Construction of a m6A RNA Methylation Regulator-related Signature for Prognosis and Immune Response in Lung Adenocarcinoma

Xiao Wu
Zhejiang university https://orcid.org/0000-0002-7330-9100
Hongxu Sheng
Zhejiang University School of Medicine First Affiliated Hospital
Pinghui Xia
Zhejiang University School of Medicine First Affiliated Hospital
Yiqing Wang
Zhejiang University School of Medicine First Affiliated Hospital
Li Yu
Zhejiang University School of Medicine First Affiliated Hospital
Wang Lv
Zhejiang University School of Medicine First Affiliated Hospital
Jian Hu (dr_hujian@zju.edu.cn)
Department of Thoracic Surgery, First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China. https://orcid.org/0000-0002-9494-9828

Primary research

Keywords: lung adenocarcinoma, m6A RNA methylation, prognostic signature, nomogram, immune response

Posted Date: August 21st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-60079/v1

Abstract

Background

N6-methyladenosine (m6A) RNA methylation is related to cancer pathogenesis and development. However, few studies have investigated the role of m6A regulator genes in lung adenocarcinoma (ADC).

Methods

Gene expressions with clinicopathological data were downloaded from The Cancer Genome Atlas (TCGA), and Gene Expression Omnibus (GEO) database. We compared the expression of twenty m6A regulator genes between tumor and normal samples. Univariate and least absolute shrinkage and selection operator (LASSO) Cox regression model were used to derive a multi-gene signature. A signature-based nomogram was developed, and the prediction performance was validated by an exterior validation set (GSE72094). Weighted gene co-expression network analysis (WGCNA) was performed to construct a co-expression network and identify the hub genes. The association of m6A signature with immunity was examined.

Results

Seventeen differentially expressed genes were identified. A five-gene prognostic signature (*IGF2BP1, IGF2BP2, HNRNPA2B1, METTL3*, and *HNRNPC*) was determined, and demonstrated as an unfavorable prognostic factor. A signature-based nomogram was developed to predict each patient's survival probability, and the nomogram was well calibrated and showed a satisfactory discrimination. Turquoise was identified as the most risk-related module, and genes in this module were enriched in the pathway of cell cycle. Six genes were determined as the hub genes. High-risk patients had significantly higher expression of *PD-L1*, higher tumor mutational burden (TMB), higher proportion of immune checkpoint blockade (ICB) response, and lower proportion of T cells CD8.

Conclusions

In summary, the signature-based nomogram is useful for survival prediction, and high-risk patients were more sensitive to the ICB therapy.

Background

Lung cancer is the most frequent type of malignant tumors, and results in the leading cancer-related cause of death in the United States (US) [1]. According to the US cancer statistic report, about 228,820 new cases were expected to be diagnosed with lung cancer in 2020, and approximately 135,720 Americans were projected to die of this cancer [2]. Lung adenocarcinoma (ADC), the most common histological type, accounts for approximately 49.7% of all lung cancers in the US [3], and exhibits an absolute growth in incidence currently [4, 5]. Despite the great improvement has been achieved in lung cancer screening and personalized treatment modalities (precision medicine) in current years, the 5-year

survival rates still remain dismal. As for stage I, the 5-year survival probability is 68–92%, while for stage IV, it decreased to 0–10% [6, 7]. Therefore, exploring a prognostic molecular signature and developing a prediction nomogram are essential for selecting the optimum therapeutic strategies and improving the adverse prognosis for lung ADC patients.

For eukaryotic species, N6-methyladenosine (m6A), a methylation modification at the N6 position of adenosine, is regarded as the most common, and conserved internal transcriptional modification detected in mRNAs, microRNAs (miRNAs), and long non-coding RNAs (lncRNAs) [8–11]. It is highly aggregated around stop codon and 3' untranslated terminal region (3'UTR), and influences RNA transcription, variable splicing, translation, and metabolism [12, 13]. Its methylation process is dynamically mediated by three categories of m6A regulator genes, which are termed as "readers" (binding proteins), "writers" (methyltransferase), and "erasers" (demethylase) [11, 14, 15]. Readers, including *IGF2BP1, IGF2BP2, IGF2BP3, YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, HNRNPG*, *HNRNPA2B1*, and *HNRNPC*, are responsible for identifying the m6A-modified RNAs, thereby producing the corresponding biological effects [16–19]. Writers, involved in a process of adding methylated modifications to RNAs, are composed of *ZC3H13, RBM15, METTL3, METTL14, METTL16, KIAA1429*, and *WTAP* [18–20]. Erasers, regulating the process of RNA demethylation, are consisted of *FTO* and *ALKBH5* [14, 18, 19, 21]. Under the impact of the above three regulatory elements, m6A methylation is actually a dynamic reversible process, thereby taking effect in mediating the expression of various genes [11].

Meanwhile, it is discovered that the changes of m6A regulator genes can exert important impacts on various physiological processes, such as self-renewal capability, circadian periods, embryonic development, stem cell differentiation, and cell death [22]. Moreover, increasing evidence has confirmed that aberrant m6A methylation modification is intimately related to cancer pathogenesis and development [11, 23]. For instance, it is reported that *METTL3* is generally up-regulated in clear cell renal cell carcinoma, and the deletion of *METTL3* is significantly correlated with a worse survival [24]. In addition, another study also demonstrates the overexpression of *METTL3* can dramatically suppress proliferation, migration and invasion for colorectal carcinoma (CRC) cells [25]. Thus, m6A regulator genes are not only promising to become the potential molecular markers for cancer diagnosis and prognosis, but also qualified to serve as the targets for the design of targeted anticancer drugs, which has been demonstrated in several literatures [26–28]. Several studies have investigated the role of m6A regulator genes in the survival of cancers, such as glioma [29], gastric cancer [30], and head and neck squamous cell carcinoma [31]. However, they did not include the newly discovered m6A regulator genes, seldom constructed a m6A-based nomogram, and never investigated the association with immunity

Therefore, in this study, we systematically evaluated the expression of twenty m6A regulator genes in 535 lung ADC and 59 normal samples, from which we screened out seventeen differentially expressed genes for further study. Meanwhile, we derived a multi-gene prognostic signature, and further assessed its performance in predicting the survival probability of lung ADC patients. Furthermore, we developed a prognostic nomogram to serve as a simple and reliable clinical tool to forecast each lung ADC patient's survival probability. To verify the model reliability, another independent cohort available from the GEO

was used as an exterior validation set. Weighted gene co-expression network analysis (WGCNA) was performed to identify the risk-related modules, physiological biology, pathways, and hub genes. In addition, we also evaluated the association of m6A prognostic signature with immune microenvironment, immune infiltration, immune checkpoint molecules, tumor mutational burden (TMB), and the response of immune checkpoint blockade (ICB).

Methods

Datasets

The whole study was designed and conducted based on the following flow diagram (Fig. 1). The RNA-seq transcriptome data, clinicopathological data, and simple nucleotide variation data were downloaded from the TCGA database (https://portal.gdc.cancer.gov). For RNA-seq data, the expression value of each gene was normalized by the Fragment Per Kilobase Million (FPKM) method. To conduct an external validation, GSE72094 [32] was used as a validation set. The gene expression profile with clinicopathological data was accessed from the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/). As only a small number of patients were classified as “Asian” or “American Indian/ Alaska Native”, we mainly concentrated on white and black patients.

Selection of m6A regulator genes

After comprehensively searching for m6A regulator genes from previous literatures [18, 19, 29, 31], a total of twenty recently known genes, including *FTO, ALKBH5, ZC3H13, RBM15, METTL3, METTL14, METTL16, KIAA1429, WTAP, YTHDF1, YTHDF2, YTHDF3, YTHDC1, YTHDC2, IGF2BP1, IGF2BP2, IGF2BP3, HNRNPC, HNRNPG*, and *HNRNPA2B1*, were incorporated into our study. Then, we systematically compared the expression of twenty m6A regulator genes between 535 lung ADC samples and 59 normal samples, from which differentially expressed genes were identified and included into further study. Accordingly, the raw p-value was corrected utilizing the false discovery rate (FDR) approach. Furthermore, heatmaps and boxplots were utilized to visualize the differential expression of these genes.

Bioinformatic analysis

Gene set enrichment analysis (GSEA) was carried out utilizing GSEA v3 software to functionally annotate a series of genes which were shown significant differences between different subgroups. Cancer-related pathways were extracted based on the following criteria: (1) |NES|>1; (2) normalized $P < 0.05$; (3) FDR q-val < 0.25 [33, 34]. Interactions among m6A regulator genes were calculated by virtue of the STRING database (https://string-db.org/).

Statistical analysis

Wilcoxon-Mann-Whitney test was utilized to compare the expression level of each m6A regulator gene between 535 lung ADC samples and 59 normal samples. In the meantime, we also utilized Wilcoxon-Mann-Whitney test to compare the expression level of each m6A regulator gene for age, gender, race, and

smoking history, and Kruskal-wallis test for American Joint Committee on Cancer (AJCC) stage. In addition, spearman correlation analysis was performed to investigate the correlations among different m6A regulator genes.

To verify our model performance of the training set, we employed an independent GEO cohort (GSE72094) as an exterior validation set. Furthermore, univariate Cox model was employed to assess the association of m6A regulator genes with overall survival (OS) in the training cohort, from which we screened and identified six genes significantly correlated with survival ($P < 0.05$). Then, we further included these genes into the least absolute shrinkage and selection operator (LASSO) Cox regression model to construct an optimum prognostic signature. Eventually, five m6A regulator genes with their coefficients were ascertained based on lambda.min (the lambda value corresponding to the minimum mean error) via 10-fold cross validation, and thereby a five-gene prognostic signature (*IGF2BP1, IGF2BP2, HNRNPA2B1, METTL3*, and *HNRNPC*) was derived. In order to figure out each patient's risk score in the training cohort and validation cohort, we performed a sum over each gene' score, which was acquired through multiplying the expression level of each gene by its coefficient. Patients were dichotomized into low-risk and high-risk groups according to the median value of risk scores. The prediction accuracy of the prognostic signature was assessed using Receiver operating characteristic (ROC) curves.

Chi-square tests were employed to compare the frequency distributions of grouped variables (including age, gender, race, smoking history, and AJCC stage) between the two risk groups. Additionally, survival curves between different risk groups were plotted by Kaplan-Meier method and compared by log-rank test. Subsequently, univariate Cox model was utilized to assess the association of variables (including risk score and clinicopathologic characteristics) with OS in the training set, from which we identified the variables significantly associated with survival ($P < 0.05$) and introduced them into the multivariate Cox model. For continuous variables (such as risk score), the restricted cubic splines (RCS) with three knots located at the empirical quantiles (10%, 50%, and 90%) were fitted to relax the linearity assumption of the model [35][35][35][35]. Then, forest plots were drawn to better visualize each prognostic variable's effect on OS, and a nomogram on the basis of the coefficients was constructed to forecast the 1-year, 3-year, and 5-year survival probability of each lung ADC patient. Accordingly, concordance index (C-index), area under the ROC curve (AUC), and calibration curves were used to evaluate the model performance in the training and validation set. A decision curve analysis (DCA) was conducted to assess the clinical usefulness of the nomogram [36]. All statistical analyses were performed utilizing the statistical software R v3.5.2. A two-tailed $P < 0.05$ was considered statistically significant.

WGCNA

Differentially expressed genes (DEGs) between high-risk and low-risk groups were screened based on the following criteria: (1) false discovery rate (FDR) < 0.05; (2) |logFoldChange| > 1. We incorporated the identified DEGs into the WGCNA to construct a gene co-expression network. Based on the power value β, a weighted adjacency matrix was constructed, and then the matrix was changed into a topological overlap matrix (TOM). Subsequently, gene modules were identified using the dynamic shear approach.

According to the module significance (MS), and the correlation coefficients between Module eigengenes (MEs) and risk scores, the most risk-related module was determined. All the genes in the modules were included into the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses using the DAVID database (https://david.ncifcrf.gov/). The determination of hub genes were based on the following criteria: (1) conforming to the cutoff criteria (Module membership (MM) > 0.85 and Gene significance (GS) > 0.55); (2) the top ten percent of genes based on the connectivity of the weighted network; (3) the top ten percent of genes based on the degree of the protein-protein interaction (PPI) analysis; (4) genes significantly associated with survival ($P < 0.05$).

Immunity

We utilized ESTIMATE algorithm to calculate each patient's immune score, stromal score, and tumor purity [37]. CIBERSORT was used to determine the proportions of the 22 immune cell subtypes for each sample [38], and P-value < 0.05 was set as the threshold value. As immune checkpoint molecules were important for immunotherapy, we explored their associations with m6A prognostic signature. A total of ten immune checkpoint molecules (*PD-L1, PD-1, PD-L2, CTLA-4, IDO1, LAG3, TIM-3, TIGIT, CD27*, and *ICOS*) were examined, and the expression differences between the high-risk and low-risk groups were assessed using Wilcoxon-Mann-Whitney test. TMB is defined as the total number of the somatic coding mutations per million bases. In this study, the mutational frequency of each sample was computed based on the number of variants divided by 38 (due to the length of exons being 38 million). For purpose of evaluating the potential responses of ICB, including *CTLA-4* and *PD-1*, for the high-risk and low-risk groups, Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was adopted, and the proportion of the response was calculated [39].

Results

The differential expression of twenty m6A regulator genes between normal and lung ADC samples

Allowing for the crucial role of each m6A regulator gene in cancer pathogenesis and development, we systematically investigated the expressions of twenty m6A regulator genes between normal and lung ADC samples, and the expression level of each gene was visualized by heatmap and boxplot. As shown in the heatmap, the expression levels of m6A regulator genes were significantly different between tumor and normal samples (Fig. 2a). To be exact, *HNRNPA2B1, HNRNPC, HNRNPG, IGF2BP1, IGF2BP2, IGF2BP3, KIAA1429, METTL3, RBM15, YTHDF1, YTHDF2*, and *YTHDF3* were significantly up-regulated in lung ADC tissues, while *FTO, METTL14, METTL16, WTAP*, and *ZC3H13* were significantly up-regulated in normal control samples (Fig. 2a-b). In addition, *YTHDC1*, *YTHDC2*, and *ALKBH5* were observed with no significant difference between tumor and normal samples. Consequently, these three genes were removed, and a total of seventeen genes were included into further study. We also explored the genetic changes (simple nucleotide variation) of these seventeen m6A regulator genes, and discovered that the

mutation frequency was generally low (no more than 3%) (Fig. 2c), which denoted that the expression changes of these genes were less likely to be caused by single nucleotide mutation.

The interaction and correlation among the seventeen m6A regulator genes

For better comprehending the relationships among these m6A regulator genes, we further investigated their interactions and correlations. As shown in the interaction analysis (Fig. 2d-e), *METTL3* and *METTL14* were regarded as the hub genes of the network, while *IGF2BP1*, and *IGF2BP3* had less associations with others. Moreover, *YTHDF3, IGF2BP2*, and *METTL16* had no relationship with the others. Notably, 'writer' genes had the highest while 'readers' genes had the least associations with others (Fig. 2d-e). Interestingly, all the 'writer' genes (except for *METTL16*) were interacted with each other (Fig. 2d), while *METTL16* had no association with the others. As displayed in the correlation analysis, seventeen m6A regulator genes tended to be weakly to moderately correlated with each other. In addition, we also observed that *KIAA1429* was positively correlated with *YTHDF3* (Fig. 2f), and their correlation coefficient was the highest (r = 0.66) than the others (Fig. 2f). Meanwhile, the highest negative correlation coefficient was found between *FTO* and *HNRNPC*, whereas they merely had a low correlation (r = -0.29) (Fig. 2f). Besides, *KIAA1429* was the only gene significantly correlated with the other 16 m6A regulator genes (Fig. 2f), suggesting a critical influence in the network.

Development of a prognostic signature and identification of the low-risk and high-risk groups

To assess the association of m6A regulator genes with OS, univariate Cox regression model was employed in the training set. Our findings revealed that six out of seventeen genes were identified as the prognostic genes ($P < 0.05$), that is, *IGF2BP1, IGF2BP2, IGF2BP3, HNRNPA2B1, METTL3*, and *HNRNPC* (Fig. 3a). Then, we incorporated these genes into the LASSO Cox regression model to develop an optimum prognostic signature. Ultimately, five genes were identified, and their coefficients were calculated based on the lambda.min via 10-fold cross validation (Fig. 3b-d). Thereby, a five-gene prognostic signature (*IGF2BP1, IGF2BP2, HNRNPA2B1, METTL3*, and *HNRNPC*) was determined. Each patient's risk score was computed via the following formula: Risk score = 0.037**IGF2BP1* + 0.004**IGF2BP2* + 0.007**HNRNPA2B1* + 0.022**HNRNPC* + (-0.125)**METTL3*, and patients were dichotomized into high-risk and low-risk groups on the basis of the median risk score (value = 0.060). Additionally, we also explored the distributions of risk scores and survival status, and our results revealed that low-risk subgroup had more alive cases and a higher OS than high-risk subgroup (Fig. 3e-f).

According to the results of the survival curve, we found that the high-risk group had a significantly poorer OS than low-risk group ($P < 0.0001$) (Fig. 4a), which indicated that the signature could successfully distinguish high-risk from low-risk lung ADC patients. Meanwhile, ROC curve demonstrated that the AUC of the five-gene prognostic signature were 0.684 and 0.646 for 3-year and 5-year survival rate of lung ADC patients, respectively, which was slightly less than the "acceptable discrimination" (Fig. 4b) [40].

Moreover, we visualized the association of the risk groups with clinicopathologic characteristics using heatmap. We discovered significant differences between the low-risk and high-risk groups for gender (P = 0.041), and AJCC stage (P = 0.028) (Fig. 4c). To investigate the functional enrichment of the differentially expressed genes between low-risk and high-risk subgroups, GSEA was conducted, and the results indicated that high-risk subgroup was strongly associated with cancer-related pathways, such as mismatch repair (NES = 2.045, normalized P = 0.004, FDR q-val = 0.005), cell cycle (NES = 2.382, normalized P < 0.0001, FDR q-val < 0.0001), nucleotide excision repair (NES = 2.017, normalized P = 0.006, FDR q-val = 0.006), p53 signaling pathway (NES = 2.069, normalized P < 0.0001, FDR q-val = 0.004), and pathways in cancer (NES = 1.585, normalized P = 0.021, FDR q-val = 0.109) (Fig. 4d).

Identification of the risk score as a prognostic indicator and development of a predictive nomogram

For purpose of ascertaining whether the five-gene signature-derived risk score could be served as a prognostic marker, both univariate and multivariate Cox model were employed in the training cohort (n = 398). The clinicopathologic characteristics of the training cohort were displayed in Table 1. In the univariate Cox model, AJCC stage (HR, 1.629; 95% confidence interval (CI), 1.400-1.894; P < 0.0001), and risk score (HR, 3.332; 95% CI, 2.242–4.952; P < 0.0001) were determined to be significantly related to the OS for lung ADC patients (Fig. 5a). Further included into the multivariate Cox model, AJCC stage (HR, 1.574; 95% CI, 1.353–1.831; P < 0.0001) and risk score (HR, 3.119; 95% CI, 2.059–4.724; P < 0.0001) were still significantly related to survival, and could be perceived as the independent unfavorable prognostic factors (Fig. 5b).

Table 1
Clinicopathological characteristics of patients in the training cohort and validation cohort.

Clinical Characteristics	TCGA (training Cohort) No. (%)	GEO (validation Cohort) No. (%)
Total	398	320
Age at diagnosis		
Median (IQR), years	65 (58–72)	71 (64–77)
Gender		
Female	215 (54.0)	170 (53.1)
Male	183 (46.0)	150 (46.9)
Race		
Black	45 (11.3)	9 (2.8)
White	353 (88.7)	311 (97.2)
AJCC Stage		
Stage I	219 (55.0)	212 (66.2)
Stage II	92 (23.1)	52 (16.2)
Stage III	64 (16.1)	46 (14.4)
Stage IV	23 (5.8)	10 (3.1)
Smoking History		
No	60 (15.1)	30 (9.4)
Yes	338 (84.9)	290 (90.6)
Survival Status		
Alive	257 (64.6)	233 (72.8)
Dead	141 (35.4)	87 (27.2)
Survival Time (months)	21.0 (13.9–35.2)	28.1 (19.3–34.4)
Abbreviations: IQR: interquartile range; AJCC: American Joint Committee on Cancer.		

Moreover, a nomogram based on the coefficients from multivariate Cox model was derived to visually forecast the 1-year, 3-year and 5-year survival probability for each lung ADC patient (Fig. 5c). According to the variable values (AJCC stage, or risk score) of each patient, drew the vertical lines from the value points of corresponding variable row to the “Points” row. Identifying the intersection points of vertical

lines with the “Points” row, each variable’ score was acquired. By performing a sum over each variable’ score, a total score was determined. Based on the derived total score, continued to draw a vertical line from the “Total Points” row to the “1-Year Survival”, “3-Year Survival” or “5-Year Survival” row, then we would obtain a corresponding predicted survival probability.

In order to determine the discrimination efficacy and prediction accuracy of the nomogram in the training set, related indicators, such as C-index, AUC, and calibration curves, were evaluated. Our results revealed that the derived nomogram was well calibrated, as the curves were close to the diagonal line (Fig. 5d). Besides, the C-index of the model was 0.71 (CI, 0.66–0.76) (Fig. 5d), and the AUC fluctuated around 0.75 in the follow-up period, which indicated a satisfactory discrimination power (Fig. 5e). We also discovered that the nomogram had a higher discrimination power than the conventional prognostic index—AJCC stage (Fig. 5e). Subsequently, the DCA showed that the clinical net benefit gained from the nomogram was higher than that in the either hypothetical treat-all-patients or treat-none scenarios, when the threshold probabilities were within the range of 0.06–0.48, 0.18–0.81, and 0.42–0.80 for 1-year, 3-year, and 5-year OS, respectively (Fig. 5f). Therefore, the developed nomogram had a favorable clinical net benefit for predicting 1-year, 3-year OS, and 5-year OS.

Model validation in GEO dataset

To further verify the model reliability, a total of 320 lung ADC samples available from the GSE72094 were used as a validation cohort. The clinicopathologic characteristics of the validation cohort were displayed in Table 1. Based on the formula mentioned above, each patient’s risk score was computed, and patients were dichotomized into high-risk and low-risk groups based on the determined cutoff value (value = 0.060). As shown in Fig. 6a-b, high-risk subgroup had more dead cases, and a lower OS than low-risk subgroup. Similar findings were also observed in the survival curve (P = 0.001) (Fig. 6c**)**. When further performing univariate and multivariate Cox analyses, we found risk score was still perceived as an independent unfavorable prognostic factor for lung ADC patients, both P < 0.0001 (Fig. 6d-e). Meanwhile, ROC curve indicated that the AUC of the prognostic signature were 0.695 and 0.656 for 3-year and 5-year OS, respectively, which was - less than the “acceptable discrimination” (Fig. 6f). Furthermore, the 1-year, 3-year, and 5-year nomogram calibration curves were still close to the diagonal line, and the C-index was 0.72 (CI, 0.67–0.77), both denoting a satisfactory model performance (Fig. 6g). In addition, we also found the AUC of the model fluctuated around 0.75 in the follow-up period (from 1-year to 5-year follow-up), and was superior to the AJCC stage (Fig. 5e).

WGCNA

After differential expression analysis, a total of 741 DEGs (335 up-regulated and 406 down-regulated) were identified (Fig. 7a), and we further included them into the WGCNA to develop a co-expression network. For purpose of constructing a scale-free network, the soft threshold power was set as 8, which achieved a higher scale-free topology fit index (> 0.8) accompanied with a higher mean connectivity (Fig. 7b-c). By means of the average linkage hierarchical clustering, a total of three modules (blue, brown, and turquoise) were identified (Fig. 7d). To investigate the associations of clinical characteristics with

modules, correlation coefficients between MEs and clinical characteristics were calculated. Turquoise module was identified with the highest correlation with risk score (Fig. 7e-f).

To functionally annotate the genes in the turquoise module, both GO and KEGG enrichment analyses were conducted, and we found the genes were primarily enriched in biological processes, including cell division, cell proliferation, and cell cycle, as well as in signaling pathways, including cell cycle, and p53 signaling pathway (Fig. 8a-b).

A total of 22 genes highly correlated with risk score (GS > 0.55) and turquoise module (MM > 0.85) were determined as the potential candidates of hub genes (Fig. 8c). Based on the connectivity of the weighted network, top fifteen genes (top 10%) were identified as the hub gene candidates. Furthermore, we performed a PPI analysis for all genes in the turquoise module using STRING database and Cytoscape software, top 15 genes (top 10%) with the highest degree were regarded as the hub gene candidates (Fig. 8d). By intersecting the above three lists of candidate genes, a total of six common genes were identified and included into the further survival analysis (Fig. 8e). Moreover, to investigate the association of six genes with the survival of lung ADC patients, univariate Cox model was adopted, and our findings revealed that the genes were significantly related to survival (Fig. 8f). Eventually, these six genes (*CCNA2, CCNB1, BUB1B, BUB1, KIF2C*, and *KIF11*) were all determined as the real hub genes.

The association of m6A prognostic signature with immunity

ESTIMATE algorithm [37] was utilized to evaluate tumor immune microenvironment, including immune score, stromal score, and tumor purity for each lung ADC sample. In general, there were no significant differences between the low-risk and high-risk groups for immune score, stromal score, and tumor purity (all P > 0.05, Fig. 9a), and the correlation coefficients between the signature-based risk scores and immune microenvironment scores were all below 0.2 (Fig. 9b) which was regarded as negligible correlation [41, 42]. In addition, we also investigated the association of m6A prognostic signature with immune infiltration. As for the 21 immune cell types, high-risk patients had significantly lower proportions of Macrophages M1, T cells CD4 memory activated, T cells CD8, and Neutrophils (Fig. 9c). As immune checkpoint molecules were important for immunotherapy, we explored their associations with m6A prognostic signature. As shown in Fig. 9d, high-risk patients had significantly higher expression levels of *PD-L1* and *PD-L2*. As TMB is regarded as an important biomarker of response for *PD-1/PD-L1* blockade, we examined the association of the signature with TMB. In the present study, there was a low correlation between risk score and TMB (r = 0.28) (Fig. 9e), and high-risk patients had significantly higher TMB (Fig. 9f), suggesting a superior response to ICBs among high-risk patients. Finally, in order to assess the potential responses of ICB, including *CTLA-4* and *PD-1*, for different risk groups, TIDE algorithm was adopted. We found high-risk patients had a significantly higher proportion of ICB response than low-risk patients (Fig. 9g). In summary, all the above findings indicated that high-risk patients might be more sensitive to the ICB therapy.

Discussion

It is generally believed that patients diagnosed with late stage lung ADC tend to have a poorer clinical prognosis [2, 6, 43, 44]. Therefore, seeking for a robust prognostic signature to distinguish high-risk from low-risk patients and a nomogram to predict each patient's survival probability, is essential for determining the optimum therapeutic strategies and reversing the adverse prognosis for lung ADC patients. Accumulating evidence proved that abnormal m6A RNA methylation modifications were intimately related to the pathogenesis and development of various cancers [11, 23, 29, 31, 45–47], whereas the studies investigating their role in lung ADC are few. In this study, we discovered a significantly abnormal expression of m6A regulator genes in lung ADC. Besides, a robust five-gene prognostic signature was developed and regarded as an unfavorable prognostic factor for lung ADC. Based on the signature, we constructed a prognostic nomogram to predict each patient's survival probability. To further assess the model performance, an exterior validation set (GSE72094) was used, and the nomogram was demonstrated with a satisfactory discrimination and calibration performance.

Our results revealed that seventeen out of twenty m6A regulator genes were observed with either up-regulation or down-regulation in lung ADC samples, indicating these genes might be involved in the oncogenic activities and prognosis for lung ADC patients. Unanimous with our findings, the abnormal expression of m6A regulator genes was also found in other cancers. For instance, it is reported that the writer *WTAP* was found to be up-regulated in various cancers, including hepatocellular carcinoma (HCC) [48], acute myelogenous leukemia (AML) [49], and glioblastoma [50], and was identified as an oncogene. The eraser *ALKBH5* was reported to be highly expressed in ovarian cancer [51], and the silencing of *ALKBH5* significantly improved autophagy and suppressed proliferation and invasion. The reader *YTHDF1* was observed with an up-regulation in CRC [52], and its knockdown significantly inhibited the tumorigenicity and colonosphere formation ability. Therefore, expression dysregulation of m6A regulator genes was prevalently found in various cancers, and was strongly related to tumorigenesis, progression, and prognosis.

When GSEA was adopted, we found m6A regulator genes were significantly associated with some important malignancy-related pathways, including p53 signaling pathway, cell cycle, mismatch repair, nucleotide excision repair, and pathways in cancer. Comparable results were also reported in other studies. It is reported that m6A regulator genes were extremely important in several cancer-related pathways, such as p53 signaling pathway [29, 53], cell cycle [29, 30, 54], Ras [30], inflammatory response [29, 53], and PPAR signaling pathway [53].

In the present study, a five-gene prognostic signature (*IGF2BP1, IGF2BP2, HNRNPA2B1, METTL3*, and *HNRNPC*) was identified and displayed with a great prediction value for patients with lung ADC. Furthermore, a distinct survival difference was found between different risk groups, which denoted that clinicians could successfully stratify the lung ADC patients into low-risk and high-risk groups based on the risk scores derived from the prognostic signature, thereby conducing them to make wiser clinical decisions. For the five identified m6A regulator genes, *IGF2BP1, IGF2BP2, HNRNPA2B1*, and *HNRNPC* were determined as adverse prognostic genes, while *METTL3* was determined as a favorable prognostic gene. Interestingly, all these genes were highly expressed in lung ADC patients. Yan M et al. demonstrated

that *HNRNPC* was drastically up-regulated in non-small cell lung cancer (NSCLC) tissues, which was in agreement with our research. Then, the researchers also discovered the overexpression of *HNRNPC* dramatically accelerated the proliferation, migration, and invasion of lung cancer cells. Moreover, the elevated expression of *HNRNPC* was significantly related to advanced tumor stages, metastasis, and shorter survival time [55]. *IGF2BP1*, a target of miR-491-5p, was reported to be significantly increased in the expression of NSCLC cells, and promoted tumor cell proliferation, migration, and invasion [56]. The same trend was also discovered in liver cancer [57]. Zhu J et al. revealed that *METTL3* was significantly up-regulated in lung adenocarcinoma, and patients with high expression of *METTL3* were observed with significantly better OS [24]. Similarly, another study reported that *METTL3* was regarded as a tumor suppressor gene for CRC, and the up-regulation of *METTL3* could dramatically suppress tumor cell proliferation, migration and invasion. Furthermore, patients with elevated expression of *METTL3* were observed with a significantly better survival [25, 45]. Overall, their findings were in agreement with ours. However, other studies demonstrated an entirely opposite function of *METTL3* for bladder cancer [58], NSCLC [59], and liver cancer [24]. Du M et al. proved that *METTL3* was targeted by miR-33a and the down-regulated *METTL3* could inhibit the proliferation of lung cancer cells [59]. In addition, another study found *METTL3* was significantly up-regulated in bladder cancer and its knockdown could significantly suppress bladder cancer cell proliferation, invasion, and survival in vitro and tumorigenicity in vivo [58]. These findings demonstrated that *METTL3* plays different (or even opposite) roles in different cancers [60]. *IGF2BP2*, a direct target of miR-485-5p, was discovered to be significantly up-regulated in the expression of NSCLC, and its depletion significantly suppressed tumor cell proliferation and invasion [61]. Similarly, *IGF2BP2* was overexpressed in pancreatic cancer and patients with the overexpression of *IGF2BP2* were observed with significantly worse OS. Furthermore, the up-regulated *IGF2BP2* expression promoted tumor cell growth by activating the PI3K/Akt signaling pathway [62]. *HNRNPA2B1*, involved in RNA-binding and pre-RNA processing, was high expression in NSCLC patients and associated with a worse prognosis. Furthermore, the overexpression of *HNRNPA2B1* accelerated NSCLC cell growth by activating COX-2 signaling pathway [63].

Based on the multivariate Cox model, we developed a prognostic nomogram to predict each lung ADC patient's 1-year, 3-year and 5-year OS probability. As far as we know, this was the first study incorporating the m6A regulator genes into the model to construct a prognostic nomogram for lung ADC patients. Notably, the nomogram performed well both in the training and validation set, indicating a robust prediction performance. To evaluate the clinical usefulness, we utilized DCA to ascertain whether the nomogram-based decisions could improve patients' outcomes. The DCA showed that the threshold probabilities were within the range of 0.06–0.48, 0.18–0.81, and 0.42–0.80 for 1-year, 3-year, and 5-year OS, respectively. If the threshold probabilities of lung ADC patients were within the above ranges, adopting the nomogram to predict OS added more benefits than either hypothetical treat-all-patients or treat-none scenarios. In addition, we also found the nomogram had a higher predictive accuracy than the traditional prognostic index—AJCC stage. In summary, our five-gene signature-based nomogram was effective in predicting each lung ADC patient's survival probability and could offer a better reference for treatment guidance than single traditional clinical index.

To explore the potential reasons for risk signature being an independent prognostic factor for lung ADC patients, we conducted a WGCNA to identify the related modules, physiological biology, pathways, and hub genes. Our results revealed that cell cycle, an important cancer-related pathway, was significantly associated with the risk signature. Furthermore, a total of six genes (*CCNA2, CCNB1, BUB1B, BUB1, KIF2C*, and *KIF11*) were determined as the real hub genes. Notably, these hub genes all played an important role in the proceeding of cell cycle [64–70]. As our prognostic signature was consisted of five m6A regulator genes (*IGF2BP1, IGF2BP2, HNRNPA2B1, METTL3*, and *HNRNPC*), we inferred that these five m6A-related genes might act with the six hub genes, thereby influencing the process of cell cycle. Further experimental studies were warranted to ascertain the real molecular mechanisms.

ICB therapy has revolutionized traditional treatment strategies for NSCLC and other cancers, and advanced patients treated with anti-*PD-1* and anti-*CTLA-4* agents have been demonstrated with better prognosis [71–73]. In addition, previous literatures have reported that m6A is extremely important in immunoregulation and autoimmune diseases [74, 75]. Therefore, it is essential to investigate the association of the developed m6A signature with immunity. To our knowledge, this is the first study assessing the relationship between m6A signature with immunity. It is reported that higher *PD-L1* expression in tumor cells was closely associated with improved efficacy of immunotherapy [76, 77], and the stimulation of *PD-1*/ *PD-L1* pathway inhibits CD8 + T cell proliferation and promotes its apoptosis [78, 79]. In this study, we found high-risk lung ADC patients had significantly higher expressions of *PD-L1* and *PD-L2*, and lower proportion of CD8 + T, which suggested high-risk lung ADC patients had a superior response to ICBs. Additionally, we also found high-risk patients had significantly higher TMB. Several reports have demonstrated that TMB serves as a useful biomarker of response for *PD-1/PD-L1* blockade across various cancers, and higher TMB is associated with superior progression-free survival and objective response rates [77, 80–82]. Finally, after using TIDE algorithm, we also found high-risk patients had a significantly higher proportion of ICB response than low-risk patients. All the above findings indicated that high-risk patients might be more sensitive to the ICB therapy and more likely to benefit from immunotherapy. With the aid of the signature, clinicians can be easier to identify the potential beneficiaries from immunotherapy.

Several limitations in the present study should be noted. First, all our data were derived from the existing public database. Although an excellent performance in the prediction of the survival for lung ADC patients was observed in our study, a prospective, multicenter study was still warranted to further validate our results. Second, as the study populations were mainly from the US, and we only focused on white and black patients, our findings might not be generalized to other countries and races. Third, the C-index of our nomogram was merely around 0.7 and the AUC of the signature was within 0.6–0.7. Therefore, incorporating some acknowledged prognostic factors, such as tumor grade, radiation, chemotherapy, operation modes, and immunotherapy, might be conducive to enhancing the prediction accuracy of the present model. Finally, all our results were based on the data mining, and consequently an experimental study is essential for better ascertaining the associated molecular mechanisms.

Conclusions

In conclusion, our studies systematically evaluated the expression and prognosis of m6A regulator genes in lung ADC. A five-gene prognostic signature was identified as an unfavorable prognosis of lung ADC, and the signature could successfully distinguish high-risk from low-risk lung ADC patients. Furthermore, a signature-based nomogram was developed and was effective in predicting each lung ADC patient's survival probability. In addition, we performed a WGCNA and inferred that the genes in the signature might act with the six hub genes, thereby influencing the process of cell cycle. We also found high-risk lung ADC patients had a significantly higher proportion of ICB response and were more sensitive to the ICB therapy. Future clinical and experimental research is warranted to further verify our findings.

Abbreviations

ADC
adenocarcinoma; AJCC:American Joint Committee on Cancer; AUC:area under the ROC curve; C-index:concordance index; DCA:decision curve analysis; DEGs:differentially expressed genes; FDR:false discovery rate; GEO:Gene Expression Omnibus; GO:Gene Ontology; GS:Gene significance; GSEA:Gene set enrichment analysis; ICB:immune checkpoint blockade; KEGG:Kyoto Encyclopedia of Genes and Genomes; LASSO:least absolute shrinkage and selection operator; lncRNAs:long non-coding RNAs; m6A:N6-methyladenosine; MEs:Module eigengenes; miRNAs:microRNAs; MM:Module membership; MS:module significance; OS:overall survival; PPI:protein-protein interaction; RCS:restricted cubic splines; ROC:receiver operating characteristic; TCGA:The Cancer Genome Atlas; TIDE:Tumor Immune Dysfunction and Exclusion; TMB:tumor mutational burden; US:United States; WGCNA:Weighted gene co-expression network analysis.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The research data will be shared upon request.

Competing interests

The authors declare no conflict of interest.

Funding

This work was supported by The National Key Research and Development Program of China (No. 2017YFC0113500), Zhejiang Province Major Science and Technology Special Project (2014C03032), Zhejiang Provincial Key Discipline of Traditional Chinese Medicine (2017-XK-A33), and Zhejiang Lung Cancer Center (JBZX-202007).

Author Contributions

Wu X, Sheng HX, and Hu J designed and wrote the manuscript; Wu X, Xia PH, Wang YQ, Yu L, Lv W, and Hu J participated in literature search, data acquisition, data analysis, or data interpretation; Wu X, Sheng HX, and Hu J contributed to the revision of manuscript. All authors approved the final version to be published.

Acknowledgements

The authors are sincerely grateful to the availability of data from the TCGA and GEO.

References

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017;67(1):7–30.
2. American Cancer Society. Cancer facts & Figs. 2020 [http://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/cancer-facts-and-figures-2020.pdf (accessed 21 February 2020)].
3. National Cancer Institute and Surveillance Epidemiology and End Results (SEER). Program. Cancer of the Lung and Bronchus. https://seer.cancer.gov/csr/1975_2016/sections.html (accessed 21 February 2020).
4. Lewis DR, Check DP, Caporaso NE, Travis WD, Devesa SS. US lung cancer trends by histologic type. Cancer. 2014;120(18):2883–92.
5. Meza R, Meernik C, Jeon J, Cote ML. Lung cancer incidence trends by gender, race and histology in the United States, 1973–2010. PLoS One. 2015;10(3):e0121323.
6. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, Nicholson AG, Groome P, Mitchell A, Bolejack V. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. J Thorac Oncol. 2016;11(1):39–51.
7. Wang BY, Huang JY, Chen HC, Lin CH, Lin SH, Hung WH, Cheng YF. The comparison between adenocarcinoma and squamous cell carcinoma in lung cancer patients. J Cancer Res Clin Oncol. 2020;146(1):43–52.
8. Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci U S A. 1974;71(10):3971–75.

9. Alarcón CR, Lee H, Goodarzi H, Halberg N, Tavazoie SF. N6-methyladenosine marks primary microRNAs for processing. Nature. 2015;519(7544):482–85.
10. Patil DP, Chen C-K, Pickering BF, Chow A, Jackson C, Guttman M, Jaffrey SR. m(6)A RNA methylation promotes XIST-mediated transcriptional repression. Nature. 2016;537(7620):369–73.
11. Chen X-Y, Zhang J, Zhu J-S. The role of mA RNA methylation in human cancer. Mol Cancer. 2019;18(1):103.
12. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell. 2012;149(7):1635–46.
13. Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, Haripal B, Zucker-Scharff I, Moore MJ, Park CY, et al. A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. Genes Dev. 2015;29(19):2037–53.
14. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, Yi C, Lindahl T, Pan T, Yang Y-G, et al. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol. 2011;7(12):885–87.
15. Yang Y, Hsu PJ, Chen Y-S, Yang Y-G. Dynamic transcriptomic mA decoration: writers, erasers, readers and functions in RNA metabolism. Cell Res. 2018;28(6):616–24.
16. Haussmann IU, Bodi Z, Sanchez-Moran E, Mongan NP, Archer N, Fray RG, Soller M. mA potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. Nature. 2016;540(7632):301–04.
17. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature. 2015;518(7540):560–64.
18. Liu L, Wang Y, Wu J, Liu J, Qin Z, Fan H. N(6)-Methyladenosine: A Potential Breakthrough for Human Cancer. Mol Ther Nucleic Acids. 2019;19:804–13.
19. Sun T, Wu R, Ming L. The role of m6A RNA methylation in cancer. Biomed Pharmacother. 2019;112:108613.
20. Schöller E, Weichmann F, Treiber T, Ringle S, Treiber N, Flatley A, Feederle R, Bruckmann A, Meister G. Interactions, localization, and phosphorylation of the mA generating METTL3-METTL14-WTAP complex. RNA (New York). 2018;24(4):pp. 499–512.
21. Tang C, Klukovich R, Peng H, Wang Z, Yu T, Zhang Y, Zheng H, Klungland A, Yan W. ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. Proc Natl Acad Sci U S A. 2018;115(2):E325-E33.
22. Liu N, Pan T. N6-methyladenosine-encoded epitranscriptomics. Nat Struct Mol Biol. 2016;23(2):98–102.
23. Wang S, Sun C, Li J, Zhang E, Ma Z, Xu W, Li H, Qiu M, Xu Y, Xia W, et al. Roles of RNA methylation by means of N-methyladenosine (mA) in human cancers. Cancer Lett. 2017;408:112–20.
24. Chen M, Wei L, Law C-T, Tsang FH-C, Shen J, Cheng CL-H, Tsang L-H, Ho DW-H, Chiu DK-C, Lee JM-F, et al. RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through

YTHDF2-dependent posttranscriptional silencing of SOCS2. Hepatology. 2018;67(6):2254–70.

25. Deng R, Cheng Y, Ye S, Zhang J, Huang R, Li P, Liu H, Deng Q, Wu X, Lan P, et al. mA methyltransferase METTL3 suppresses colorectal cancer proliferation and migration through p38/ERK pathways. Onco Targets Ther. 2019;12:4391–402.
26. He L, Li H, Wu A, Peng Y, Shu G, Yin G. Functions of N6-methyladenosine and its role in cancer. Mol Cancer. 2019;18(1):176.
27. Huang Y, Su R, Sheng Y, Dong L, Dong Z, Xu H, Ni T, Zhang ZS, Zhang T, Li C, et al. Small-Molecule Targeting of Oncogenic FTO Demethylase in Acute Myeloid Leukemia. Cancer Cell. 2019;35(4):677–91. e10.
28. Huang Y, Yan J, Li Q, Li J, Gong S, Zhou H, Gan J, Jiang H, Jia GF, Luo C, et al. Meclofenamic acid selectively inhibits FTO demethylation of m6A over ALKBH5. Nucleic Acids Res. 2015;43(1):373–84.
29. Chai RC, Wu F, Wang QX, Zhang S, Zhang KN, Liu YQ, Zhao Z, Jiang T, Wang YZ, Kang CS. m(6)A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. Aging. 2019;11(4):1204–25.
30. Su Y, Huang J, Hu J. m(6)A RNA Methylation Regulators Contribute to Malignant Progression and Have Clinical Prognostic Impact in Gastric Cancer. Front Oncol. 2019;9:1038.
31. Zhao X, Cui L. Development and validation of a mA RNA methylation regulators-based signature for predicting the prognosis of head and neck squamous cell carcinoma. Am J Cancer Res. 2019;9(10):2156–69.
32. Schabath MB, Welsh EA, Fulp WJ, Chen L, Teer JK, Thompson ZJ, Engel BE, Xie M, Berglund AE, Creelan BC, et al. Differential association of STK11 and TP53 with KRAS mutation-associated gene expression, proliferation and immune surveillance in lung adenocarcinoma. Oncogene. 2016;35(24):3209–16.
33. Zeng L, Fan X, Wang X, Deng H, Zhang K, Zhang X, He S, Li N, Han Q, Liu Z. Bioinformatics Analysis based on Multiple Databases Identifies Hub Genes Associated with Hepatocellular Carcinoma. Curr Genomics. 2019;20(5):349–61.
34. Xu Z, Wang N, Xu Y, Hua L, Zhou D, Zheng M, Deng X. Effects of chronic PM exposure on pulmonary epithelia: Transcriptome analysis of mRNA-exosomal miRNA interactions. Toxicol Lett. 2019;316:49–59.
35. Harrell F. Jr Harrell. Regression modeling strategies 2016.
36. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inf Decis Making. 2008;8(1):53.
37. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.
38. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453–57.

39. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24(10):1550–58.
40. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. Journal of thoracic oncology: official publication of the International Association for the Study of Lung Cancer. 2010;5(9):1315–16.
41. Crane AK, Geller EJ, Myers EM, Fenderson JL, Wells E, Jannelli M, Connolly A, Matthews CA. Implementation of a standardized digital rectal exam to improve the accuracy of rectocele diagnosis. Int Urogynecol J. 2015;26(1):107–11.
42. Prasad K, Dash D, Kumar A. Validation of the Hindi version of National Institute of Health Stroke Scale. Neurol India. 2012;60(1):40–4.
43. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ, Wu Y-L, Paz-Ares L. Lung cancer: current therapies and new targeted treatments. The Lancet. 2017;389(10066):299–311.
44. Shao Q, Liu S, Wang W, Zhang Y, Li F, Li J. Clinical investigation into the initial diagnosis and treatment of 539 patients with stage IV lung cancer. Onco Targets Ther. 2017;10:535–41.
45. Zhu J, Wang M, Hu D. Deciphering N(6)-Methyladenosine-Related Genes Signature to Predict Survival in Lung Adenocarcinoma. Biomed Res Int. 2020;2020:2514230.
46. Ding Y, Qi N, Wang K, Huang Y, Liao J, Wang H, Tan A, Liu L, Zhang Z, Li J, et al. FTO Facilitates Lung Adenocarcinoma Cell Progression by Activating Cell Migration Through mRNA Demethylation. Onco Targets Ther. 2020;13:1461–70.
47. Zhang Y, Liu X, Liu L, Li J, Hu Q, Sun R. Expression and Prognostic Significance of m6A-Related Genes in Lung Adenocarcinoma. Medical science monitor: international medical journal of experimental clinical research. 2020;26:e919644.
48. Chen Y, Peng C, Chen J, Chen D, Yang B, He B, Hu W, Zhang Y, Liu H, Dai L, et al. WTAP facilitates progression of hepatocellular carcinoma via m6A-HuR-dependent epigenetic silencing of ETS1. Mol Cancer. 2019;18(1):127.
49. Bansal H, Yihua Q, Iyer SP, Ganapathy S, Proia DA, Proia D, Penalva LO, Uren PJ, Suresh U, Carew JS, et al. WTAP is a novel oncogenic protein in acute myeloid leukemia. Leukemia. 2014;28(5):1171–74.
50. Jin D-I, Lee SW, Han M-E, Kim H-J, Seo S-A, Hur G-Y, Jung S, Kim B-S, Oh S-O. Expression and roles of Wilms' tumor 1-associating protein in glioblastoma. Cancer Sci. 2012;103(12):2102–09.
51. Zhu H, Gan X, Jiang X, Diao S, Wu H, Hu J. ALKBH5 inhibited autophagy of epithelial ovarian cancer through miR-7 and BCL-2. J Exp Clin Cancer Res. 2019;38(1):163.
52. Bai Y, Yang C, Wu R, Huang L, Song S, Li W, Yan P, Lin C, Li D, Zhang Y. YTHDF1 Regulates Tumorigenicity and Cancer Stem Cell-Like Activity in Human Colorectal Carcinoma. Front Oncol. 2019;9:332.
53. Wang J, Zhang C, He W, Gou X. Effect of m(6)A RNA Methylation Regulators on Malignant Progression and Prognosis in Renal Clear Cell Carcinoma. Front Oncol. 2020;10:3.

54. Zhou Y, Yin Z, Hou B, Yu M, Chen R, Jin H, Jian Z. Expression profiles and prognostic significance of RNA N6-methyladenosine-related genes in patients with hepatocellular carcinoma: evidence from independent datasets. Cancer Manag Res. 2019;11:3921–31.
55. Yan M, Sun L, Li J, Yu H, Lin H, Yu T, Zhao F, Zhu M, Liu L, Geng Q, et al. RNA-binding protein KHSRP promotes tumor growth and metastasis in non-small cell lung cancer. J Exp Clin Cancer Res. 2019;38(1):478.
56. 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(2):485–95.
57. Gutschner T, Hämmerle M, Pazaitis N, Bley N, Fiskin E, Uckelmann H, Heim A, Groβ M, Hofmann N, Geffers R, et al. Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) is an important protumorigenic factor in hepatocellular carcinoma. Hepatology. 2014;59(5):1900–11.
58. Cheng M, Sheng L, Gao Q, Xiong Q, Zhang H, Wu M, Liang Y, Zhu F, Zhang Y, Zhang X, et al. The mA methyltransferase METTL3 promotes bladder cancer progression via AFF4/NF-κB/MYC signaling network. Oncogene. 2019;38(19):3667–80.
59. Du M, Zhang Y, Mao Y, Mou J, Zhao J, Xue Q, Wang D, Huang J, Gao S, Gao Y. MiR-33a suppresses proliferation of NSCLC cells via targeting METTL3 mRNA. Biochem Biophys Res Commun. 2017;482(4):582–89.
60. Zheng W, Dong X, Zhao Y, Wang S, Jiang H, Zhang M, Zheng X, Gu M. Multiple Functions and Mechanisms Underlying the Role of METTL3 in Human Cancers. Front Oncol. 2019;9:1403.
61. Huang RS, Zheng YL, Li C, Ding C, Xu C, Zhao J. MicroRNA-485-5p suppresses growth and metastasis in non-small cell lung cancer cells by targeting IGF2BP2. Life Sci. 2018;199:104–11.
62. Xu X, Yu Y, Zong K, Lv P, Gu Y. Up-regulation of IGF2BP2 by multiple mechanisms in pancreatic cancer promotes cancer proliferation by activating the PI3K/Akt signaling pathway. J Exp Clin Cancer Res. 2019;38(1):497.
63. Xuan Y, Wang J, Ban L, Lu J-J, Yi C, Li Z, Yu W, Li M, Xu T, Yang W, et al. hnRNPA2/B1 activates cyclooxygenase-2 and promotes tumor growth in human lung cancers. Mol Oncol. 2016;10(4):610–24.
64. Zwicker J, Lucibello FC, Wolfraim LA, Gross C, Truss M, Engeland K, Müller R. Cell cycle regulation of the cyclin A, cdc25C and cdc2 genes is based on a common mechanism of transcriptional repression. EMBO J. 1995;14(18):4514–22.
65. Roberts JM. Evolving ideas about cyclins. Cell. 1999;98(2):129–32.
66. Bharadwaj R, Yu H. The spindle checkpoint, aneuploidy, and cancer. Oncogene. 2004;23(11):2016–27.
67. Logarinho E, Bousbaa H. Kinetochore-microtubule interactions "in check" by Bub1, Bub3 and BubR1: The dual task of attaching and signalling. Cell cycle (Georgetown Tex). 2008;7(12):1763–68.
68. Sacristan C, Kops GJPL. Joined at the hip: kinetochores, microtubules, and spindle assembly checkpoint signaling. Trends Cell Biol. 2015;25(1):21–8.

69. Ganguly A, Bhattacharya R, Cabral F. Cell cycle dependent degradation of MCAK: evidence against a role in anaphase chromosome movement. Cell cycle (Georgetown Tex). 2008;7(20):3187–93.
70. Rath O, Kozielski F. Kinesins and cancer. Nature reviews Cancer. 2012;12(8):527–39.
71. Hellmann MD, Paz-Ares L, Bernabe Caro R, Zurawski B, Kim SW, Carcereny Costa E, Park K, Alexandru A, Lupinacci L, de la Mora Jimenez E, et al. Nivolumab plus Ipilimumab in Advanced Non-Small-Cell Lung Cancer. N Engl J Med. 2019;381(21):2020–31.
72. Escudier B, Motzer RJ, Sharma P, Wagstaff J, Plimack ER, Hammers HJ, Donskov F, Gurney H, Sosman JA, Zalewski PG, et al. Treatment Beyond Progression in Patients with Advanced Renal Cell Carcinoma Treated with Nivolumab in CheckMate 025. Eur Urol. 2017;72(3):368–76.
73. Li J, Kan H, Zhao L, Sun Z, Bai C. Immune checkpoint inhibitors in advanced or metastatic mucosal melanoma: a systematic review. Ther Adv Med Oncol. 2020;12:1758835920922028.
74. Paramasivam A, Priyadharsini JV, Raghunandhakumar S. Implications of m6A modification in autoimmune disorders. Cell Mol Immunol. 2020;17(5):550–51.
75. Zhang C, Fu J, Zhou Y. A Review in Research Progress Concerning m6A Methylation and Immunoregulation. Front Immunol. 2019;10:922.
76. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, Patnaik A, Aggarwal C, Gubens M, Horn L, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. N Engl J Med. 2015;372(21):2018–28.
77. Hellmann MD, Ciuleanu T-E, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, Minenza E, Linardou H, Burgers S, Salman P, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. N Engl J Med. 2018;378(22):2093–104.
78. Boussiotis VA. Molecular and Biochemical Aspects of the PD-1 Checkpoint Pathway. N Engl J Med. 2016;375(18):1767–78.
79. Leng C, Li Y, Qin J, Ma J, Liu X, Cui Y, Sun H, Wang Z, Hua X, Yu Y. Relationship between expression of PD-L1 and PD-L2 on esophageal squamous cell carcinoma and the antitumor effects of CD8 + T cells. Oncol Rep. 2016;35(2):699–708.
80. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, Stephens PJ, Daniels GA, Kurzrock R. Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. Mol Cancer Ther. 2017;16(11):2598–608.
81. Wang Z, Duan J, Cai S, Han M, Dong H, Zhao J, Zhu B, Wang S, Zhuo M, Sun J, et al. Assessment of Blood Tumor Mutational Burden as a Potential Biomarker for Immunotherapy in Patients With Non-Small Cell Lung Cancer With Use of a Next-Generation Sequencing Cancer Gene Panel. JAMA Oncol. 2019;5(5):696–702.
82. Leslie M. Mutation Burden Predicts Anti-PD-1 Response. Cancer Discov. 2018;8(3):258.

Figures

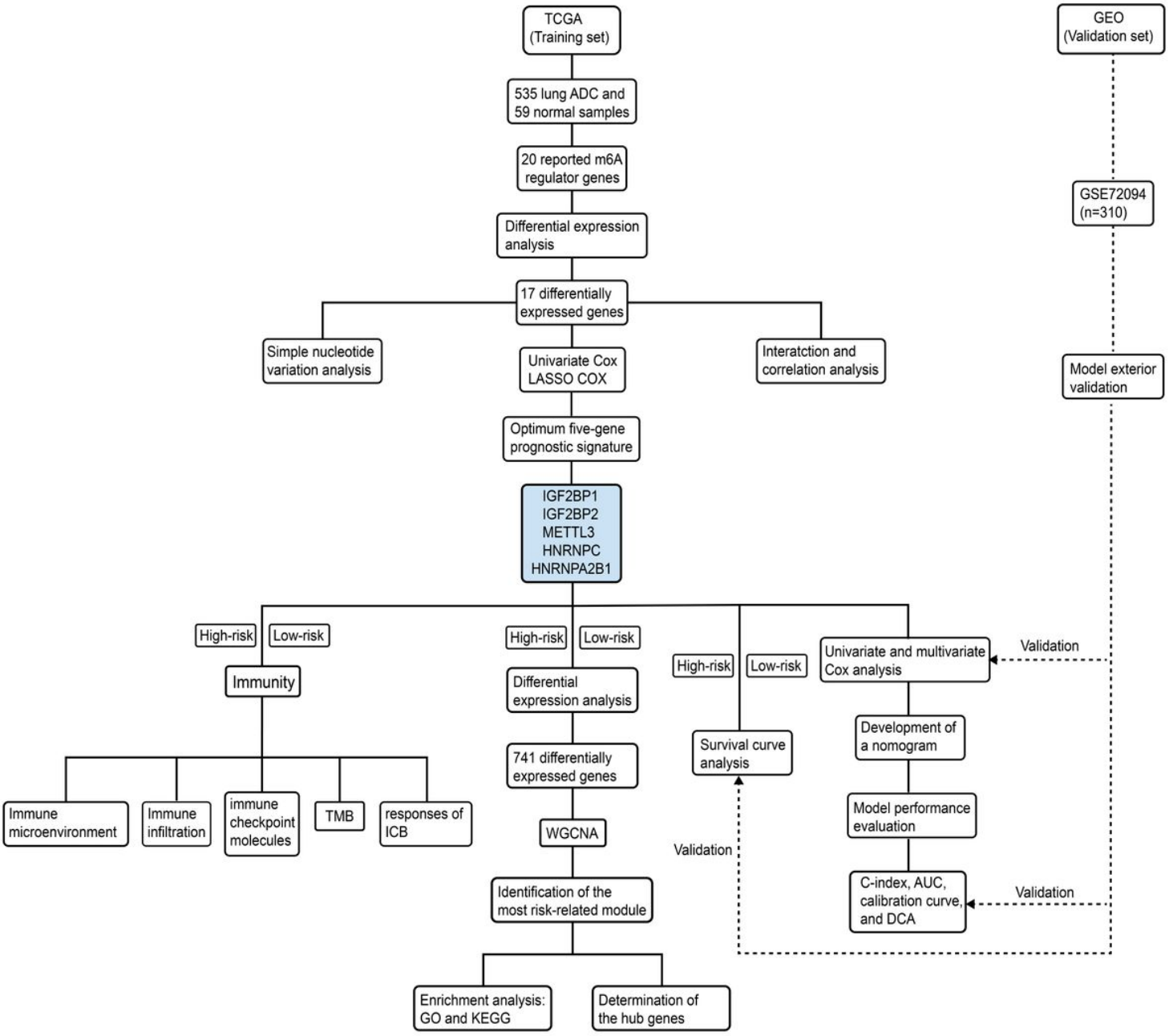

Figure 1

Flow diagram of the whole study The whole study mainly included five sections: bioinformatics analysis of m6A regulator genes, identification of a five-gene signature and nomogram, and the exterior validation using GEO dataset (GSE72094), WGCNA, and the association of the signature with immunity.

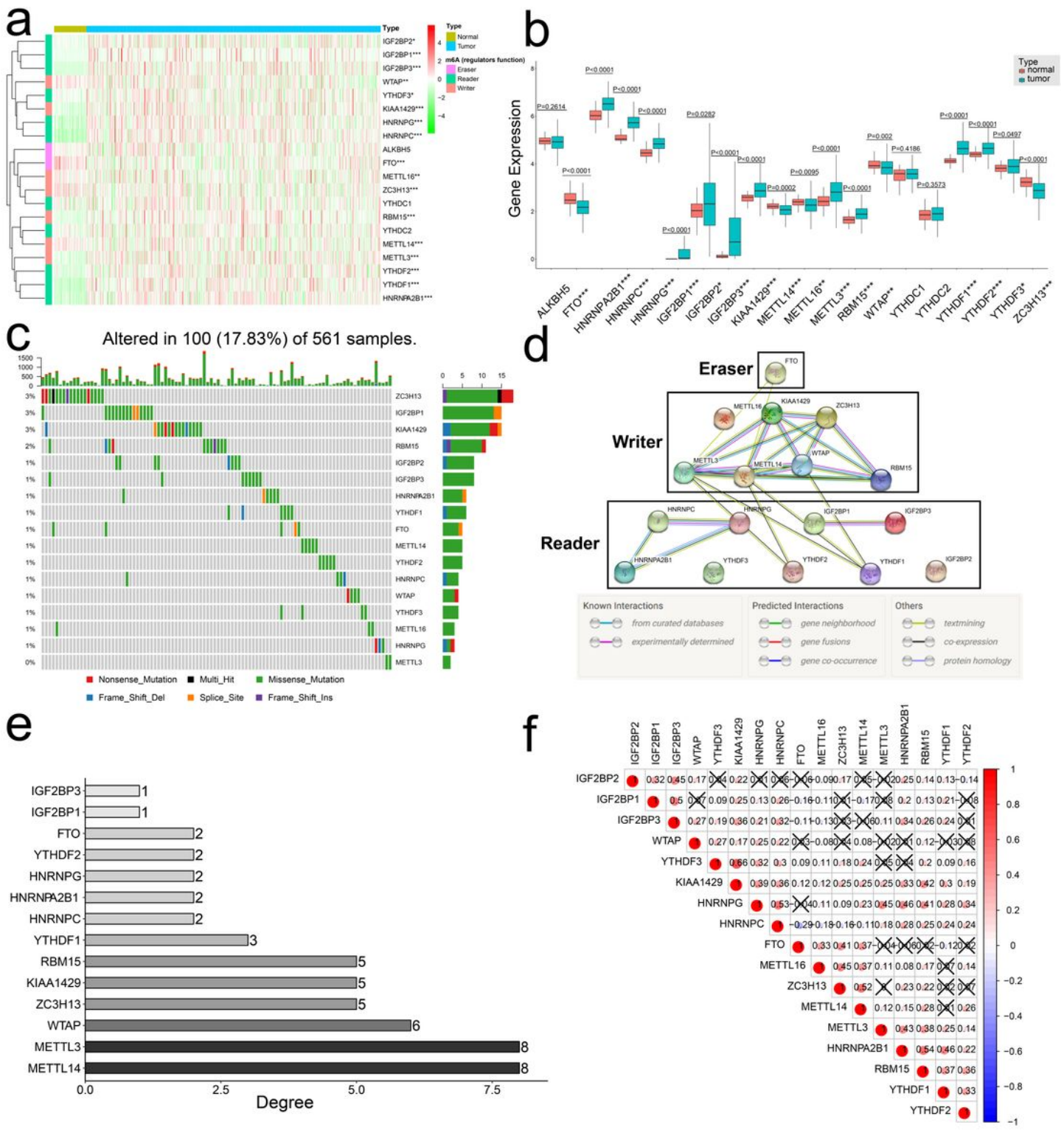

Figure 2

Expression of twenty m6A regulator genes and interaction among them (a) The heatmap of twenty m6A regulator genes in 59 normal and 535 lung ADC tissues, “*” referring to 0.01 ≤ P < 0.05, “**” referring to 0.001 ≤ P < 0.01, “***” referring to P < 0.001. (b) Boxplot visualizing the differential expression of twenty m6A regulator genes between normal and tumor tissues, “*” referring to 0.01 ≤ P < 0.05, “**” referring to 0.001 ≤ P < 0.01, “***” referring to P < 0.001. (c)Simple nucleotide variation of seventeen m6A regulator

genes in 561 tumor tissues. (d) PPI network of seventeen m6A regulator genes derived from the STRING website. (e) The quantity of interacted nodes of seventeen m6A regulator genes. (f)Spearman correlation analysis of seventeen m6A regulator genes.

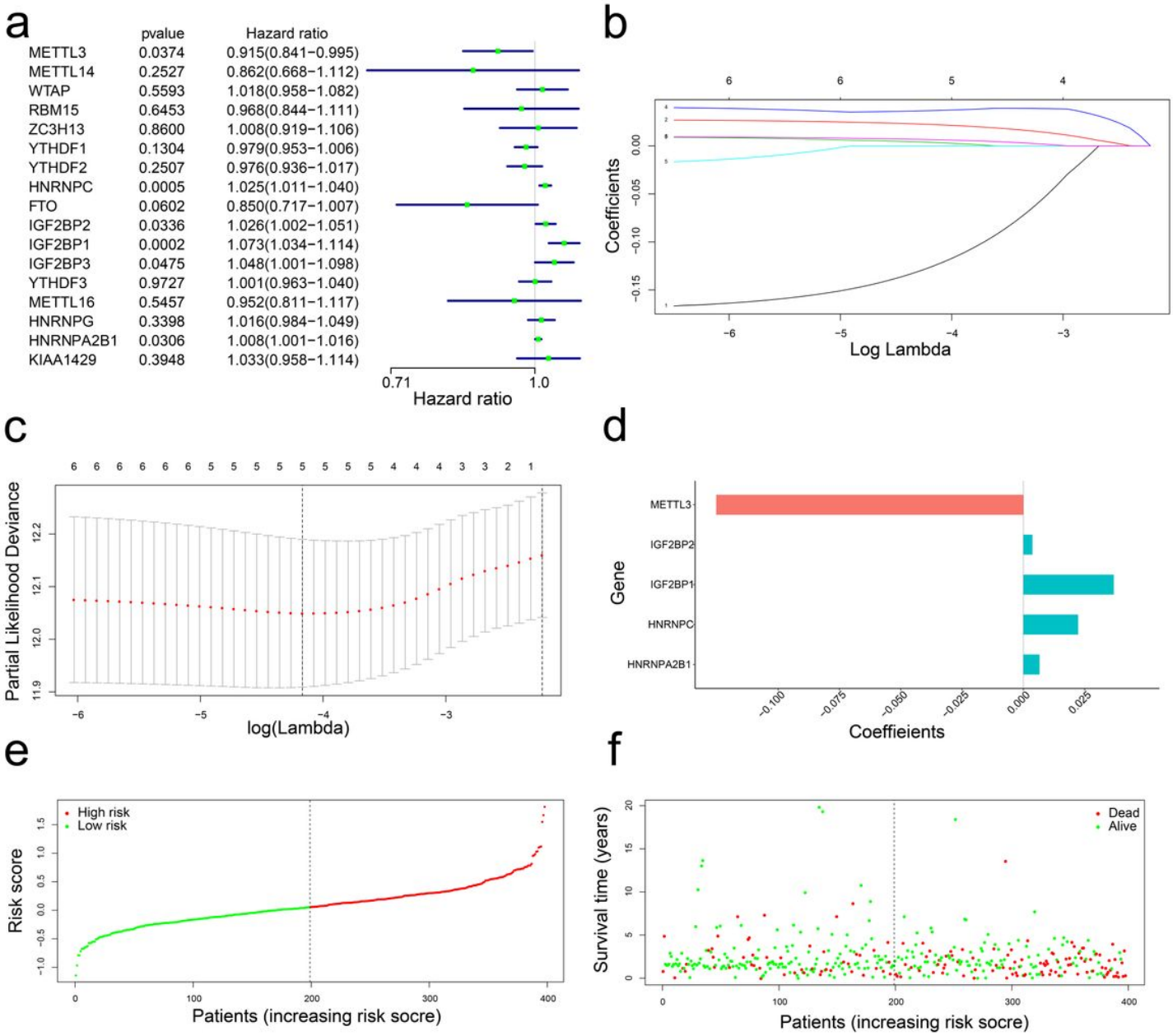

Figure 3

Development of a five-gene prognostic signature (a) The forest plot visualizing the association of seventeen m6A regulator genes with OS status determined by univariate Cox regression model. (b)LASSO coefficients profiles for six m6A regulatorgenes. (c)Five genes with coefficients obtained based on the lambda.min via 10-fold cross validation. (d) Coefficient distributions of the five identified genes. (e) The distribution of each patient's risk score in the training set (TCGA). (f) The distributions of risk score and OS status in the training set (TCGA).

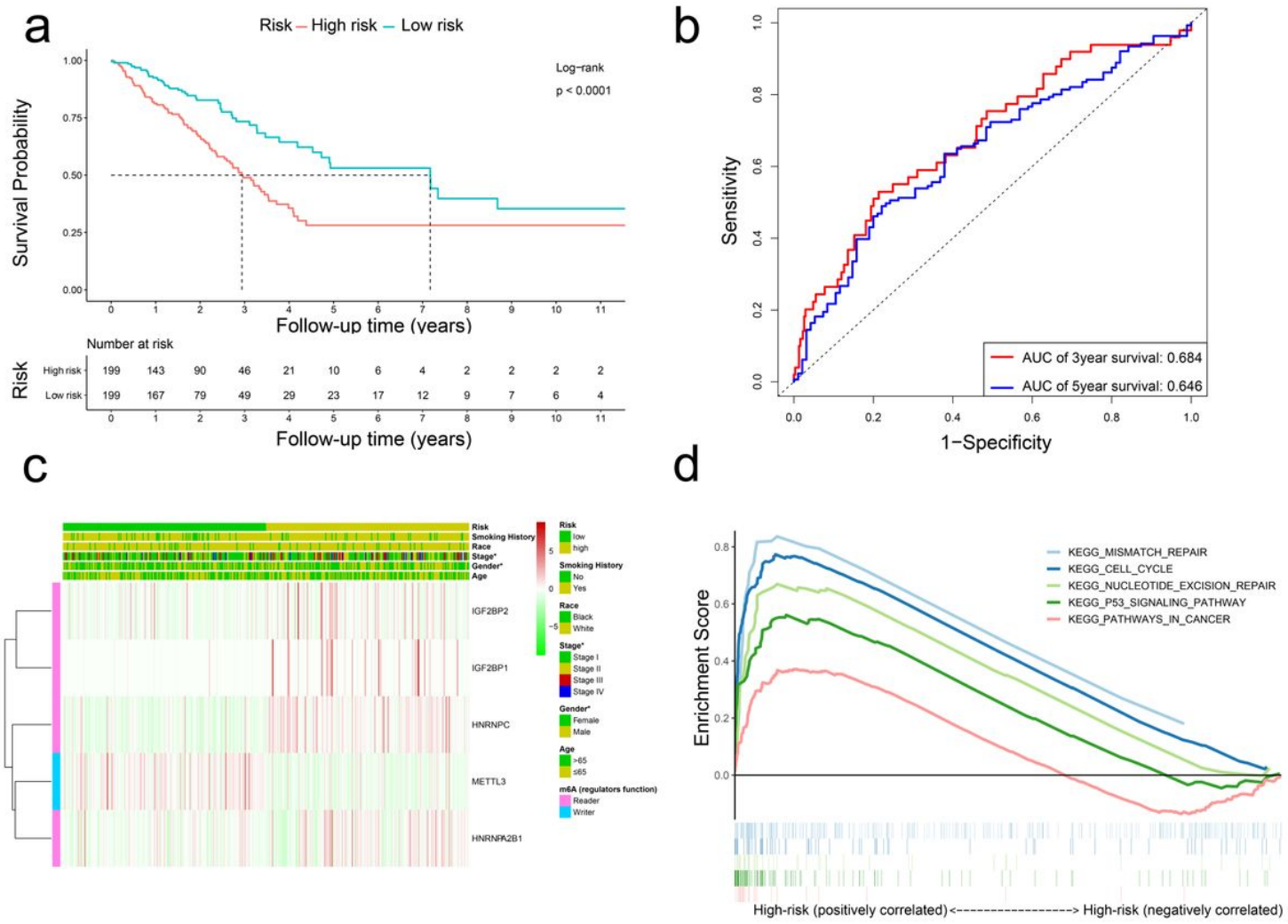

Figure 4

The association of risk score with OS and clinicopathologic characteristics (a) A Kaplan-Meier curve between the high-risk and low-risk subgroups in the training set (TCGA)(b) A ROC curve for the prediction accuracy of prognostic signature in the training set (TCGA). (c) The Heatmap visualizing the expression of five genes and the association of clinicopathologic characteristicswith risk subgroups, “*” referring to $0.01 \leq P < 0.05$, “**” referring to $0.001 \leq P < 0.01$, “***” referring to $P < 0.001$. (d) Cancer-related pathways positively correlated with high-risk subgroup.

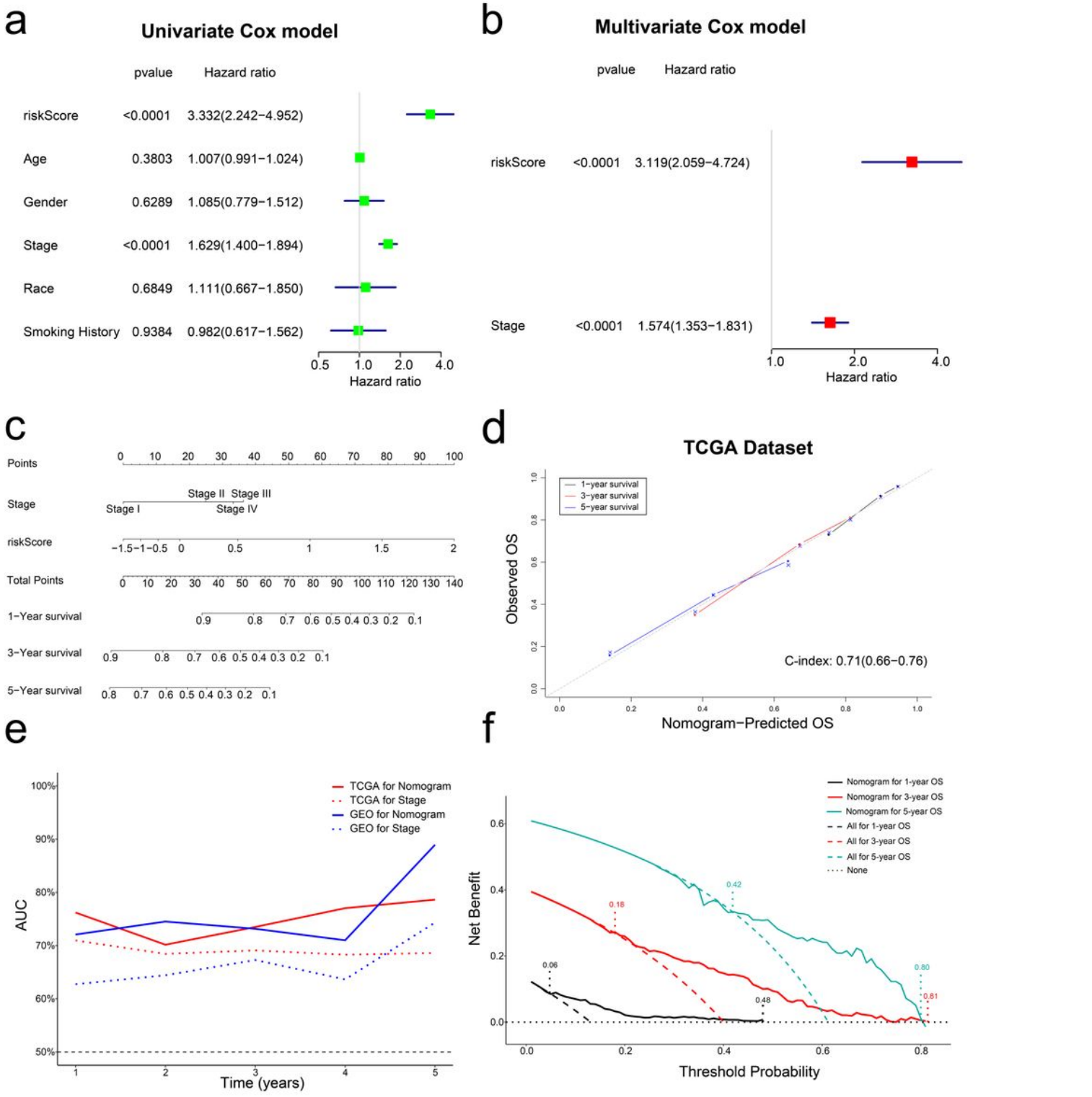

Figure 5

Construction of a signature-based nomogram and the model performance evaluation in TCGA dataset (a-b) The forest plot visualizing the association of risk score and clinicopathologic characteristics with OS status assessed by univariate and multivariate Cox regression model in the TCGA dataset. (c) Asignature-based nomogram predicting each lung ADC patient's 1-year, 3-year and 5-year OS probability. (d) The 1-year, 3-year, and 5-year nomogram calibration curveswith C-index in the training set (TCGA). (e) AUC

variations of the training set and validation set throughout the follow-up period (from 1-year to 5-year follow-up) for nomogram and AJCC stage. (f) The DCA of the nomogram for 1-year, 3-year and 5-year OS.

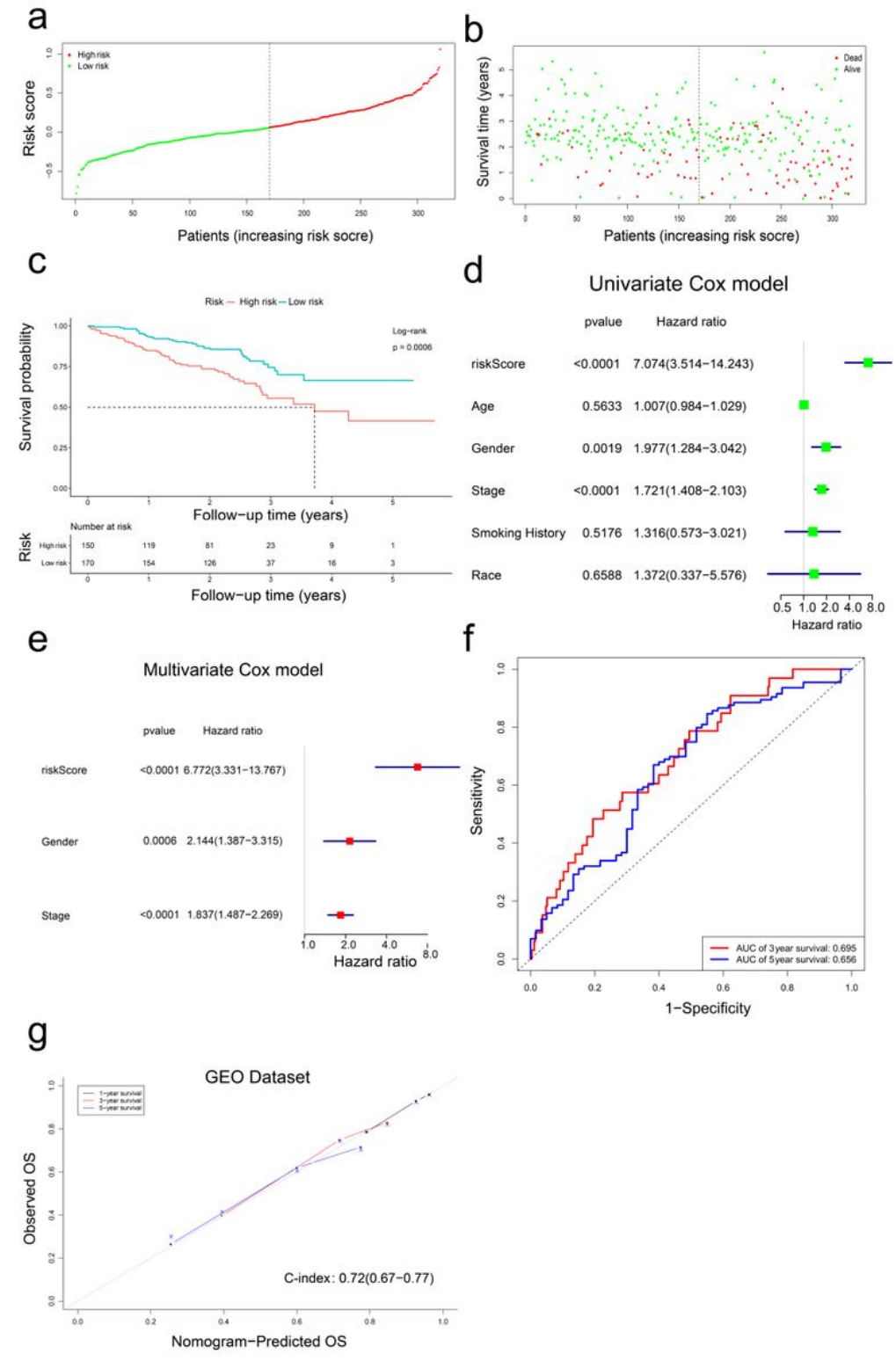

Figure 6

The model validation in GEO dataset (a)The distribution of each patient's risk score in the validation set (GEO). (b) The distributions of risk score and OS status in the validation set (GEO). (c) A Kaplan-Meier curve between the high-risk and low-risk subgroups inthe validation set (GEO). (d-e) The forest plot

visualizing the association of risk score and clinicopathologic characteristics with OS status assessed by univariate and multivariate Cox regression model in the validation set (GEO). (f) A ROC curve for the prediction accuracy of prognostic signature in the validation set (GEO). (g)The 1-year, 3-year, and 5-year nomogram calibration curveswith C-index in the validation set (GEO).

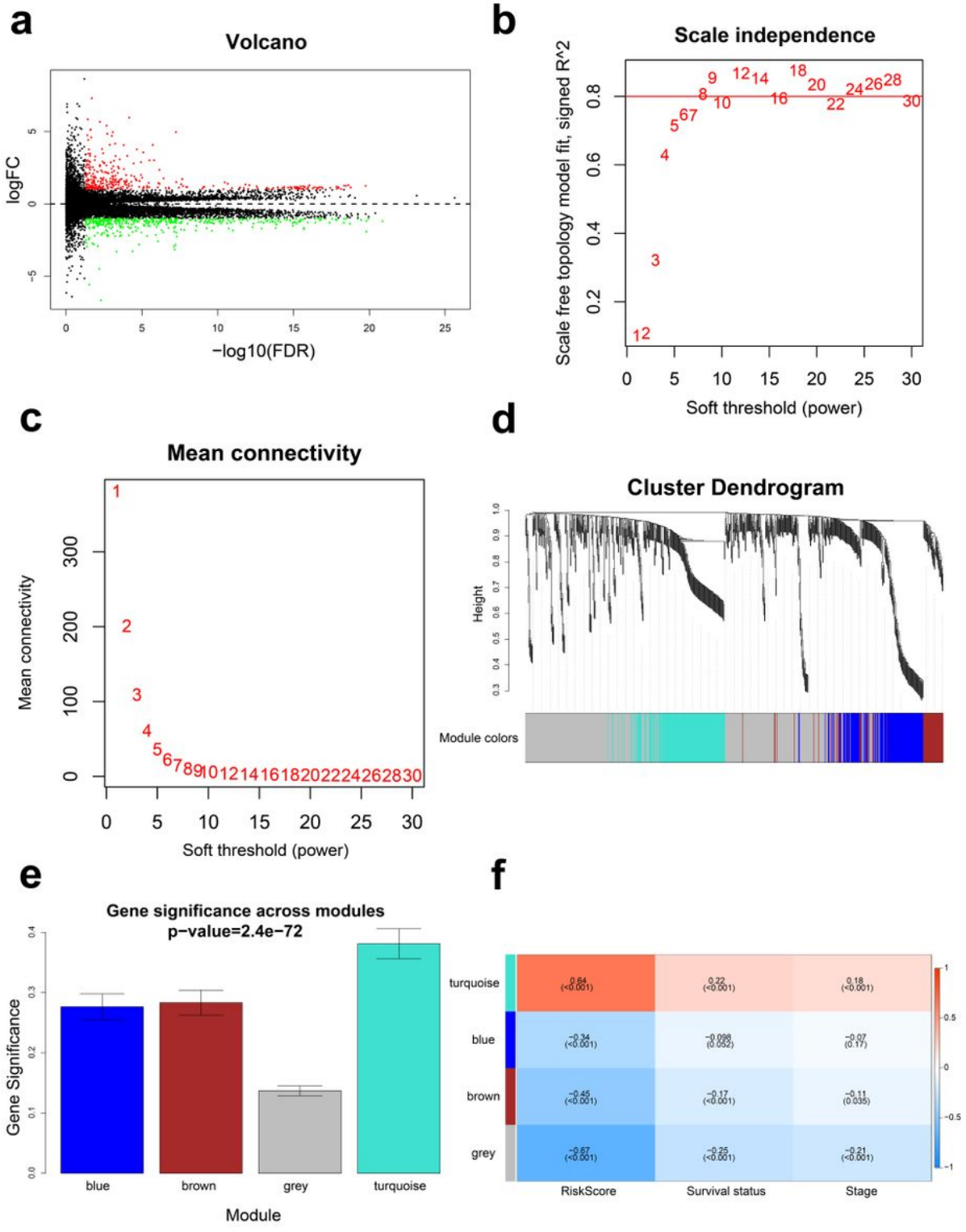

Figure 7

Identification of the gene modules using WGCNA (a) Volcano plot visualizing741 DEGs (335 up-regulated and 406 down-regulated) between the high-risk and low-risk groups. (b)Scale-free topology fit index with different soft threshold powers. (c) Mean connectivity with different soft threshold powers. (d) Clustering dendrogram of the gene modules on the basis of a topological overlap-derived dissimilarity, and different colors referring to different modules. (e) Gene significances of different modules. (f) The correlation coefficients with p values between MEs and clinical characteristics.

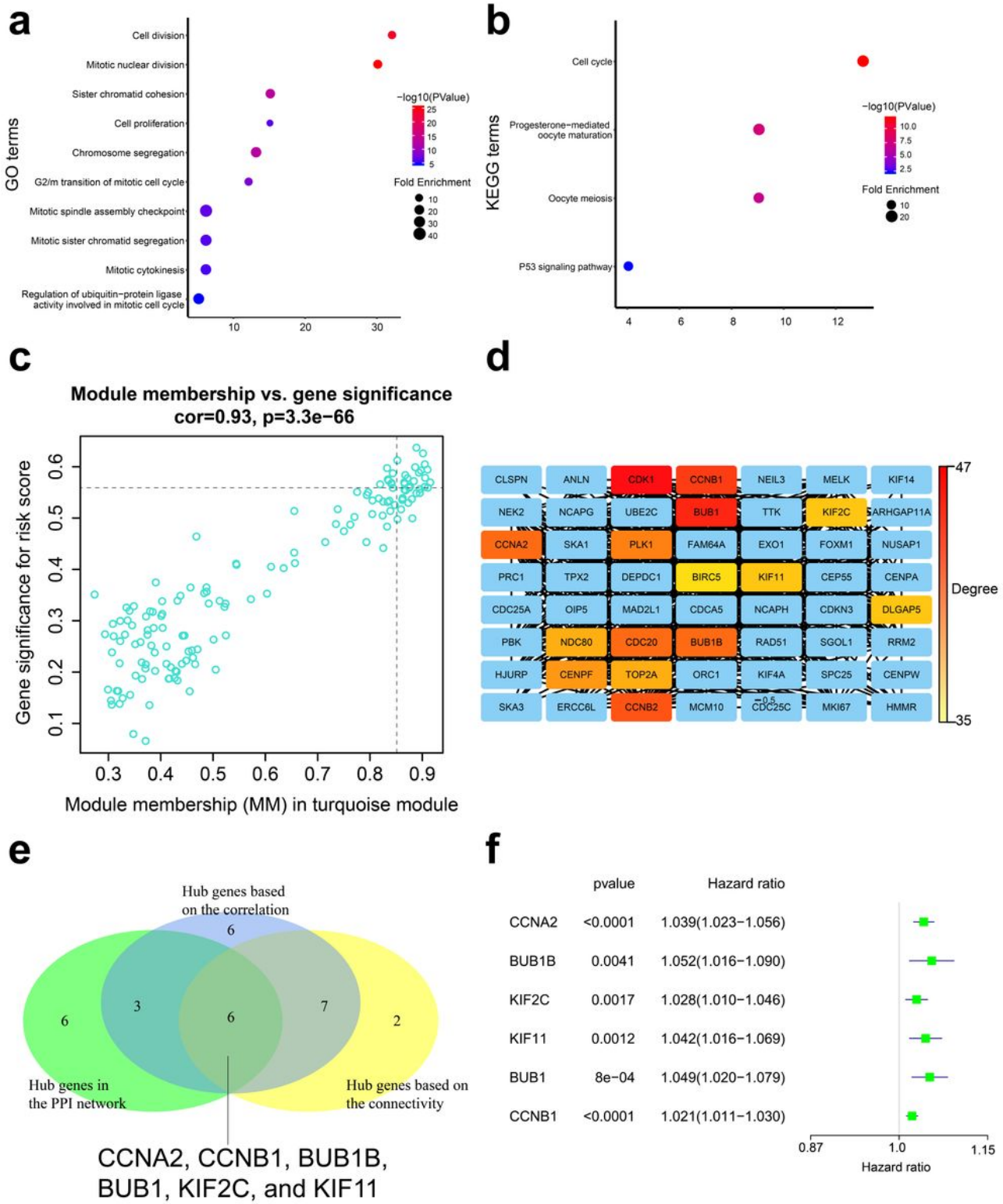

Figure 8

Enrichment analyses of the genes in the turquoise module, and the determination of hub genes (a) Go enrichment analysis for genes in the turquoise module. (b)KEGG enrichment analysis for genes in the turquoise module. (c)Genes in the turquoise module correlated with risk score and turquoise module, with cutoff value set as GS > 0.55 and MM > 0.85. (d) PPI analysis for all genes in the turquoise module using STRING database and Cytoscape software, and the identification of the top 15 genes with the highest degree. (e) Venn diagram visualizing the six common genes when three lists of genes were intersected. (f)The forest plot visualizing the association of six hub genes with OS status assessed by univariate Cox regression model.

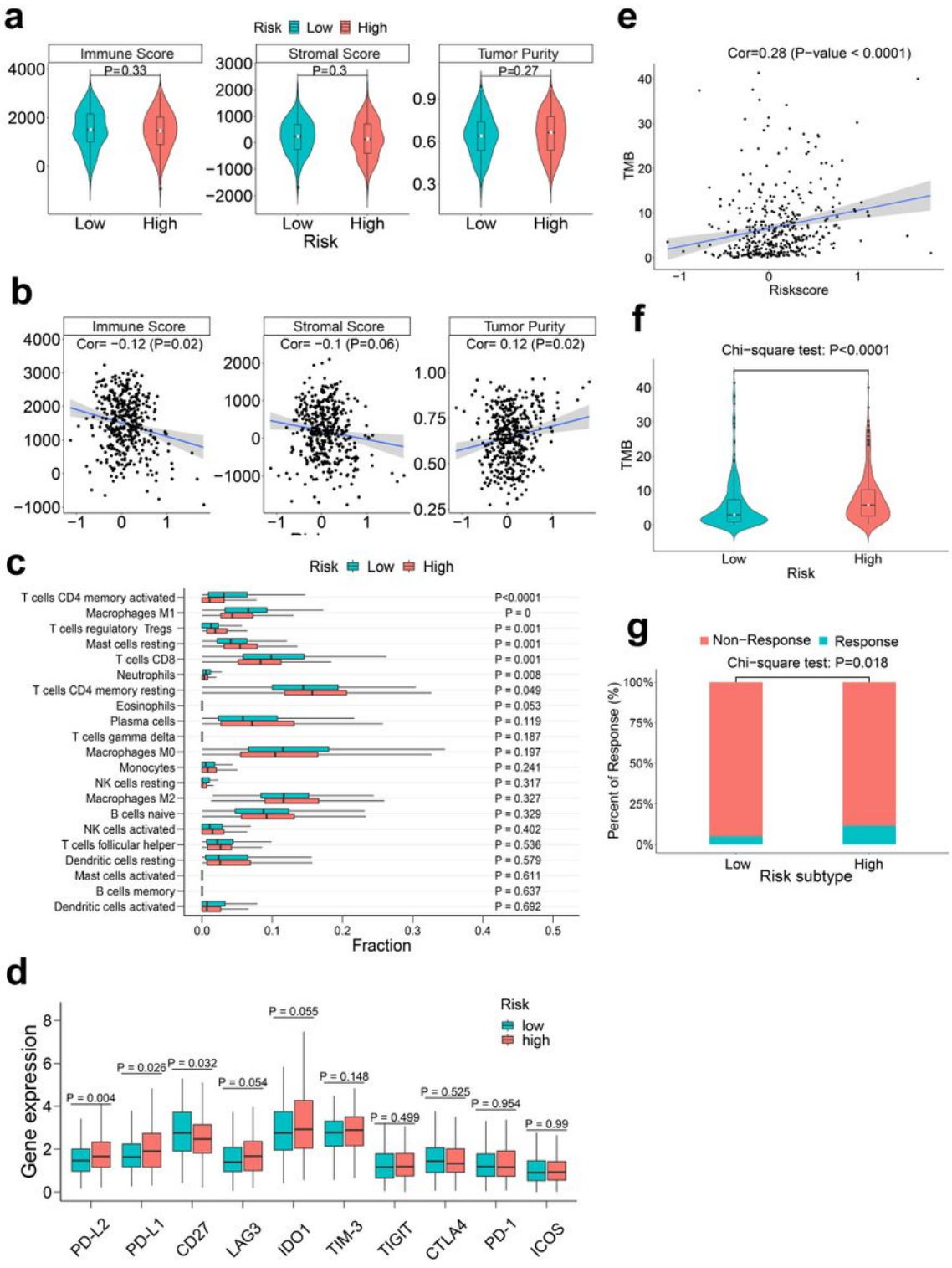

Figure 9

The association of the m6A signature with immunity (a) Boxplot visualizing the differences of the immune microenvironment (immune score, stromal score and tumor purity) between the high-risk and low-risk groups.(b) Spearman correlation analysis between the immune microenvironment (immune score, stromal score and tumor purity) and risk scores. (c)Relative proportion of immune infiltration between the high-risk and low-risk groups.(d) Boxplot visualizing the differences of the immune

checkpoint molecules between the high-risk and low-risk groups. (e)Spearman correlation analysis between TMB and risk scores.(f)Violin plot visualizing the difference of TMB between the high-risk and low-risk groups. (g) Boxplot visualizing the proportions of response between the high-risk and low-risk groups.