ETHICAL STATEMENT AND CONSENT

This study was approved by the Ethics Committee of People's Hospital of Deyang (LWH-OP-006-A04-V2.0). This study complied with the Declaration of Helsinki. All methods were carried out in accordance with relevant guidelines and regulations. Written informed consent was obtained from all participants.

FUNDING

No funding was provided for this study.

REFERENCES

1. Gridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N, Mok T, Petrella F, Spaggiari L, Rosell R. Non-small-cell lung cancer. Nat Rev Dis Primers. 2015; 1:15009. https://doi.org/10.1038/nrdp.2015.9 PMID:27188576

2. Sun GZ, Zhao TW. Lung adenocarcinoma pathology stages related gene identification. Math Biosci Eng. 2019; 17:737–46. https://doi.org/10.3934/mbe.2020038 PMID:31731374

3. Denisenko TV, Budkevich IN, Zhivotovsky B. Cell death-based treatment of lung adenocarcinoma. Cell Death Dis. 2018; 9:117. https://doi.org/10.1038/s41419-017-0063-y PMID:29371589

4. Bao X, Shi R, Zhao T, Wang Y. Immune landscape and a novel immunotherapy-related gene signature associated with clinical outcome in early-stage lung adenocarcinoma. J Mol Med (Berl). 2020; 98:805–18. https://doi.org/10.1007/s00109-020-01908-9 PMID:32333046

5. Skřičková J, Kadlec B, Venclíček O, Merta Z. Lung cancer. Cas Lek Cesk. 2018; 157:226–36. PMID:30441934

6. Sun J, Zhao T, Zhao D, Qi X, Bao X, Shi R, Su C. Development and validation of a hypoxia-related gene signature to predict overall survival in early-stage lung adenocarcinoma patients. Ther Adv Med Oncol. 2020; 12:1758835920937904. https://doi.org/10.1177/1758835920937904 PMID:32655701

7. Chen H, Carrot-Zhang J, Zhao Y, Hu H, Freeman SS, Yu S, Ha G, Taylor AM, Berger AC, Westlake L, Zheng Y, Zhang J, Ramachandran A, et al. Genomic and immune profiling of pre-invasive lung adenocarcinoma. Nat Commun. 2019; 10:5472. https://doi.org/10.1038/s41467-019-13460-3 PMID:31784532

8. Lavin Y, Kobayashi S, Leader A, Amir ED, Elefant N, Bigenwald C, Remark R, Sweeney R, Becker CD, Levine JH, Meinhof K, Chow A, Kim-Shulze S, et al. Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. Cell. 2017; 169:750–65.e17. https://doi.org/10.1016/j.cell.2017.04.014 PMID:28475900

9. Yang Y, Hsu PJ, Chen YS, Yang YG. Dynamic transcriptomic m^{6}A decoration: writers, erasers, readers and functions in RNA metabolism. Cell Res. 2018; 28:616–24. https://doi.org/10.1038/s41422-018-0040-8 PMID:29789545

10. Li N, Kang Y, Wang L, Huff S, Tang R, Hui H, Agrawal K, Gonzalez GM, Wang Y, Patel SP, Rana TM. ALKBH5 regulates anti-PD-1 therapy response by modulating lactate and suppressive immune cell accumulation in tumor microenvironment. Proc Natl Acad Sci USA. 2020; 117:20159–70. https://doi.org/10.1073/pnas.1918986117 PMID:32747553

11. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. m^{6}A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. Mol Cancer. 2020; 19:53. https://doi.org/10.1186/s12943-020-01170-0 PMID:32164750

12. Wang L, Hui H, Agrawal K, Kang Y, Li N, Tang R, Yuan J, Rana TM. m^6 A RNA methyltransferases METTL3/14 regulate immune responses to anti-PD-1 therapy. EMBO J. 2020; 39:e104514. https://doi.org/10.15252/embj.2020104514 PMID:32964498

13. Lin S, Choe J, Du P, Triboulet R, Gregory RI. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. Mol Cell. 2016; 62:335–45. https://doi.org/10.1016/j.molcel.2016.03.021 PMID:27117702

14. Xu F, Chen JX, Yang XB, Hong XB, Li ZX, Lin L, Chen YS. Analysis of Lung Adenocarcinoma Subtypes Based on Immune Signatures Identifies Clinical Implications for Cancer Therapy. Mol Ther Oncolytics. 2020; 17:241–9. https://doi.org/10.1016/j.omto.2020.03.021 PMID:32346613

15. Zhou B, Gao S. Comprehensive Analysis of Clinical Significance, Immune Infiltration and Biological Role of m^{6}A Regulators in Early-Stage Lung Adenocarcinoma. Front Immunol. 2021; 12:698236. https://doi.org/10.3389/fimmu.2021.698236 PMID:34650549

16. Wang S, Sun C, Li J, Zhang E, Ma Z, Xu W, Li H, Qiu M, Xu Y, Xia W, Xu L, Yin R. Roles of RNA methylation by means of N^6-methyladenosine (m^{6}A) in human cancers. Cancer Lett. 2017; 408:112–20. https://doi.org/10.1016/j.canlet.2017.08.030 PMID:28867248

17. Kwok CT, Marshall AD, Rasko JE, Wong JJ. Genetic alterations of m^{6}A regulators predict poorer survival in acute myeloid leukemia. J Hematol Oncol. 2017; 10:39. https://doi.org/10.1186/s13045-017-0410-6 PMID:28153030

18. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010; 26:1572–3. https://doi.org/10.1093/bioinformatics/btq170 PMID:20427518

19. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. 2017; 18:248–62. https://doi.org/10.1016/j.celrep.2016.12.019 PMID:28052254

20. Ye L, Zhang T, Kang Z, Guo G, Sun Y, Lin K, Huang Q, Shi X, Ni Z, Ding N, Zhao KN, Chang W, Wang J, et al. Tumor-Infiltrating Immune Cells Act as a Marker for Prognosis in Colorectal Cancer. Front Immunol. 2019; 10:2368. https://doi.org/10.3389/fimmu.2019.02368 PMID:31681276

21. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005; 102:15545–50. https://doi.org/10.1073/pnas.0506580102 PMID:16199517

22. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013; 4:2612. https://doi.org/10.1038/ncomms3612 PMID:24113773

23. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14:7. https://doi.org/10.1186/1471-2105-14-7 PMID:23323831

24. Šimundić AM. Measures of Diagnostic Accuracy: Basic Definitions. EJIFCC. 2009; 19:203–11. PMID:27683318

25. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001; 25:402–8. https://doi.org/10.1006/meth.2001.1262 PMID:11846609

26. Han D, Liu J, Chen C, Dong L, Liu Y, Chang R, Huang X, Liu Y, Wang J, Dougherty U, Bissonnette MB, Shen B, Weichselbaum RR, et al. Anti-tumour immunity controlled through mRNA m^6A methylation and YTHDF1 in dendritic cells. Nature. 2019; 566:270–4. https://doi.org/10.1038/s41586-019-0916-x PMID:30728504

27. Xie L, Dai R, Wang X, Xie G, Gao Z, Xu X. Comprehensive Analysis Revealed the Potential Implications of m6A Regulators in Lung Adenocarcinoma. Front Mol Biosci. 2022; 9:806780. https://doi.org/10.3389/fmolb.2022.806780 PMID:35419413

28. Yi M, Li A, Zhou L, Chu Q, Luo S, Wu K. Immune signature-based risk stratification and prediction of immune checkpoint inhibitor's efficacy for lung adenocarcinoma. Cancer Immunol Immunother. 2021; 70:1705–19. https://doi.org/10.1007/s00262-020-02817-z PMID:33386920

29. Tekpli X, Lien T, Røssevold AH, Nebdal D, Borgen E, Ohnstad HO, Kyte JA, Vallon-Christersson J, Fongaard M, Due EU, Svartdal LG, Sveli MA, Garred Ø, et al, and OSBREAC. An independent poor-prognosis subtype of breast cancer defined by a distinct tumor immune microenvironment. Nat Commun. 2019; 10:5499. https://doi.org/10.1038/s41467-019-13329-5 PMID:31796750

30. Sun S, Guo W, Wang Z, Wang X, Zhang G, Zhang H, Li R, Gao Y, Qiu B, Tan F, Gao Y, Xue Q, Gao S, He J. Development and validation of an immune-related prognostic signature in lung adenocarcinoma. Cancer Med. 2020; 9:5960–75. https://doi.org/10.1002/cam4.3240 PMID:32592319

31. Song C, Guo Z, Yu D, Wang Y, Wang Q, Dong Z, Hu W. A Prognostic Nomogram Combining Immune-Related Gene Signature and Clinical Factors Predicts Survival in Patients With Lung Adenocarcinoma. Front Oncol. 2020; 10:1300. https://doi.org/10.3389/fonc.2020.01300 PMID:32850406

32. Dermani FK, Samadi P, Rahmani G, Kohlan AK, Najafi R. PD-1/PD-L1 immune checkpoint: Potential target for cancer therapy. J Cell Physiol. 2019; 234:1313–25. https://doi.org/10.1002/jcp.27172 PMID:30191996

33. Ai L, Xu A, Xu J. Roles of PD-1/PD-L1 Pathway: Signaling, Cancer, and Beyond. Adv Exp Med Biol. 2020; 1248:33–59. https://doi.org/10.1007/978-981-15-3266-5_3 PMID:32185706

34. Ramasami S, Kerr KM, Chapman AD, King G, Cockburn JS, Jeffrey RR. Expression of CD44v6 but not E-cadherin or beta-catenin influences prognosis in primary pulmonary adenocarcinoma. J Pathol. 2000; 192:427–32. https://doi.org/10.1002/1096-9896(2000)9999:9999<::AID-PATH741>3.0.CO;2-Z PMID:11113858

35. Tanaka S, Baba Y. B Cell Receptor Signaling. Adv Exp Med Biol. 2020; 1254:23–36. https://doi.org/10.1007/978-981-15-3532-1_2 PMID:32323266

36. Zhang X, Yao J, Shi H, Gao B, Zhang L. LncRNA TINCR/microRNA-107/CD36 regulates cell proliferation and apoptosis in colorectal cancer via PPAR signaling pathway based on bioinformatics analysis. Biol Chem. 2019; 400:663–75. https://doi.org/10.1515/hsz-2018-0236 PMID:30521471

37. Meng X, Lu P, Bai H, Xiao P, Fan Q. Transcriptional regulatory networks in human lung adenocarcinoma. Mol Med Rep. 2012; 6:961–6. https://doi.org/10.3892/mmr.2012.1034 PMID:22895549

38. Wang S, Xu G, Chao F, Zhang C, Han D, Chen G. HNRNPC Promotes Proliferation, Metastasis and Predicts Prognosis in Prostate Cancer. Cancer Manag Res. 2021; 13:7263–76. https://doi.org/10.2147/CMAR.S330713 PMID:34584453

39. Xu F, Zhang H, Chen J, Lin L, Chen Y. Immune signature of T follicular helper cells predicts clinical prognostic and therapeutic impact in lung squamous cell carcinoma. Int Immunopharmacol. 2020; 81:105932. https://doi.org/10.1016/j.intimp.2019.105932 PMID:31836430

40. Jiang L, Zhang M, Wu J, Wang S, Yang X, Yi M, Zhang X, Fang X. Exploring diagnostic m6A regulators in endometriosis. Aging (Albany NY). 2020; 12:25916–38. https://doi.org/10.18632/aging.202163 PMID:33232273

41. Tang R, Zhang Y, Liang C, Xu J, Meng Q, Hua J, Liu J, Zhang B, Yu X, Shi S. The role of m6A-related genes in the prognosis and immune microenvironment of pancreatic adenocarcinoma. PeerJ. 2020; 8:e9602.

https://doi.org/10.7717/peerj.9602
PMID:33062408

42. Guo W, Tan F, Huai Q, Wang Z, Shao F, Zhang G, Yang Z, Li R, Xue Q, Gao S, He J. Comprehensive Analysis of PD-L1 Expression, Immune Infiltrates, and m6A RNA Methylation Regulators in Esophageal Squamous Cell Carcinoma. Front Immunol. 2021; 12:669750. https://doi.org/10.3389/fimmu.2021.669750 PMID:34054840

43. Vianello F, Sambado L, Goss A, Fabris F, Prandoni P. Dabigatran antagonizes growth, cell-cycle progression, migration, and endothelial tube formation induced by thrombin in breast and glioblastoma cell lines. Cancer Med. 2016; 5:2886–98. https://doi.org/10.1002/cam4.857 PMID:27600331

44. Shi K, Damhofer H, Daalhuisen J, Ten Brink M, Richel DJ, Spek CA. Dabigatran potentiates gemcitabine-induced growth inhibition of pancreatic cancer in mice. Mol Med. 2017; 23:13–23. https://doi.org/10.2119/molmed.2016.00214 PMID:28182192

45. Waite MMA, Martinelli AW, Preston SD, Gudgin E, Symington E, Rintoul RC, Peryt A, Coughlin P, Hayes P, Gilligan D, Besser M. A hypercoagulable state leading to venous limb gangrene associated with occult lung adenocarcinoma. Clin Case Rep. 2019; 7:888–92. https://doi.org/10.1002/ccr3.2106 PMID:31110709

46. Zhang J, Hu K, Yang YQ, Wang Y, Zheng YF, Jin Y, Li P, Cheng L. LIN28B-AS1-IGF2BP1 binding promotes hepatocellular carcinoma cell progression. Cell Death Dis. 2020; 11:741. https://doi.org/10.1038/s41419-020-02967-z PMID:32917856

47. Zhang L, Wan Y, Zhang Z, Jiang Y, Gu Z, Ma X, Nie S, Yang J, Lang J, Cheng W, Zhu L. IGF2BP1 overexpression stabilizes PEG10 mRNA in an m6A-dependent manner and promotes endometrial cancer progression. Theranostics. 2021; 11:1100–14. https://doi.org/10.7150/thno.49345 PMID:33391523

48. Wang Q, Zhang Q, Li Q, Zhang J, Zhang J. Clinicopathological and immunological characterization of RNA m^6 A methylation regulators in ovarian cancer. Mol Genet Genomic Med. 2021; 9:e1547. https://doi.org/10.1002/mgg3.1547 PMID:33225598

49. Gong F, Ren P, Zhang Y, Jiang J, Zhang H. MicroRNAs-491-5p suppresses cell proliferation and invasion by inhibiting IGF2BP1 in non-small cell lung cancer. Am J Transl Res. 2016; 8:485–95. PMID:27158341

50. Huang H, Wang D, Guo W, Zhuang X, He Y. Correlated low IGF2BP1 and FOXM1 expression predicts a good prognosis in lung adenocarcinoma. Pathol Res Pract. 2019; 215:152433. https://doi.org/10.1016/j.prp.2019.152433 PMID:31085008

51. Yang Z, Wang T, Wu D, Min Z, Tan J, Yu B. RNA N6-methyladenosine reader IGF2BP3 regulates cell cycle and angiogenesis in colon cancer. J Exp Clin Cancer Res. 2020; 39:203. https://doi.org/10.1186/s13046-020-01714-8 PMID:32993738

52. Huang W, Li Y, Zhang C, Zha H, Zhou X, Fu B, Guo J, Wang G. IGF2BP3 facilitates cell proliferation and tumorigenesis via modulation of JAK/STAT signalling pathway in human bladder cancer. J Cell Mol Med. 2020; 24:13949–60. https://doi.org/10.1111/jcmm.16003 PMID:33094561

53. Liu J, Zhou Z, Ma L, Li C, Lin Y, Yu T, Wei JF, Zhu L, Yao G. Effects of RNA methylation N6-methyladenosine regulators on malignant progression and prognosis of melanoma. Cancer Cell Int. 2021; 21:453. https://doi.org/10.1186/s12935-021-02163-9 PMID:34446007

54. Jin L, Chen C, Yao J, Yu Z, Bu L. The RNA N^6 -methyladenosine modulator HNRNPA2B1 is involved in the development of non-small cell lung cancer. Clin Exp Pharmacol Physiol. 2022; 49:329–40. https://doi.org/10.1111/1440-1681.13608 PMID:34717005

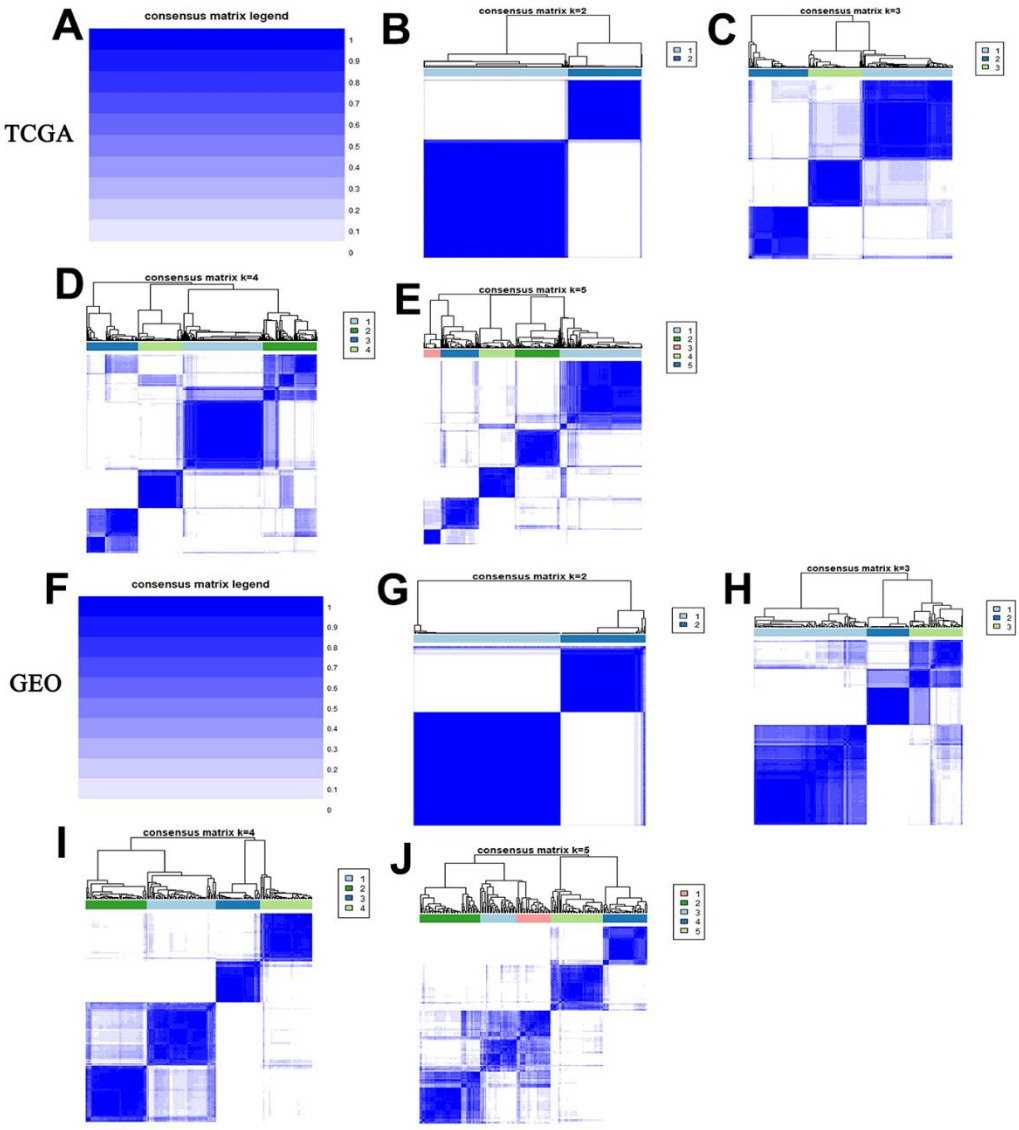
55. Yang S, Liu T, Cheng Y, Bai Y, Liang G. Immune cell infiltration as a biomarker for the diagnosis and prognosis of digestive system cancer. Cancer Sci. 2019; 110:3639–49. https://doi.org/10.1111/cas.14216 PMID:31605436

56. Wang W, Ren S, Wang Z, Zhang C, Huang J. Increased expression of TTC21A in lung adenocarcinoma infers favorable prognosis and high immune infiltrating level. Int Immunopharmacol. 2020; 78:106077. https://doi.org/10.1016/j.intimp.2019.106077 PMID:31812070

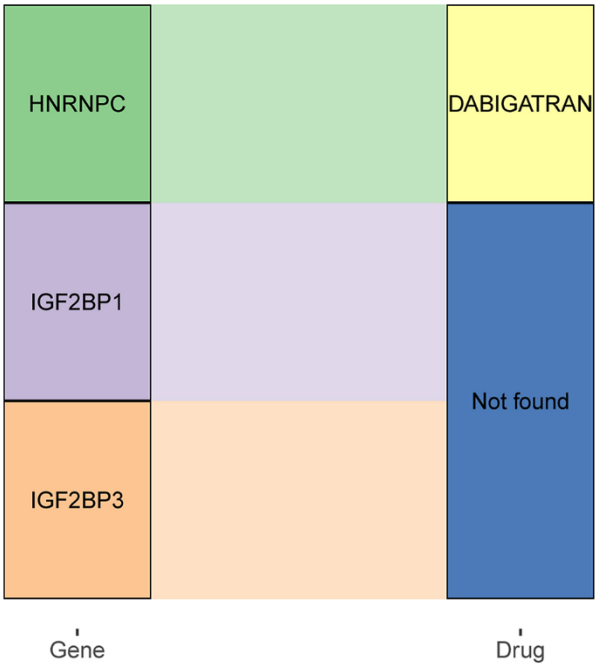
57. Yu Y, Zeng D, Ou Q, Liu S, Li A, Chen Y, Lin D, Gao Q, Zhou H, Liao W, Yao H. Association of Survival and Immune-Related Biomarkers With Immunotherapy in Patients With Non-Small Cell Lung Cancer: A Meta-analysis and Individual Patient-Level Analysis. JAMA Netw Open. 2019; 2:e196879. https://doi.org/10.1001/jamanetworkopen.2019.6879 PMID:31290993

SUPPLEMENTARY MATERIALS

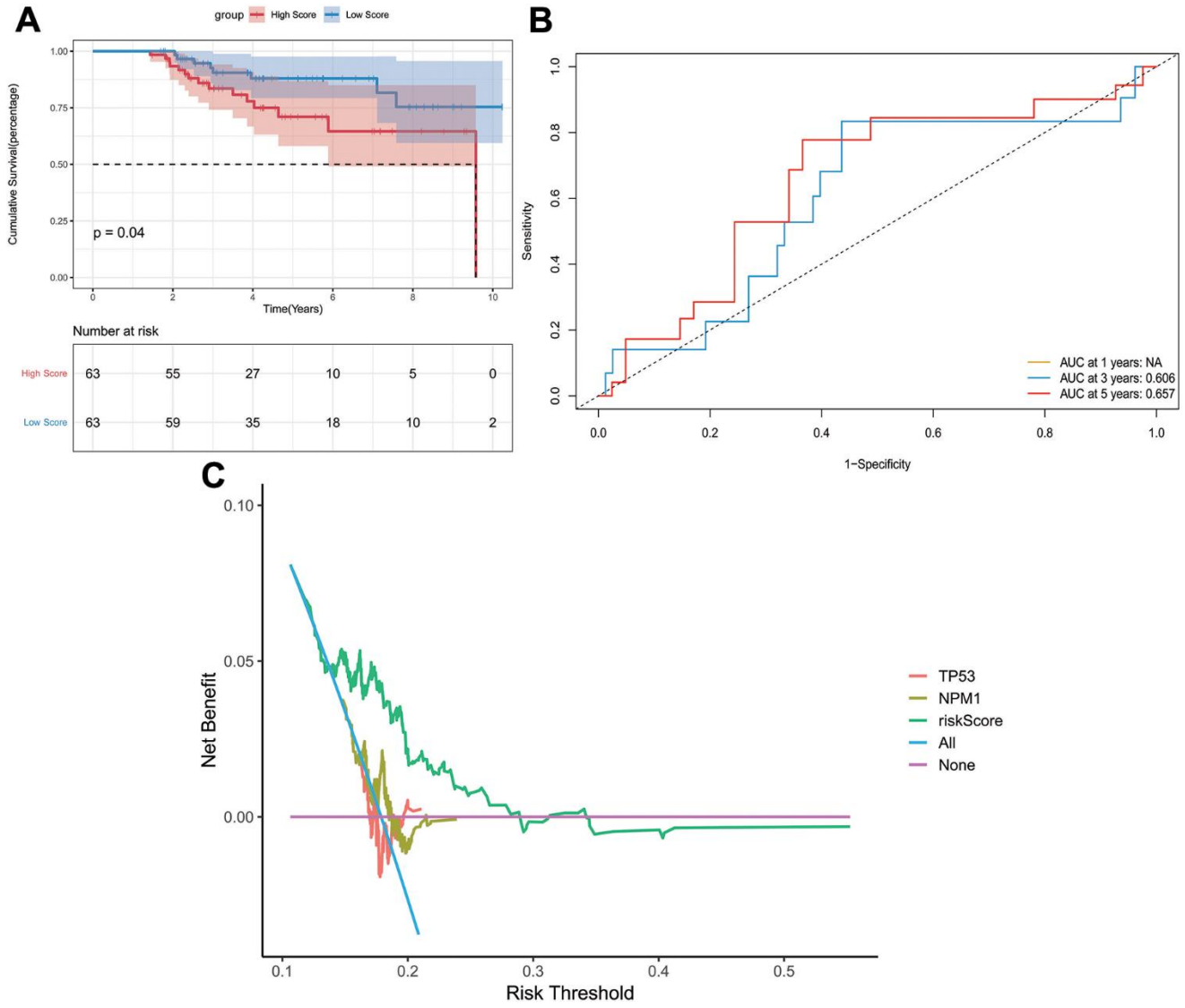
Supplementary Figures



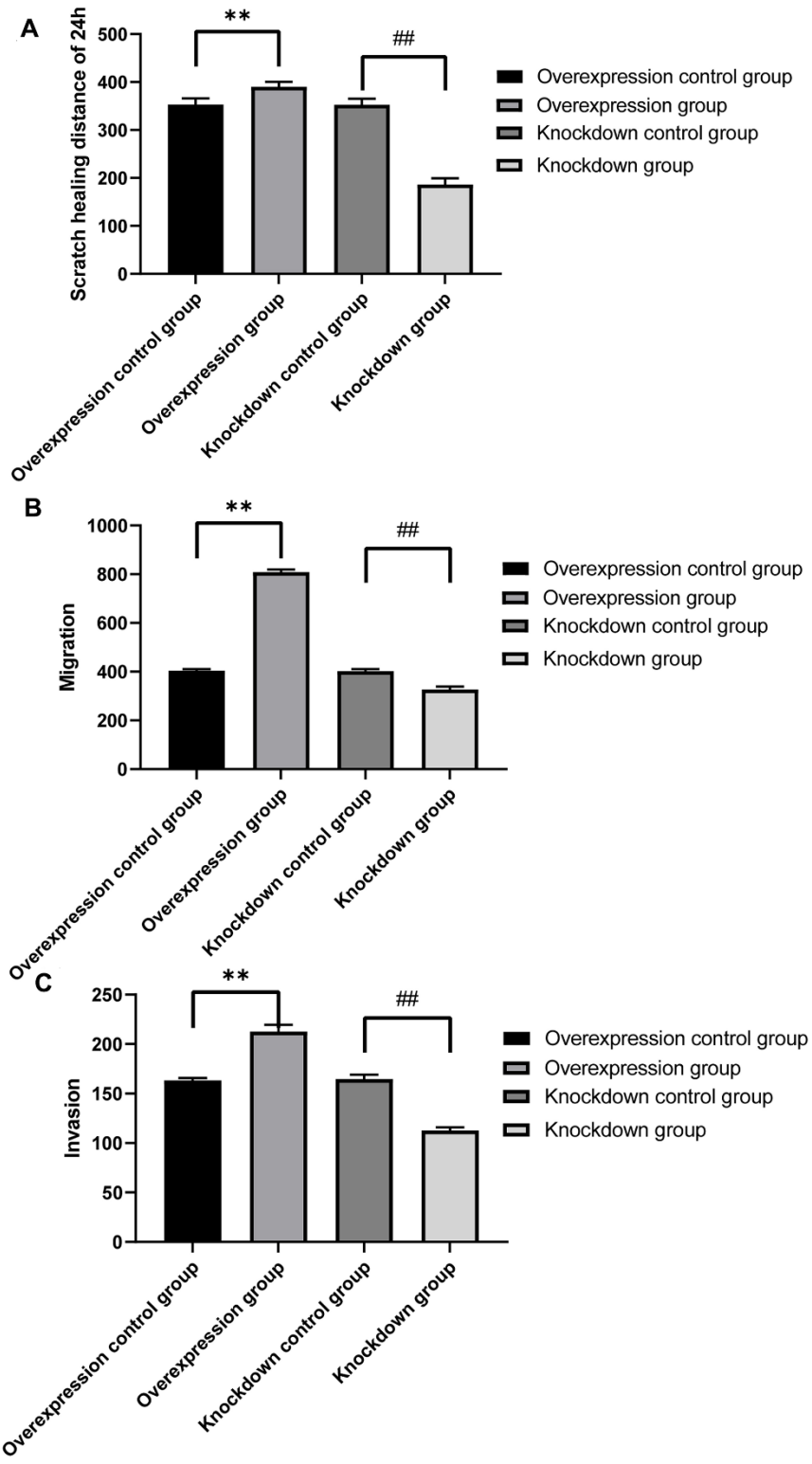
Supplementary Figure 1. Consensus clustering matrix analysis based on m^{6}A methylation regulators. (**A**) Color legend of consensus clustering matrix heat map in TCGA. The values of the consistency matrix from 0 (impossible to cluster together) to 1 (always cluster together) were represented by white to dark blue. (**B**) Consistent clustering matrix for k = 2 in TCGA. (**C**) Consistent clustering matrix for k = 3 in TCGA. (**D**) Consistent clustering matrix for k = 4 in TCGA. (**E**) Consistent clustering matrix for k = 5 in TCGA. (**F**) Color legend of consensus clustering matrix heat map in GEO. The values of the consistency matrix from 0 (impossible to cluster together) to 1 (always cluster together) were represented by white to dark blue. (**G**) Consistent clustering matrix for k = 2 in GEO. (**H**) Consistent clustering matrix for k = 3 in GEO. (**I**) Consistent clustering matrix for k = 4 in GEO. (**J**) Consistent clustering matrix for k = 5 in GEO. When taking different k values, the clustering effect was obviously different. The cleaner the clustering matrix, the better the effect.



Supplementary Figure 2. Prediction of HNRNPC, IGF2BP1 and IGF2BP3 related drugs based on DGIdb database.



Supplementary Figure 3. Kaplan-Meier survival and time-dependent ROC analyses based on cBioPortal, and DCA based on TCGA. (**A**) Kaplan-Meier survival analysis of early-stage LUAD in cBioPortal high and low risk score group; (**B**) Time-dependent ROC analysis measure the predictive value of the risk score in cBioPortal; (**C**) Prognostic indicator genes TP53 and NPM1 for LUAD and risk score were selected to perform DCA based on TCGA.



Supplementary Figure 4. One-way ANOVA was used to analyze the effect of HNRNPC on the migration and invasion of NCI-H1299 cells by SPSS17.0 software. (**A**) Statistical analysis of 24 h scratch healing distance of NCI-H1299 cells after HNRNPC overexpression and knockdown; (**B**) Statistical analysis of NCI-H1299 cells migration after HNRNPC overexpression and knockdown; (**C**) Statistical analysis of NCI-H1299 cells invasion after HNRNPC overexpression and knockdown. **P<0.01; ##P<0.01.

Supplementary Tables

Supplementary Table 1. Basic information of datasets.

Accession number /Source	Platform	Number of patients	Survival data
GEO: GSE31210	Affymetrix Human Genome U133 Plus 2.0 Array	Control: LUAD= 20:226	Overall Survival
TCGA: early-stage LUAD	Illumina RNAseq	Control: LUAD = 59:398	Overall Survival

GEO, Gene Expression Omnibus; TCGA, The Cancer Genome Atlas; LUAD, Lung adenocarcinoma.

Supplementary Table 2. Clinical information of patients in the RT-PCR.

Number	Sex	Age	Cancer metastasis	Pathological grade (stage I, II, III, IV)	Clinical stages (T stage, N stage, M stage)			Smoking history	Radiotherapy, chemotherapy and other treatment history	History of other malignancies	Family history of cancer
1	Female	50	No	IB	T2	N0	M0	No	No	No	No
2	Female	53	No	IA1	T1a	N0	M0	No	No	No	No
3	Male	43	No	IA1	T1a	N0	M0	Yes	No	No	No
4	Male	59	No	IA2	T1b	N0	M0	Yes	No	No	No
5	Male	57	No	IA1	T1a	N0	M0	Yes	No	No	No
6	Female	54	No	IA1	T1a	N0	M0	No	No	No	No
7	Female	47	No	IA1	T1a	N0	M0	No	No	No	No
8	Female	49	No	IB	T2a	N0	N0	No	No	No	No
9	Female	55	No	IA1	T1a	N0	N0	No	No	No	No
10	Male	66	No	IA2	T1b	N0	N0	No	No	No	No
11	Female	66	No	IA1	T1b	N0	N0	No	No	No	No
12	Female	52	No	IB	T2a	N0	N0	No	No	No	No
13	Male	52	No	IA2	T1b	N0	N0	Yes	No	No	No

Supplementary Table 3. Primer sequence in the RT-PCR.

Primer name	Primer sequence (5' to 3')
ACTB-F (internal reference)	5-CATGTACGTTGCTATCCAGGC-3
ACTB-R (internal reference)	5-CTCCTTAATGTCACGCACGAT-3
HNRNPC-F	5-GTTACCAACAAGACAGATCCTCG-3
HNRNPC-R	5-AGGCAAAGCCCTTATGAACAG-3
IGF2BP1-F	5-TGCAGTTTGGGTTGTGGACT-3
IGF2BP1-R	5-GATAGCTCCCAACTGCCTCC-3
IGF2BP3-F	5-CATCGAGGCGCTTTCAGGTA-3
IGF2BP3-R	5-CTCACAGCTCTCCACCACTC-3