



N6-Methyladenosine RNA Methylation Regulator-Related Alternative Splicing (AS) Gene Signature Predicts Non-Small Cell Lung Cancer Prognosis

Zhenyu Zhao^{1,2}, Qidong Cai^{1,2}, Pengfei Zhang^{1,2}, Boxue He^{1,2}, Xiong Peng^{1,2}, Guangxu Tu^{1,2}, Weilin Peng^{1,2}, Li Wang^{1,2}, Fenglei Yu^{1,2} and Xiang Wang^{1,2*}

¹Department of Thoracic Surgery, The Second Xiangya Hospital of Central South University, Changsha, China, ²Hunan Key Laboratory of Early Diagnosis and Precise Treatment of Lung Cancer, The Second Xiangya Hospital of Central South University, Changsha, China

OPEN ACCESS

Edited by:

Arzucan Ozgur,
Boğaziçi University, Turkey

Reviewed by:

Junguk Hur,
University of North Dakota,
United States
Pinar Pir,
Gebze Technical University, Turkey

*Correspondence:

Xiang Wang
wangxiang@csu.edu.cn

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 22 January 2021

Accepted: 12 May 2021

Published: 11 June 2021

Citation:

Zhao Z, Cai Q, Zhang P, He B, Peng X,
Tu G, Peng W, Wang L, Yu F and
Wang X (2021) N6-Methyladenosine
RNA Methylation Regulator-Related
Alternative Splicing (AS) Gene
Signature Predicts Non-Small Cell
Lung Cancer Prognosis.
Front. Mol. Biosci. 8:657087.
doi: 10.3389/fmolb.2021.657087

Aberrant N6-methyladenosine (m6A) RNA methylation regulatory genes and related gene alternative splicing (AS) could be used to predict the prognosis of non-small cell lung carcinoma. This study focused on 13 m6A regulatory genes (METTL3, METTL14, WTAP, KIAA1429, RBM15, ZC3H13, YTHDC1, YTHDC2, YTHDF1, YTHDF2, HNRNPC, FTO, and ALKBH5) and expression profiles in TCGA-LUAD ($n = 504$) and TCGA-LUSC ($n = 479$) datasets from the Cancer Genome Atlas database. The data were downloaded and bioinformatically and statistically analyzed, including the gene ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses. There were 43,948 mRNA splicing events in lung adenocarcinoma (LUAD) and 46,020 in lung squamous cell carcinoma (LUSC), and the data suggested that m6A regulators could regulate mRNA splicing. Differential HNRNPC and RBM15 expression was associated with overall survival (OS) of LUAD and HNRNPC and METTL3 expression with the OS of LUSC patients. Furthermore, the non-small cell lung cancer prognosis-related AS events signature was constructed and divided patients into high- vs. low-risk groups using seven and 14 AS genes in LUAD and LUSC, respectively. The LUAD risk signature was associated with gender and T, N, and TNM stages, but the LUSC risk signature was not associated with any clinical features. In addition, the risk signature and TNM stage were independent prognostic predictors in LUAD and the risk signature and T stage were independent prognostic predictors in LUSC after the multivariate Cox regression and receiver operating characteristic analyses. In conclusion, this study revealed the AS prognostic signature in the prediction of LUAD and LUSC prognosis.

Keywords: non-small cell lung cancer, m6A, alternative splicing, The Cancer Genome Atlas, prognostic signature

INTRODUCTION

Lung cancer is still the most significant health burden in the world, accounting for 14% of all newly diagnosed cancer cases as the second most common cancer and 18% of all cancer-related deaths as the leading cause of cancer death globally in 2018 and 2020 (de Martel et al., 2020; Sung et al., 2021). Lung cancer is also prevalent and the leading cause of cancer death in men (Sung et al., 2021). Histologically, lung cancer can be divided into small cell lung cancer and non-small cell lung cancer (NSCLC), and the latter accounts for 85% of all lung cancer cases, and the overall 5-year survival rate of lung cancer remains to be approximately 15% (Balata et al., 2019). NSCLC can be further classified as lung squamous cell carcinoma (LUSC), lung adenocarcinoma (LUAD), and larger cell carcinoma; however, LUAD and LUSC are the main histological subtypes of NSCLC (Tanoue et al., 2015) and major contributors to NSCLC morbidity and mortality (Hirsch et al., 2017). The outcome data were from our most recent advancement and improvement in early detection, prevention, improved surgical procedures, neoadjuvant therapy, immunotherapy, and targeted therapy. To date, treatment of NSCLC is dependent on the stage of disease at diagnosis, and early-staged NSCLC could be surgically cured, whereas the advanced staged diseases can only be subjected to chemotherapy, radiation therapy, immunotherapy, and/or targeted therapy (Maconachie et al., 2019; Planchard et al., 2018) and their prognosis is, therefore, still poor, approximately less than 5–7% at the best according to the American Cancer Society data (<https://www.cancer.org/cancer/lung-cancer/detection-diagnosis-staging/survival-rates.html>) or after advanced therapy (Zhang et al., 2021). Thus, the search and development of biomarkers for early detection and prediction of prognosis and treatment outcome are urgently needed to effectively conquer this now deadly disease clinically.

Newly transcribed RNA could undergo different chemical modifications and N6-methyladenosine (m6A) is the most prevailing one in polyadenylated RNAs (Bokar et al., 1997). Methylation of the adenosine is directed in cells by a large m6A methyltransferase complex containing METTL3 as the SAM-binding subunit (Bokar et al., 1997). The biological functions of m6A are through a group of RNA-binding proteins that can specifically recognize the methylated adenosine on RNA molecules to regulate cell activities (Ji et al., 2018), for example, N6-methyladenosine (m6A) RNA modification could regulate RNA splicing, stability, translocation, and translation and therefore, to influence gene expression and functions in cells (Deng, et al., 2018). These binding proteins to m6A are regarded as the m6A readers, and m6A methyltransferases are considered as the writers, whereas demethylases are considered as the erasers. Altogether, these proteins form a complex mechanism of m6A regulation in which writers and erasers determine the distributions of m6A on RNA, whereas readers mediate m6A-dependent functions (Liu et al., 2014; Wang et al., 2014). Deregulation of the m6A on an RNA molecule has been implicated in the development of various human cancers (Liu et al., 2014; Wang et al., 2014). According to the recent studies,

there were 13 m6A regulator genes confirmed to affect cancer progression, including the “writer” (KIAA1429, METTL3, METTL14, RBM15, WTAP, and ZC3H13), the “readers” (HNRNPC, YTHDC1, YTHDC2, YTHDF1, and YTHDF2), and the “erasers” (ALKBH5 and FTO) (Zhang et al., 2020c; Zhang et al., 2020b). Further studies of the m6A regulator genes showed that the m6A regulator genes were also the mRNA splicing factors for gene alternative splicing (GAS) and the m6A regulator genes could interact with the AS events (Kasowitz et al., 2018; Yoshimi et al., 2019). Human cancer cells frequently showed the GAS events, which were regulated by the m6A regulators (Dai et al., 2018). For example, METTL3 was able to regulate the mRNA alternative splicing by the p53 pathway (Alarcón et al., 2015). YTHDC1 could recruit SRSF10 to its target mRNA regions and modulate their exon skipping (Xiao et al., 2016). Abnormal splicing factor expression in normal cells could lead to the formation of the specific pro-oncogenic splicing subtypes and carcinogenesis (Kasowitz et al., 2018).

Indeed, gene alternative splicing (GAS), a posttranscriptional process, subjects a single pre-mRNA molecule to splice into different exons for coding and expression of various protein isoforms (David and Manley, 2008). A molecular structure called a spliceosome is assembled on the pre-mRNA to join the exons together at the splicing site to form a particular mRNA molecule, while the introns are discarded (Papasaiakas and Valcárcel, 2016). The assembly of spliceosomes on pre-mRNA is usually affected by the SF and some exons (alternative exons) are variably incorporated into mRNA; thus, under different alternative splicing patterns (including exon skip, retained intron, alternate donor site, alternate acceptor site, alternate promoter, alternate terminator, and mutually exclusive exons), the whole exons of a gene could be spliced into mRNA or excluded (Hanahan and Weinberg, 2011). The different GAS events could lead to the diversity of protein functions and normal GAS will maintain normal cell functions, which is mediated by the production of the diverse and multifunctional proteome to ensure “normal” RNA molecules to maintain normal cell functions; however, abnormal GAS will promote tumorigenesis and cancer development (Bonnal et al., 2020), which could be mainly due to the up or downregulation of the related splicing factors, for example, alterations in the upstream signaling pathways or mutations in the splicing site sequences all lead to abnormal mRNA splicing (Li et al., 2019). Accumulating evidence suggests the contribution of abnormal GAS to cancer phenotypes, like increases in cell proliferation, angiogenesis, but inhibition of apoptosis and drug resistance, and the GAS events form a novel and separate hallmark in cancer (Bonnal et al., 2020; Urbanski et al., 2018). For example, in gastric cancer, abnormal GAS could lead to activation of tumor cell invasion and metastasis (Pio and Montuenga, 2009; Sun and Ma, 2019), while in breast cancer, abnormal GAS results in drug resistance (Yang et al., 2019). In lung cancer, the GAS events could be used as biomarkers for tumor diagnosis (Sholl, 2017). Aberrant *BCL2L1*, *MDM2*, *MDM4*, *NUMB*, and *MET* mRNA splicing occurred in lung cancer and altered cell apoptosis, proliferation, and cohesion (Cooper, Black). Thus, further

investigation related to the abnormal GAS events to lung tumorigenesis (Coomer et al., 2019) and novel strategy for cancer targeting therapy (Frankiw et al., 2019) as well as biomarkers for various human cancers, including NSCLC (Li et al., 2017; Paschalis et al., 2018; Yang et al., 2019; Liu et al., 2020; Zhang et al., 2020a).

In this study, we focused on the m6A-related splicing factors for aberrant expression and AS events to associate them with NSCLC clinicopathological and prognostic data from patients using the online The Cancer Genome Atlas (TCGA) data. We then explored the role of the abnormally expressed m6A-related splicing factors in the regulation of the GAS events and constructed the risk signature of these factors to predict NSCLC prognosis after the gene ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. This study could provide a novel insight into the discovery of biomarkers in the prediction of NSCLC prognosis and possibly the underlying molecular mechanisms of NSCLC oncogenesis and development.

MATERIALS AND METHODS

Data Download and Analysis

In this study, we first searched and downloaded differential gene expression profiles in LUAD and LUSC tissue specimens from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>). The corresponding clinicopathological data were subsequently downloaded from the University of California Santa Cruz database (<https://xena.ucsc.edu/>), which included 514 LUAD and 488 LUSC tissue samples. However, patients with incomplete clinical information and follow-up duration less than 30 days were excluded from our data analysis, resulting in 504 LUAD and 479 LUSC samples in this study. Moreover, the gene alternative splicing (GAS) events in LUAD and LUSC were download from TCGA Splice Seq (<https://bioinformatics.mdanderson.org/TCGASpliceSeq/PSIdownload>) and then calculated for the percent spliced in index (PSI) value, a quantifiable GAS indicator after the comparison of single and multiple samples between subgroups, that is, calculation of the percentage of GAS value for each GAS event, which was typically used to quantify GAS events according to a previous study (Lin and Krainer, 2019). We downloaded the contents that included seven main GAS types, that is, the exon skip (ES), retained intron (RI), alternate donor site (AD), alternate acceptor site (AA), alternate promoter (AP), alternate terminator (AT), and mutually exclusive exons (ME).

Selection and Analysis of N6-Methyladenosine RNA Methylation Regulatory Genes

In this study, we selected 13 m6A RNA methylation regulatory genes, that is, N6-adenosine-methyltransferase 70-kDa subunit (METTL3), methyltransferase-like 14 (METTL14), Wilms' tumor-1 associated protein (WTAP), KIAA1429, RNA-binding protein 15 (RBM15), zinc finger CCCH

domain-containing protein 13 (ZC3H13), YTH domain-containing protein 1 (YTHDC1), YTHDC2, YTH domain family, member 1 (YTHDF1), YTHDF2, heterogeneous nuclear ribonucleoproteins C1/C2 (HNRNPC), fat mass and obesity-associated protein (FTO), and m6A demethylase ALKB homolog 5 (ALKBH5). We assessed their role in diagnosis, progression, and prognosis of LUAD and LUSC, that is, we first imported data on these m6A regulators into Cytoscape software [version 3.8.2 (Shannon et al., 2003)] and analyzed the data using the ClueGO plugin. After that, we performed the gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and network analysis, and the GO terms included the cellular component (CC), molecular function (MF), biological process (BP), an adjusted $p < 0.05$ as a statistically significant value. Afterward, we utilized the "LIMMA" package (Ritchie et al., 2015) to select and analyze the differentially expressed m6A RNA methylation regulatory genes in LUAD and LUSC tissue specimens using the cutoff value of $|\log_2 \text{fold change (FC)}| \geq 1$ and adjusted p -value < 0.05 . The "euclidean" and "ward.D2" methods were utilized to cluster the tumor samples, and the "p heat map" R package was used to plot the differential expression analysis results and cluster analysis results. Spearman's correlation analysis was performed to analyze the correlation between the clusters and clinical traits. We also performed the univariate and multivariate Cox regression analyses to predict the association of these m6A RNA methylated regulatory genes with the overall survival of patients using the "survival" R package (Rizvi et al., 2019).

Association of Gene Alternative Splicing Events With Overall Survival of Patients

We utilized the "WGCNA" (weighted gene co-expression network analysis) package (Langfelder and Horvath, 2008) to associate the GAS events with the overall survival of LUAD and LUSC patients. The WGCNA package is able to analyze thousands of the most varied genetic information to identify the sets of genes to associate with tumor phenotypes (Meng et al., 2019); thus, this tool allowed us to analyze the information regarding the GAS event data in association with clinical traits data and profile of m6A regulatory gene expressions, but avoided any unnecessary procedures for multiple hypothesis testing and corrections. In this analysis, we first estimated the standard scale-free network according to formula A to generate the adequate β value (appropriate soft threshold power). We then constructed the weighted adjacency matrix using the formula B and converted the data into a topological overlap matrix (TOM). After that, we utilized the dynamic tree cutting method according to the hierarchical clustering to identify the modules that highly correlated with GAS events. The GAS event then used 1-TOM as distance measurement with the depth (the cutoff value of 2) and the minimum size (the cutoff value of 60). After that, the highly similar modules were fused by clustering and height truncation of 0.3 according to previous studies (Niemira et al., 2019; Wan et al., 2018; Xie and Xie, 2019). Last, we performed the Spearman's correlation and module

eigengenes analysis of m6A regulator genes expression for association with the clinical traits and prognosis of 487 LUAD patients. Significant data on the association of this m6A regulator genes expression with clinicopathological data were further analyzed according to a previous study using the |correlation coefficients| between m6A regulators and GAS events module more than 0.4 and adjusted $p < 0.05$ (Langfelder and Horvath, 2008).

In the LUSC cohort of patient's data, we selected and further analyzed the modules of the most significant correlation between clinical features and m6A regulator genes, that is, the |correlation coefficients| between m6A regulators and GAS events module more than 0.4 and adjusted $p < 0.05$. We then performed univariate and multivariate Cox Regression analyses to screen associate the GAS events with the overall survival of LUAD and LUSC patients.

The formula A: $A_{ij} = \text{power}(S_{ij}, \beta) = |S_{mn}|^\beta$ (i and j represent the GAS event of i and j, respectively, while m and n were the numbers of node connections, and β was the appropriate soft threshold power). The formula B: $TOM_{ij} = \frac{\sum_u A_{iu}A_{ju}+A_{ij}}{\min(K_i, K_j)+1-A_{ij}}$ (i and j represent the GAS event of i and j, respectively, while the letter u represents clinical traits and prognostic information).

The Gene Ontology Term and Kyoto Encyclopedia of Genes and Genomes Pathway Analyses of Overall Survival-Related Gene Alternative Splicing Genes

After screening the m6A-related GAS events, we performed the GO terms and KEGG pathways analysis of these overall survival-related GAS genes. Specifically, we imported data on the m6A-related GAS genes and m6A regulator genes into Cytoscape software (3.8.2) and analyzed them using the ClueGO plugin. After associating the overall survival of patients, we performed the GO terms and KEGG pathway analyses of these GAS events-related genes and the GO terms included the cellular component (CC), molecular function (MF), and biological process (BP) using an adjusted $p < 0.05$ as a statistically significant value according to a previous study (Amado et al., 2014). After that, we constructed the functional network of these corresponding genes using the Cytoscape software.

Risk Model Construction

We constructed the risk model using the LASSO Cox regression analysis that could prevent any overfitting of the overall survival-related genes according to a previous study (Tang et al., 2017), and then performed the multivariate Cox regression analysis to predict the usefulness of these overall survival-related genes using the following formula:

$J = 1n \sum_i i = 1n (f(x_i) - y_i)^2 + \lambda \|w\|_1$ (the greater the value of J, the better the prediction value; the letter w indicates a globally optimal value of lost J).

Risk score = $\sum_{n=x}^n \text{coef}(X) * \text{PSI}(X)$ [Coef(X) is the coefficient of each GAS gene and PSI(X) is the PSI value of the AS genes].

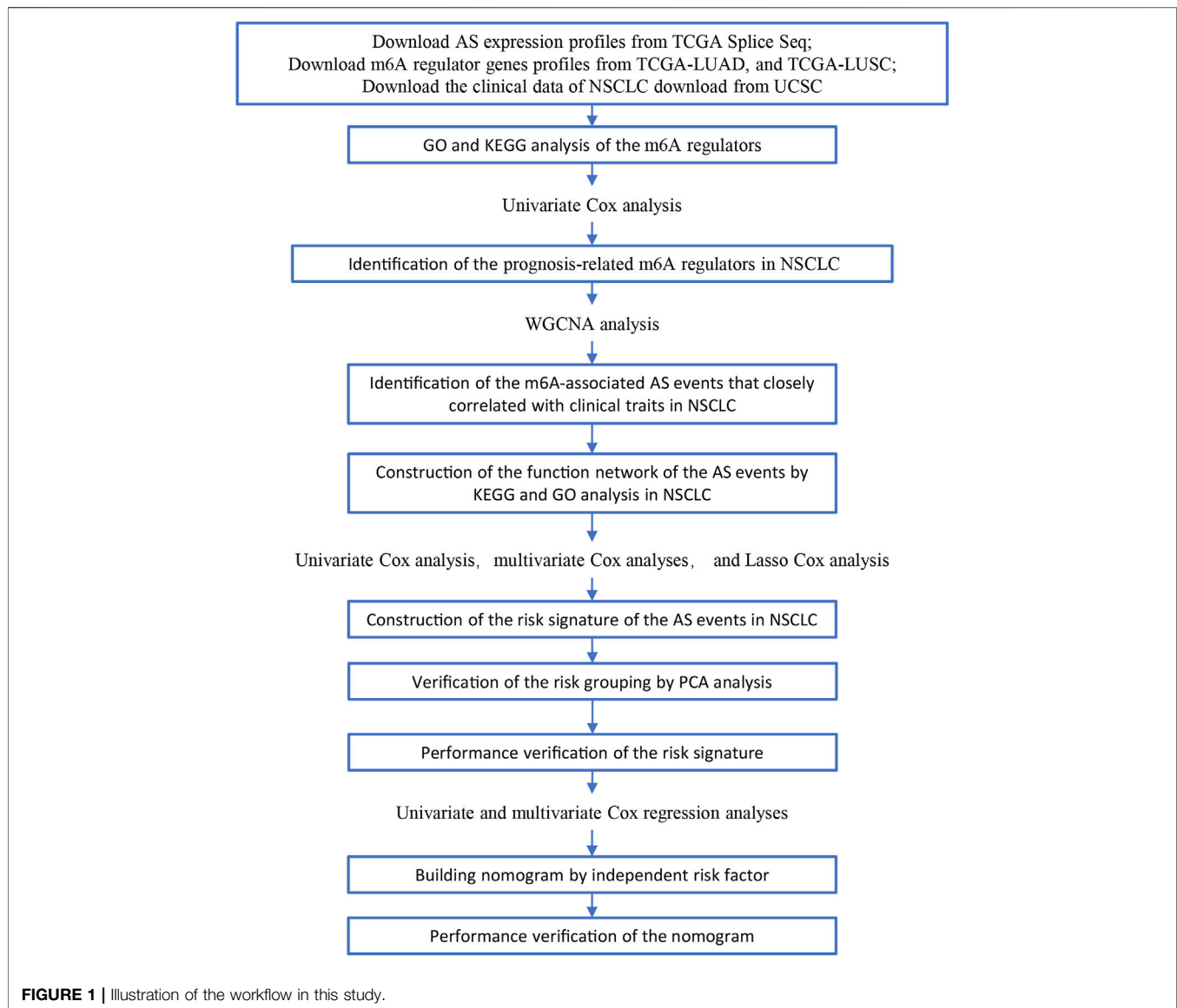
According to the hazard ratio (HR) values after the multivariate Cox regression analysis, we classified the m6A-related prognostic AS events into protective/risky AS events (HR > 1 as a risk factor; HR < 1 as a protective factor), and showed the Sankey diagram that plotted is by “ggalluvial, dplyr, and ggplot2” R packages (Graedel, 2019; Soh et al., 2019). According to the median value of the risk score of the signature of each cohort, we divided the LUAD and LUSC cohorts into two subgroups, that is, the high- and low-risk groups. We then utilized the “survival” and “survminer” package to calculate the survival significance of the high-/low-risk group in these NSCLC patients. We performed the Kaplan–Meier survival analysis and receiver operating characteristic (ROC) curves to further verify the predictive ability of the risk signature using the “survivalROC” package and “Survival” package in R (Park et al., 2004). The concordance index (C-index) was used to validate and quantify the discrimination ability of the risk signature. At last, we performed the univariate and multivariate Cox regression analyses to assess whether these risk models and clinicopathological features were independent predictors for the survival of LUAD and LUSC patients.

Predictive Nomogram Construction

After the univariate and multivariate Cox regression analyses, we used the “RMS” package to construct the nomogram of the independent risk factors (Zhang et al., 2019). We then performed the Wilcoxon rank-sum test to verify the association of the risk model with clinical characteristics (using an adjusted p -value < 0.05 as the statistical significance cutoff). For the construction of the predictive nomogram, we utilized the calibration curves to evaluate and validate the application ability of the nomogram performance.

Statistical Analysis

The “Limma” R package was utilized to analyze the difference in gene expression profiles, while the “WGCNA” package was used to select m6A-related GAS events. Moreover, the univariate, LASSO and multivariate Cox regression analyses were performed to construct the risk signature, while the “Survival,” “survivalROC,” and “survminer” packages were utilized to verify the predictive efficacy of the risk signature in patients, while the area under the curve (AUC) value (ranged between 0.5 and 0.9) was used to assess the diagnostic ability of the risk signature (larger AUC value, better diagnostic value) (Verbakel et al., 2020). The Wilcoxon rank-sum and Spearman's correlation tests were used to analyze the subgroup differences, while the “RMS” R package was utilized to plot the nomogram and calibrate the analytic data. The GO terms and KEGG pathways enrichment analysis were performed using Cytoscape software (3.8.2) and the “ggplot” R package was to plot the resulting data. All statistical analyses were performed using R software (version 3.6.1). A two-sided p value < 0.05 was considered statistically significant, while an adjusted p -value < 0.05 was applied as the threshold to avoid missing any significant changes.



RESULTS

Characteristics of Lung Adenocarcinoma and Lung Squamous Cell Carcinoma Data and mRNA Splicing Events in The Cancer Genome Atlas Datasets

In this study, we searched and downloaded expression profiles of TCGA-LUAD and TCGA-LUSC from TCGA database (<https://portal.gdc.cancer.gov/>) and clinicopathological data from the University of California Santa Cruz database (<https://xena.ucsc.edu/>). We obtained 486 LUAD and 479 LUSC cases for our data analysis (Figure 1). In LUAD samples, there were 224 men and 262 women with a median age of 64.94 years old (arranged between 33 and 88 years). The patients were at the TNM stage of I/II ($n = 381$) and III/IV ($n = 105$) and 167 cases had lymph node tumor metastasis (Table 1). In LUSC cases, there were 353 men and 126 women with a median age of 64.294 years old (arranged between 39 and 90 years).

The patients were at the TNM stage of I/II ($n = 391$) and III/IV ($n = 88$) and 169 cases had lymph node tumor metastasis (Table 1). Our data analyses identified a total of 43,948 mRNA splicing events in LUAD tissue samples and 46,020 mRNA splicing events in LUSC samples. The exon skip (ES) events were the most GAS events in both LUAD and LUSC groups of samples (Supplementary Figure S1).

Association of N6-Methyladenosine RNA Methylation Regulatory Gene Expressions With Lung Adenocarcinoma and Lung Squamous Cell Carcinoma Prognosis

We focused on 13 m6A RNA methylation regulatory genes, including METTL3, METTL14, WTAP, KIAA1429, RBM15, ZC3H13, YTHDC1, YTHDC2, YTHDF1, YTHDF2, HNRNPC, FTO, and ALKBH5. We performed the GO terms and KEGG pathway analyses and found that these m6A regulators were

TABLE 1 | The univariate and multivariate cox regression analysis of clinicopathological data from TCGA-LUAD and LUSC.

Variables	Univariate analysis				Multivariate analysis			
	HR	HR.95L	HR.95H	Adj-p	HR	HR.95L	HR.95H	Adj-p
LUAD								
Age (≥ 65 / <65)	1.111	0.823	1.500	4.92E-01	1.204	0.884	1.640	2.38E-01
Gender (female/male)	1.111	0.826	1.494	4.86E-01	1.077	0.791	1.467	6.37E-01
Smoking (yes/no)	0.864	0.440	1.697	6.72E-01	1.863	0.878	3.955	1.05E-01
TNM stage (I + II/III + IV)	1.635	1.422	1.881	***	1.564	1.187	2.062	***
T	1.548	1.296	1.848	***	1.066	0.865	1.315	5.48E-01
N	1.962	1.461	2.634	***	1.108	0.722	1.701	6.39E-01
M	2.181	1.302	3.653	**	0.740	0.355	1.543	4.21E-01
Risk score	1.441	1.346	1.543	***	1.390	1.281	1.508	***
LUSC								
Age (≥ 65 / <65)	1.017	0.999	1.035	6.40E-02	1.028	1.007	1.050	*
Gender (female/male)	1.095	0.782	1.534	5.97E-01	1.225	0.834	1.800	3.01E-01
TNM stage (I + II/III + IV)	1.283	1.079	1.526	**	1.132	0.726	1.765	5.83E-01
T	1.352	1.121	1.630	**	1.434	1.083	1.897	**
N	1.170	0.952	1.438	1.35E-01	1.092	0.735	1.622	6.62E-01
M	2.455	0.905	6.659	7.77E-02	1.111	0.279	4.420	8.82E-01
Risk score	1.054	1.040	1.069	***	1.884	1.473	2.386	***

LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; TNM: tumor-node metastasis; HR: hazard ratio. * represents $p < 0.05$; ** represents $p < 0.05$; *** represents $p < 0.001$.

significantly enriched in the mRNA splicing spliceosome biology process, RNA methylation biology process, and RNA destabilization biology process (Figure 2A, Supplementary Table S1). We then performed the Wilcoxon signed-rank test and found that KIAA1429, HNRNPC, RBM15, METTL3, YTHDF1, YTHDF2, and YTHDC1 were differentially expressed between normal and LUAD tissues (Figure 2B). YTHDF1, YTHDF2, WTAP, KIAA1429, RBM15, METTL3, METTL14, FTO, HNRNPC, and ZC3H13 were differentially expressed between normal tissues and LUSC tissues (Figure 2D). Furthermore, KIAA1429, HNRNPC, RBM15, METTL3, YTHDF1, and YTHDF2 were all highly expressed in both LUAD and LUSC tissues (Figures 2B,D). The LUAD and LUSC patients were clustered into four groups according to the expression of m6A regulators in the heat map (Supplementary Figure S2A,B). The correlation analysis suggested that the differentially expressed m6A regulators were associated with status, smoking, TNM stage, and N stage in LUAD samples, and the differentially expressed m6A regulators were associated with status and age in LUSC samples (Table 2). These results suggested that m6A regulators play an important role in NSCLC development. The univariate Cox regression data revealed that HNRNPC and RBM15 expression were able to predict overall survival (OS) of LUAD patients (Figure 2C), while HNRNPC and METTL3 expression were associated with the OS of LUSC patients (Figure 2E); thus, these three m6A regulators genes were subjected to the subsequent analysis for association with OS of NSCLC patients as the splicing factors (Figures 2C,E).

Association of N6-Methyladenosine RNA Methylation Regulatory Genes With Lung Adenocarcinoma and Lung Squamous Cell Carcinoma Clinical Features

After that, we correlated the GAS events with the weighted gene co-expression network, and they were consistent with the scale-

free network (Supplementary Figure S3). The hierarchical clustering analysis of the samples using the Euclidean distance showed log₁₀-transformed RNA-seq fractional counts (Supplementary Figure S4), while the dynamic tree cutting method identified the modules with a similar expression spectrum and combine similar modules (Figures 3A,C, Supplementary Figure S5). We then utilized the “WGCNA” package to analyze the GAS events and Spearman’s correlation test to associate the expression of m6A regulator genes with clinical traits. The data showed that the MEBrown module was significantly associated with expression of the m6A regulator genes (RBM15, $p = 3e-24$, the |correlation coefficient| = -0.44), gender ($p = 0.03$, the coefficient correlation = -0.1), and tobacco smoking ($p = 0.006$, the coefficient correlation = -0.12) of LUAD patients (Figure 3B, Supplementary Figure S6A). Furthermore, the MERed, MEblue, and MERoyalblue modules were significantly associated with expression of the m6A regulator genes (MERed, HNRNPC with an adj $p = 1e-24$ and the coefficient correlation = -0.44 ; MEblue, HNRNPC with an adj $p = 2e-27$ and the coefficient correlation = -0.47 ; MERoyalblue, HNRNPC with an adj $p = 3e-33$ and the coefficient correlation = -0.51). These three modules were also associated with the age of patients (MERed, adj $p = 0.007$ and the coefficient correlation = 0.12), the TNM stage (MERed, adj $p = 0.03$ and the coefficient correlation = -0.1), and the N stage (MEblue, adj $p = 0.04$ and the coefficient correlation = -0.092 ; MERoyalblue, adj $p = 0.03$ and the coefficient correlation = 0.098 ; in Figure 3D, Supplementary Figure S6B). These results suggested that the m6A-related AS events in the MERed, MEblue, and MERoyalblue modules could predict NSCLC development and lymph node metastasis, while the age of patients might also affect the m6A-related AS events.

Furthermore, the MEBrown module included 1.102 GAS events, and the MERed, MEblue, and MERoyalblue modules included 1.5150 GAS events. The most significant enrichment

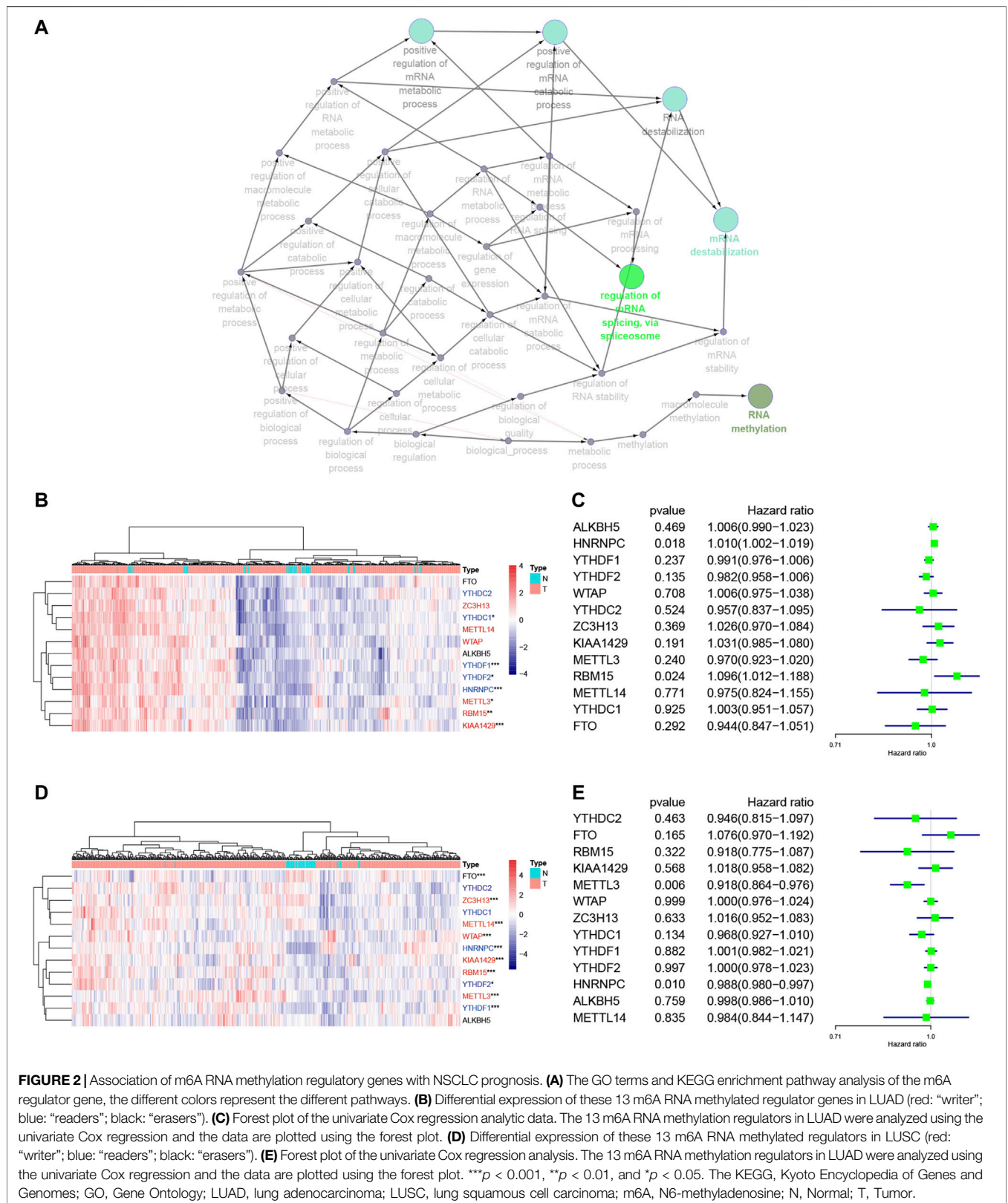


FIGURE 2 | Association of m6A RNA methylation regulatory genes with NSCLC prognosis. **(A)** The GO terms and KEGG enrichment pathway analysis of the m6A regulator gene, the different colors represent the different pathways. **(B)** Differential expression of these 13 m6A RNA methylated regulator genes in LUAD (red: “writer”; blue: “readers”; black: “erasers”). **(C)** Forest plot of the univariate Cox regression analytic data. The 13 m6A RNA methylation regulators in LUAD were analyzed using the univariate Cox regression and the data are plotted using the forest plot. **(D)** Differential expression of these 13 m6A RNA methylated regulators in LUSC (red: “writer”; blue: “readers”; black: “erasers”). **(E)** Forest plot of the univariate Cox regression analysis. The 13 m6A RNA methylation regulators in LUSC were analyzed using the univariate Cox regression and the data are plotted using the forest plot. *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$. The KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; m6A, N6-methyladenosine; N, Normal; T, Tumor.

TABLE 2 | The correlation analysis of clinical traits and m6A clusters from TCGA-LUAD and LUSC.

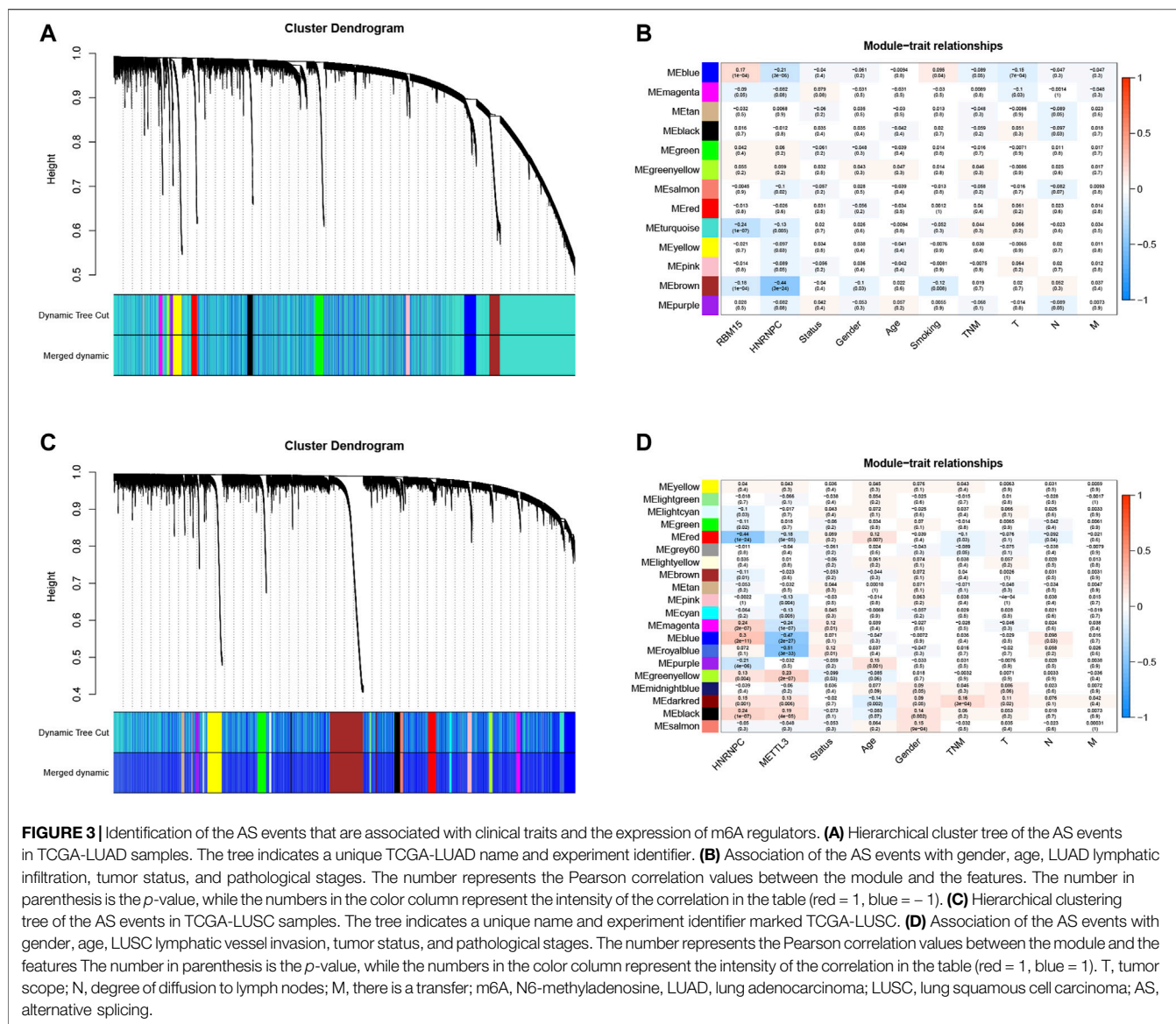
		Overall	Cluster 1	Cluster 3	Cluster 4	Cluster 4	p
LUAD							
<i>n</i>		486	148	86	129	123	
Status (%)	Alive	366 (76.0)	123 (81.8)	61 (70.9)	76 (64.9)	106 (86.2)	***
	Dead	120 (24.0)	25 (18.2)	25 (29.1)	53 (35.1)	17 (13.8)	
Gender (%)	Female	262 (53.7)	74 (50.0)	52 (60.5)	69 (52.0)	67 (54.5)	0.483
	Male	224 (46.3)	74 (50.0)	34 (39.5)	60 (48.0)	56 (45.5)	
Age (%)	<65	215 (45.5)	49 (43.9)	40 (46.5)	65 (43.0)	61 (49.6)	0.71
	≥65	271 (54.5)	99 (56.1)	46 (53.5)	64 (57.0)	62 (50.4)	
Smoking (%)	No	13 (2.6)	1 (0.7)	7 (8.1)	1 (0.7)	4 (3.3)	**
	Yes	495 (97.4)	147 (99.3)	79 (91.9)	150 (99.3)	119 (96.7)	
Stage (%)	Stage I	263 (55.7)	85 (56.1)	38 (44.2)	60 (54.3)	80 (65.0)	*
	Stage II	118 (23.4)	38 (26.4)	25 (29.1)	33 (21.9)	22 (17.9)	
	Stage III	80 (15.7)	14 (9.5)	20 (23.3)	30 (19.9)	16 (13.0)	
	Stage IV	25 (5.1)	11 (8.1)	3 (3.5)	6 (4.0)	5 (4.1)	
T (%)	T1	165 (34.1)	59 (30.4)	26 (30.2)	30 (34.4)	50 (40.7)	0.104
	T2	256 (52.8)	74 (58.1)	47 (54.7)	73 (48.3)	62 (50.4)	
	T3	44 (9.1)	7 (5.4)	10 (11.6)	16 (10.6)	11 (8.9)	
	T4	21 (4.1)	8 (6.1)	3 (3.5)	10 (6.6)	0 (0.0)	
N (%)	N0	319 (67.3)	102 (69.6)	49 (57.0)	77 (65.6)	91 (74.0)	*
	N1	91 (17.7)	29 (18.9)	19 (22.1)	25 (16.6)	18 (14.6)	
	N2	70 (13.8)	12 (8.1)	18 (20.9)	26 (17.2)	14 (11.4)	
	N3	6 (1.2)	5 (3.4)	0 (0.0)	1 (0.7)	0 (0.0)	
M (%)	M0	460 (94.9)	136 (91.9)	83 (96.5)	123 (96.0)	118 (95.9)	0.274
	M1	26 (5.1)	12 (8.1)	3 (3.5)	6 (4.0)	5 (4.1)	
LUSC							
<i>n</i>		479	84	169	125	101	
Status (%)	Alive	293 (61.3)	60 (71.4)	108 (64.1)	63 (50.4)	62 (61.4)	*
	Dead	186 (38.8)	24 (28.6)	61 (35.9)	62 (49.6)	39 (38.6)	
Gender (%)	FEMALE	126 (26.5)	24 (28.6)	37 (22.4)	36 (28.8)	29 (28.7)	0.516
	MALE	353 (73.5)	60 (71.4)	132 (77.6)	89 (71.2)	72 (71.3)	
Age (%)	<65	166 (34.8)	37 (44.0)	61 (36.5)	30 (24.0)	38 (37.6)	*
	≥65	313 (65.2)	47 (56.0)	108 (63.5)	95 (76.0)	63 (62.4)	
Stage (%)	Stage I	235 (49.2)	35 (41.7)	83 (49.4)	60 (48.0)	57 (56.4)	0.176
	Stage II	156 (32.5)	33 (39.3)	54 (31.8)	46 (36.8)	23 (22.8)	
	Stage III	81 (16.9)	14 (16.7)	32 (18.8)	16 (12.8)	19 (18.8)	
	Stage IV	7 (1.5)	2 (2.4)	0 (0.0)	3 (2.4)	2 (2.0)	
T (%)	T1	109 (22.9)	16 (19.0)	34 (20.6)	32 (25.6)	27 (26.7)	0.567
	T2	281 (58.5)	51 (60.7)	103 (60.6)	67 (53.6)	60 (59.4)	
	T3	68 (14.2)	13 (15.5)	25 (14.7)	22 (17.6)	8 (7.9)	
	T4	21 (4.4)	4 (4.8)	7 (4.1)	4 (3.2)	6 (5.9)	
N (%)	N0	310 (64.8)	54 (64.3)	106 (62.9)	82 (65.6)	68 (67.3)	0.263
	N1	125 (26.0)	23 (27.4)	49 (28.8)	35 (28.0)	18 (17.8)	
	N2	39 (8.1)	6 (7.1)	14 (8.2)	6 (4.8)	13 (12.9)	
	N3	5 (1.0)	1 (1.2)	0 (0.0)	2 (1.6)	2 (2.0)	
M (%)	M0	472 (98.5)	82 (97.6)	169 (100.0)	122 (97.6)	99 (98.0)	0.264
	M1	7 (1.5)	2 (2.4)	0 (0.0)	3 (2.4)	2 (2.0)	

LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; T: tumor; N: node; M: metastasis; * represents $p < 0.05$; ** represents $p < 0.05$; *** represents $p < 0.001$.

of the GO terms and KEGG pathway analysis of LUAD cohort revealed the m6A-related AS events were significantly enriched in the GPCR signaling pathway, DNA metabolic process, DNA repair, cellular response to DNA damage stimulus, carbon-oxygen lyase activity, and cell adhesion molecule binding pathways, while the m6A-related AS events in LUSC cohort were significantly enriched in the TGF-beta signaling pathway, peptidyl-serine phosphorylation, peptidyl-serine modification, regulation of actin cytoskeleton, intracellular signaling by second messengers, and early endosome pathway (adj $p < 0.001$; **Figure 4, Supplementary Table S2**).

Association of N6-Methyladenosine-Related Alternative Splicing Events With Lung Adenocarcinoma and Lung Squamous Cell Carcinoma Prognosis

We first performed the univariate Cox regression analysis to identify m6A-related AS events for association with LUAD and LUSC prognosis. We found 292 prognostic AS events in LUAD and 922 prognostics AS events in LUSC ($p < 0.05$; **Figures 5A,D, Supplementary Table S3**). The LASSO Cox regression analysis confirmed 13 of the prognostic AS events in LUAD (**Figures 5B,C**) and 15 in LUSC (**Figures 5E,F**), while the multivariate

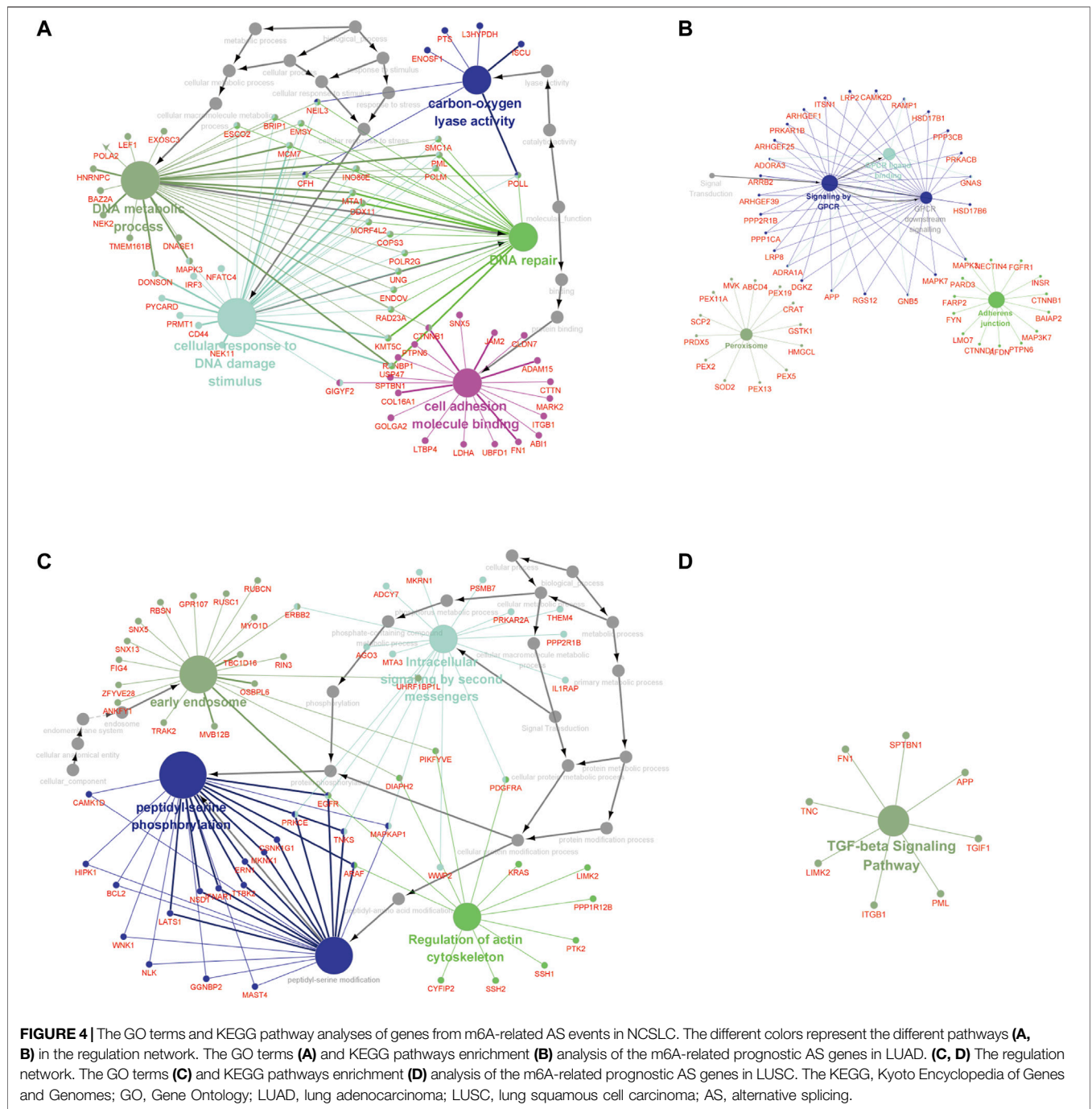


Cox regression analysis further confirmed seven of the prognostic AS events in LUAD and 14 in LUSC (Supplementary Table S4).

Non-Small Cell Lung Cancer Prognosis-Related Gene Alternative Splicing Events Signature

We utilized these seven and 14 AS genes in LUAD and LUSC, respectively, to further construct the LUAD and LUSC risk signature (Supplementary Table S5). The Sankey diagram shows that DGKZ|15540|AP and PMP22|39340|AP were the risky m6A-related AS events in LUAD (HR > 1), whereas ABCC6|34219|AT, KIAA0586|27718|ES, LDB1|12935|AP, RPS25|19054|ES, and S100A14|7729|AP were the protective m6A-related AS events in LUAD (HR < 1) (Figure 6A, Supplementary Table S5). Furthermore, AKR1E2|10639|ES

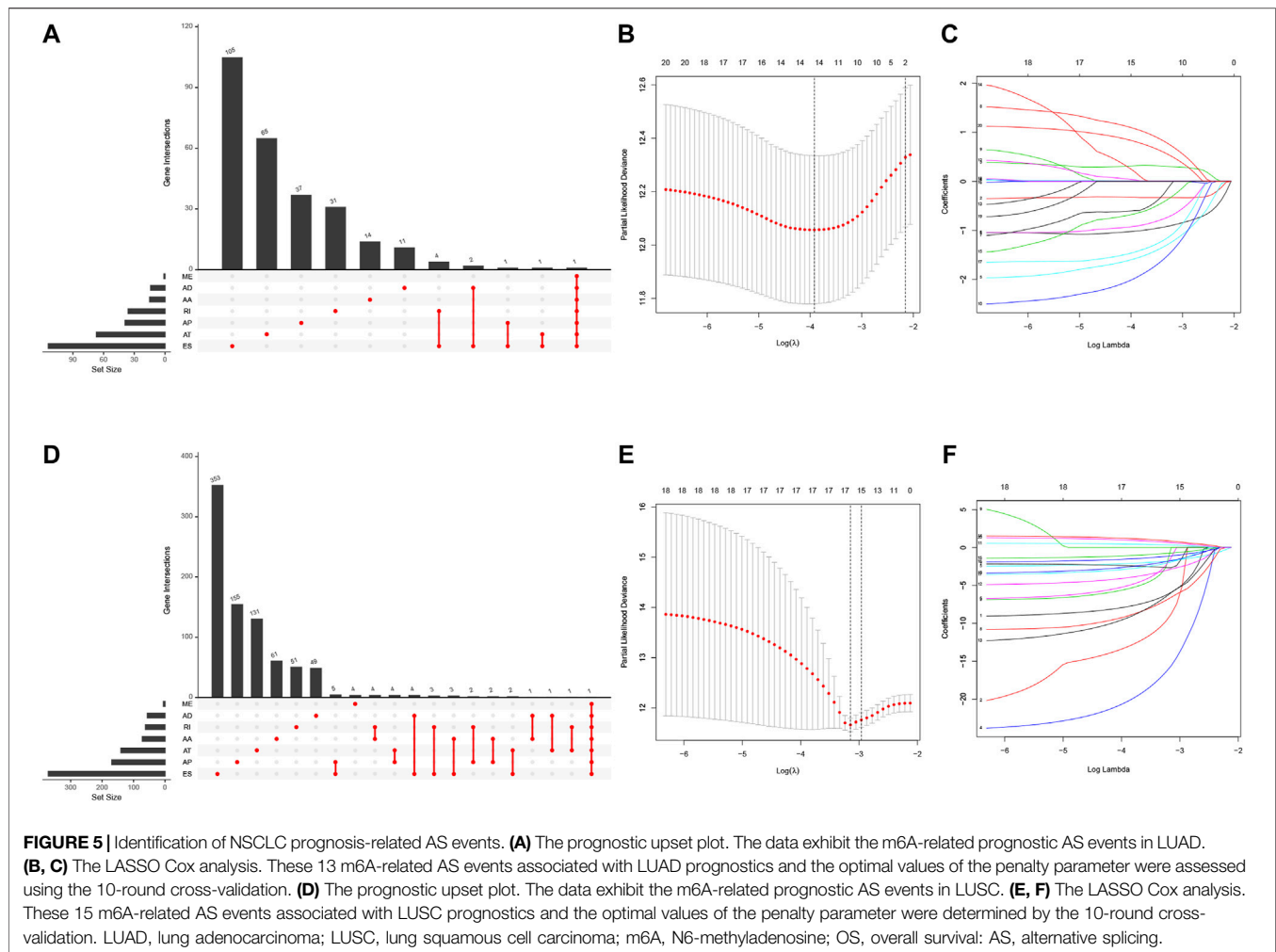
and SSH1|24258|ES were the risky m6A-related AS events in LUSC (HR > 1), whereas ALPK1|70369|ES, FAM63A|7531|AP, CHMP1A|38102|ES, TSTD2|87013|AT, KIAA1598|13239|AP, ASXL3|45046|AT, VPS37A|82796|ES, TOX2|59455|ES, ZNF544|52429|ES, NOL8|86863|ES, FAM124B|57772|AT, and PTCHD4|76446|AT were the m6A-related protective AS events in LUSC (HR < 1; Figure 6D, Supplementary Table S5). We then divided LUAD and LUSC patients into high- and low-risk groups according to their risk scores (high-risk LUAD group, *n* = 240; high-risk LUSC group, *n* = 239; low-risk LUAD group, *n* = 246; and low-risk LUSC group, *n* = 240; Supplementary Table S6). The Kaplan–Meier curve analyses showed that the high-risk group had a poorer OS than the low-risk group (*p* < 0.001; Figures 6B,E). The ROC analysis revealed that the AUC values were 0.868, 0.834, and 0.801 for 1-, 3-, and 5-year OS of LUAD, respectively, while the AUC values were 0.893, 0.824, and 0.849



for 1-, 3-, and 5-year LUSC OS, respectively (Figures 6C,F). The concordance index (C-index) of OS was 0.847 [95% confidence interval (CI): 0.788–0.847] in LUAD and 0.832 in LUSC (95% CI: 0.788–0.877; Supplementary Table S7).

Furthermore, we found that the risk signature (univariate Cox analysis, HR: 1.543, 95% CI: 1.441–1.346; $p < 0.001$; multivariate Cox analysis, HR: 1.390, 95% CI: 1.281–1.508; $p < 0.001$) and the TNM stage (univariate Cox analysis, HR: 1.635, and 95% CI: 1.422–1.881; $p < 0.001$; multivariate Cox analysis, HR: 1.564, 95% CI: 1.187–2.062; $p = 0.001$) were independent prognostic factors

of LUAD, while the risk signature (univariate Cox analysis, HR: 1.054, 95% CI: 1.040–1.069; $p < 0.001$; multivariate Cox analysis, HR: 1.884, 95% CI: 1.472–2.386; $p < 0.001$) and T stage (univariate Cox analysis, HR: 1.352, 95% CI: 1.121–1.630; $p = 0.002$; multivariate Cox analysis, HR: 1.434, 95% CI: 1.083–1.897; $p = 0.012$) were independent prognostic factors in LUSC (Figures 7A,B, Table 2). The risk curve and scatterplot of the risk score and survival of each NSCLC sample and the heat map of these AS genes in NSCLC samples are shown in Figures 7Aiii–v, iii–v.



Association of These Prognostic Signatures With Non-Small Cell Lung Cancer Clinicopathologies

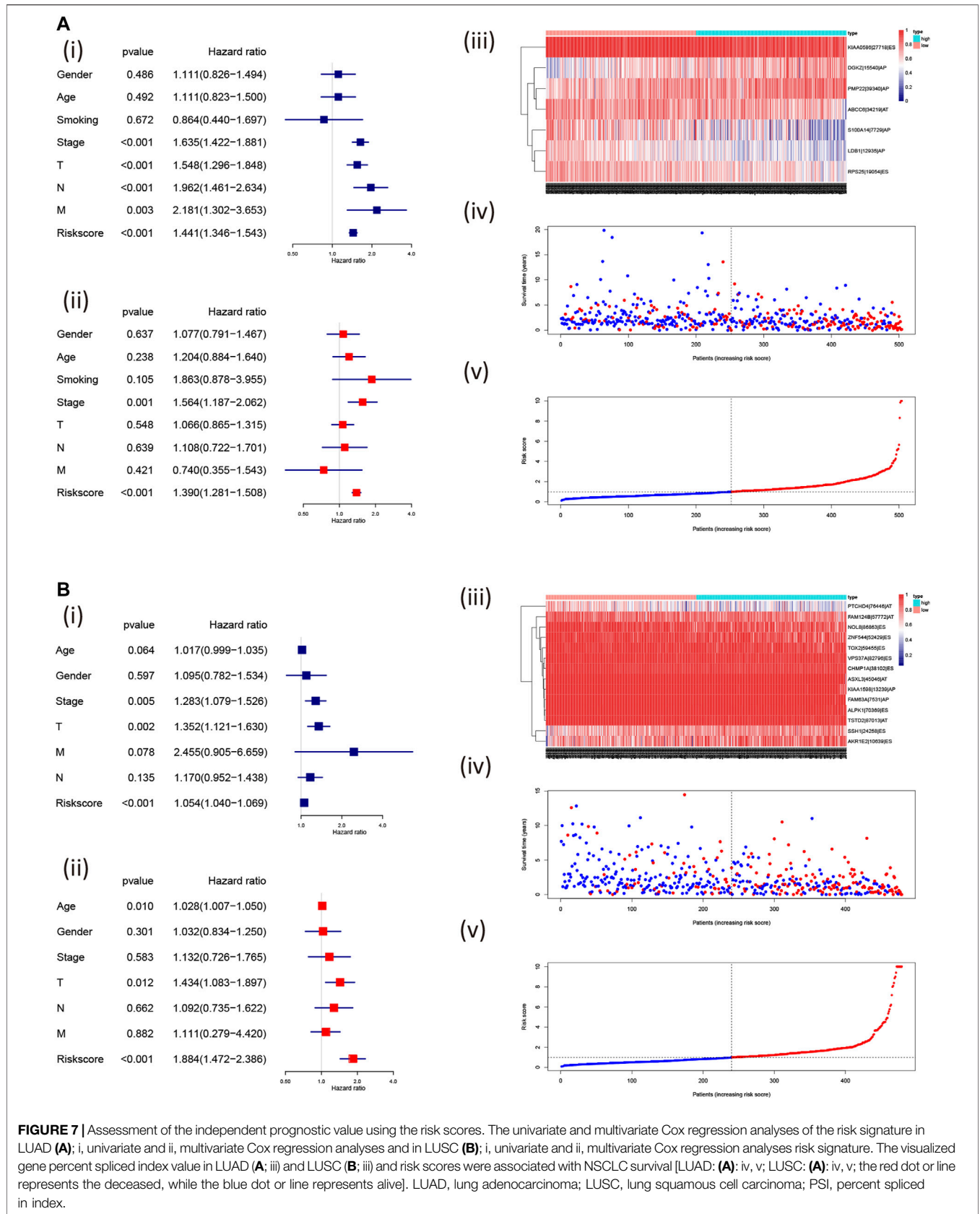
After that, we associated these prognostic signatures with NSCLC clinicopathologies and found that the LUAD risk signature was associated with the gender of patients and tumor T, N, and TNM stages (adj $p < 0.05$; **Figures 8A,C, Table 3**), although there was no association occurred between the LUSC risk signature and clinical features (**Table 3**). Furthermore, male ($n = 224$), TNM stage III-IV ($n = 105$), N stage 1-3 ($n = 167$), and T stage 3-4 ($n = 65$) of LUAD patients in had significantly higher risk scores than female ($n = 262$), TNM stage I-II ($n = 381$), N stage 0 ($n = 319$), and T stage 1-2 ($n = 421$; all adj $p < 0.05$; **Figure 8, Table 3**). Older age ($n = 271$) and M1 stage ($n = 26$) of LUAD patients also had the higher risk scores (all adj $p > 0.05$; **Figure 8B, Table 3**). Notably, male ($n = 353$), TNM stage III-IV ($n = 88$), N stage 1-3 ($n = 169$), M1 stage ($n = 7$), older age ($n = 313$), and T stage 3-4 ($n = 89$) of LUSC patients also had higher risk scores than those of the corresponding subgroups, but the differences did not appear statistically significant (**Figure 8D**).

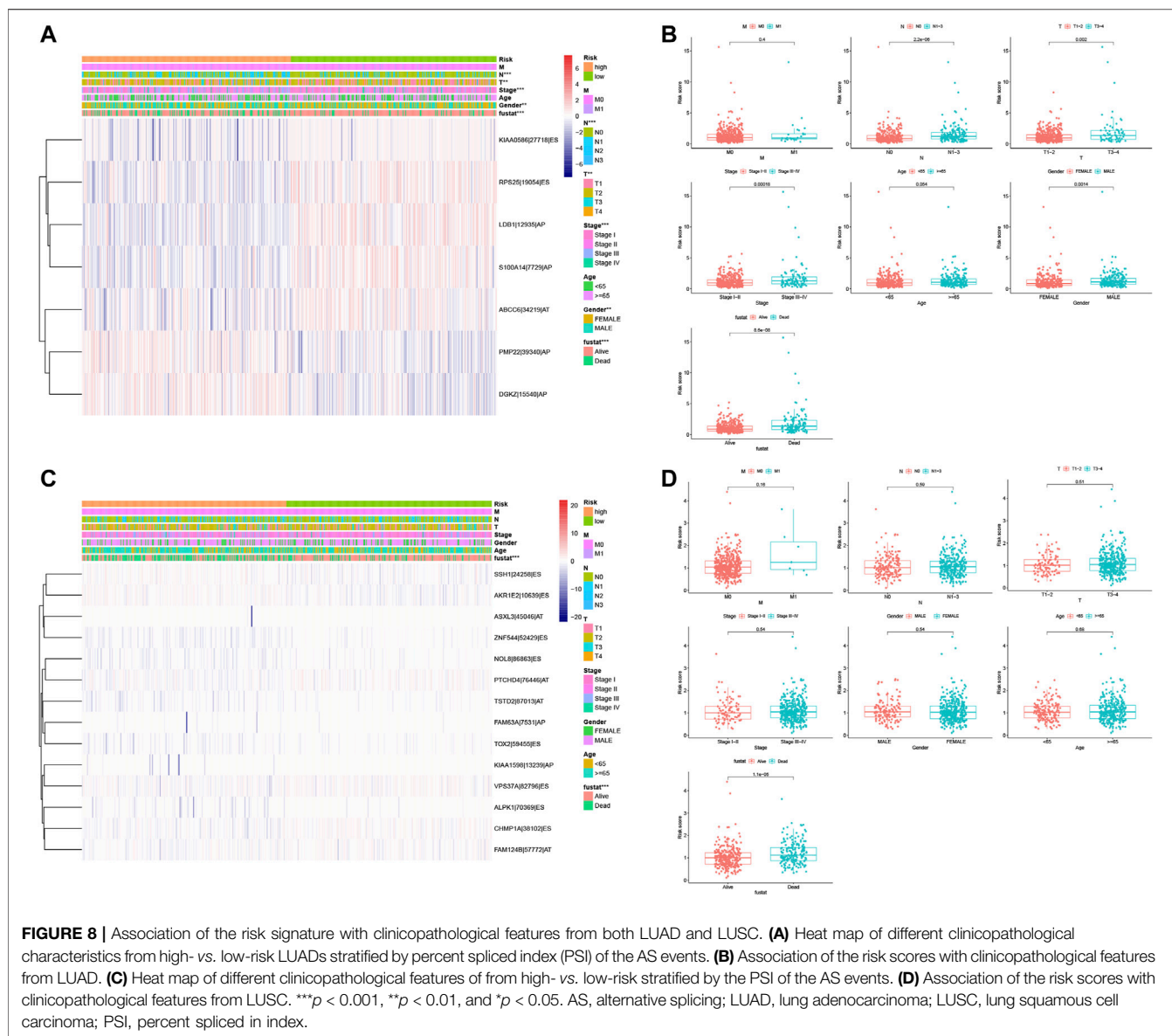
Usefulness of the Predictive Nomogram in Non-Small Cell Lung Cancer

So far, we showed the risk signature and TNM stage as independent prognostic predictors in LUAD and the risk signature and T stage as independent prognostic predictors in LUSC. We thus, constructed the nomogram using these parameters to assess and apply this risk model for NSCLC (**Figures 9A,C**) and verified the calibration curves of the nomogram (**Figures 9B,D**). We were able to use the numerous values of this risk model to predict the 1-, 3-, and 5-year survival of NSCLC patients.

DISCUSSION

In the current study, we analyzed the aberrant expression of 13 m6A regulatory genes and related GAS events to construct a risk gene signature to predict the overall survival of NSCLC patients. We found that a number of them were highly expressed in LUAD or LUSC tissues vs. their normal ones, which could be used to predict the survival of patients. Furthermore, we found 43,948





although the hazard ratios suggest that their prediction of NSCLC survival might be marginal. A previous study from Sun et al. (2020) showed the usefulness of the m6A regulators as the risk signature in LUAD (AUC, 0.65–0.82). Our results also suggested that m6A regulators play an important role in NSCLC development. Other previous studies (Lee et al., 2010; Gruber et al., 2016; Li et al., 2020) reported that HNRNPC was an RNA-binding protein (the “reader”), which could regulate RNA splicing, 3-terminal processing, and translation (Gruber et al., 2016; Lee et al., 2010; Li et al., 2020). HNRNPC overexpression was observed in a variety of human cancers, including lung cancer (Park et al., 2012) HNRNPC, as a protein-coding gene, could also interact with KHSRP to activate the IFN- α -JAK-p-STAT1 signaling pathway and promoted NSCLC cell proliferation, migration, and invasion (Yan et al., 2019). It can also regulate

the GAS as an “m6A switcher” (Alarcón et al., 2015; Dai et al., 2018; Li et al., 2017; Wang et al., 2020). Furthermore, RBM15, as a “writer,” can bind to METTL3 and WTAP and direct them to specific RNA sites for m6A modification (Wang et al., 2020), although it does not possess any catalytic functions (Chen et al., 2019). RBM15 was also shown to interact with the METTL3 complex and depletion of these adapters could also reduce the m6A level (Pendleton et al., 2017). Further investigation of RBM15 and GAS events revealed that RBM15 was able to bind to specific intron regions to recruit the splicing factor SF3B1AS (Zhang et al., 2015). In addition, *METTL3*, containing highly conserved sequences, is the most important component of the m6A methyltransferase complex and was shown to be an S-adenosyl methionine (SAM)-binding protein and catalyze m6A modification (Wang et al., 2016).

TABLE 3 | Clinicopathological features from LUAD and LUSC subgroups stratified by the AS events signature.

	LUAD	Adj-p	LUSC	Adj-p
Gender		*		8.96E-01
Male	224		353	
Female	262		126	
Age		1.38E-01		9.50E-01
<65	215		166	
≥65	271		313	
TNM stage		***		5.04E-01
I	263		235	
II	118		156	
III	80		81	
IV	25		7	
Tumor (T)		**		7.25E-01
T1	165		110	
T2	256		280	
T3	44		68	
T4	21		21	
Lymph node (N)		***		6.30E-01
N0	319		310	
N1	91		125	
N2	70		39	
N3	6		5	
Metastasis (M)		5.17E-01		9.96E-01
M0	460		472	
M1	26		7	
Status		***		**
Dead	120		293	
Alive	366		186	
Total case	486		479	

LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; TNM: tumor-node metastasis; *** represents $p < 0.001$; ** represents $p < 0.01$; * represents $p < 0.05$.

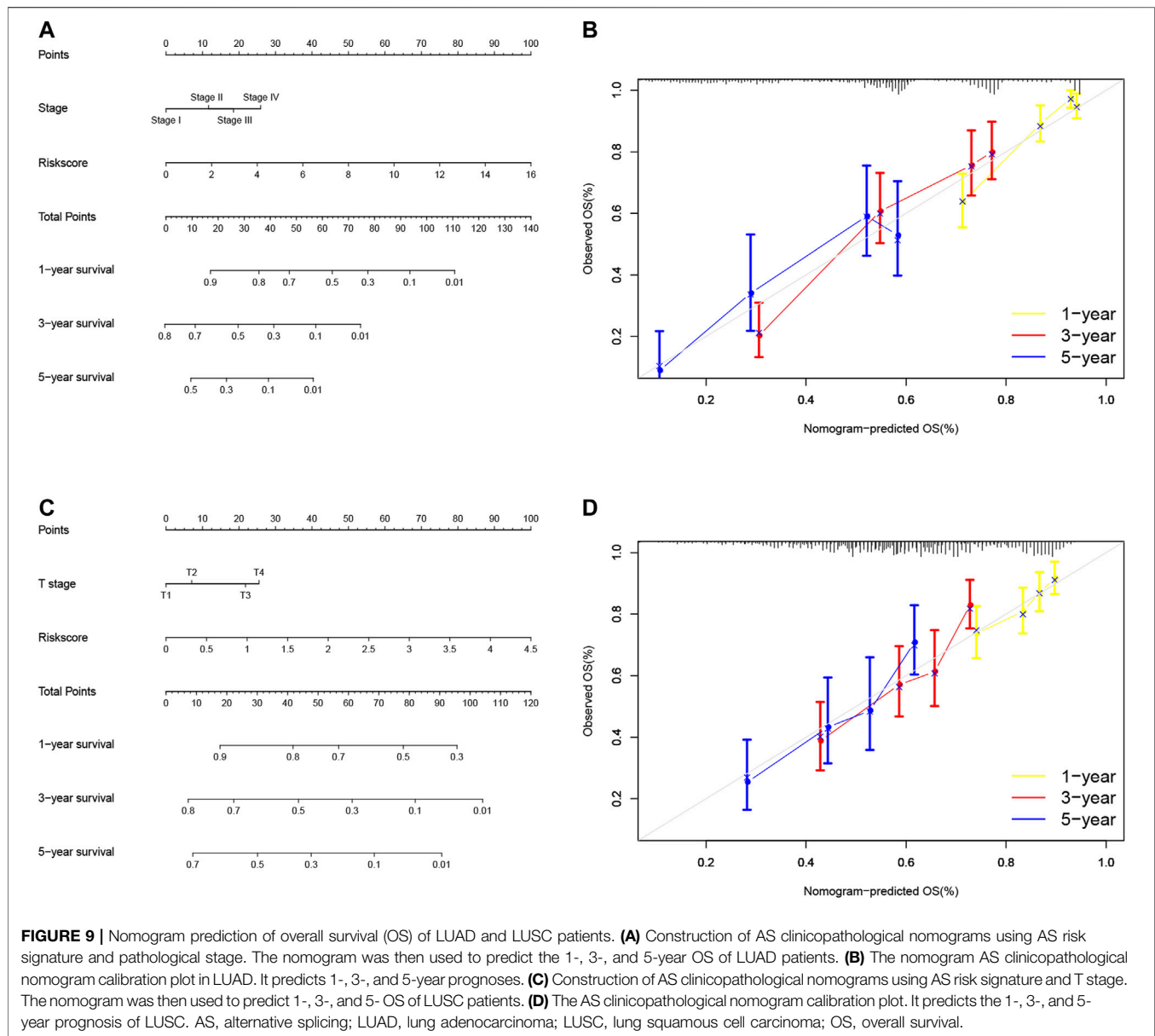
METL3 expression was high in LUAD and promoted the translation of the epidermal growth factor receptor (EGFR) mRNA and hippo pathway effector TAZ mRNA in lung cancer cells, for induction of cell growth, survival, and invasion (Lin et al., 2016). METTL3 was also shown to interact with GAS events of the skipped exons and alternative first exon (Alarcón et al., 2015), and METTL3 dysregulation was reported to indeed affect GAS events (Katz et al., 2015; Liu et al., 2014). METTL3 silence significantly affected gene expression and alternative splicing patterns, leading to modulation of the p53 pathway and cell apoptosis (Dominissini et al., 2012). Taken altogether, these three m6A RNA methylation regulatory genes were important in the regulation of GAS events in NSCLC.

Indeed, AS events is an important mRNA modification process and produce a large number of mRNA and protein isoforms with different regulatory functions (Bonnal et al., 2020; Liu et al., 2020). The prognostic value of the AS events in NSCLC has well been documented, for example, Zhao et al. (2020) built a predictive model of aberrant AS events and predicated NSCLC prognosis. Indeed, alternation in splicing factor expression could alter many AS events in NSCLC (Coomer et al., 2019). For instance, QKI was shown to one of the most downregulated splicing factors in NSCLC, while QKI-5 was able to competitively bind to NUMB with SF1 protein to induce the NUMB exon 11 skip and, therefore, inhibited the Notch signaling (Zong et al., 2014; de Miguel et al., 2016). In lung

cancer, QKI expression was significantly reduced, increasing in the abnormal splicing of num exon 11 to, in turn, activate the Notch signaling pathway and tumor cell proliferation (Zong et al., 2014; de Miguel et al., 2016). The AS events also influenced p53 expression in NSCLC and MDM2-B, an AS product of MDM2, was able to promote p53-independent cell growth and inhibition of apoptosis (Coomer et al., 2019). In this regard, the AS events are important in NSCLC development and progression (Bonnal et al., 2020; Sciarillo et al., 2020).

Furthermore, the weighted gene co-expression network analysis (WGCNA) is a widely used data mining method, especially used for studying the biological networks based on pairwise correlations between variables (Langfelder and Horvath, 2008). In the current study, we used WGCNA to select the AS events that are highly correlated with the NSCLC survival-related m6A RNA regulators. After that, we performed the GO and KEGG pathways enrichment analysis to identify genes of m6A-related AS events to significantly participate in gene pathways that play an important role in NSCLC tumorigenesis, progression, drug sensitivity, and metastasis. Indeed, some of the abnormal AS events were associated with drug sensitivity and resistance of NSCLC (Motegi et al., 2019; Pilié et al., 2019) as well as cell adhesion molecule binding process (Song et al., 2013; Hintermann and Christen, 2019). Our KEGG analysis showed that the genes in the m6A-related AS events significantly participated in the GPCR signaling in LUAD, and the latter is mediated by three major G protein subclasses and each subclass also has multiple proteins that are products due to the AS events (Kang et al., 2015; Kallifatidis et al., 2020). Similarly, we found the m6A-related AS events in the TGF- β signaling in LUSC. The TGF- β signaling pathway was frequently downregulated in human cancers (Syed, 2016), whereas this pathway activation could also promote tumorigenesis, metastasis, and chemoresistance (Colak and Ten Dijke, 2017; Zi, 2019). In this regard, genes of the m6A-related AS event-led activation of the TGF- β signaling could promote LUSC tumorigenesis. However, further study is needed to confirm this speculation.

In addition, in our current study, we constructed the risk signature using these altered genes in m6A and AS events to associate with NSCLC prognosis, and the AUC of the ROC curves showed the sensitivity and specificity of LUAD and LUSC, respectively, which are better than other recent studies (Zhao et al., 2020) (Liu et al., 2020). In these risk signatures, a previous study showed that S100A14 overexpression was able to promote LUAD cell migration and invasion (Ding et al., 2018). In all recent studies of the AS events in NSCLC, Li et al. (2017) were the first to construct an AS risk signature for the prediction of NSCLC prognosis, while Zhao et al. (2020) constructed the AS risk signature stratified by gender of patients. Liu et al. (2020) formed an AS signature for LUSC. Our current study also explored abnormal expression of the splicing factors in NSCLC as well as the C-index (**Supplementary Table S7**). However, our current study does have some limitations, for example, the AS events database is relatively simple and lacks all other relevant datasets for us to verify our data. In addition, the relationship of m6A regulators with the AS events and the mechanism by which they play a role in NSCLC development



remain; thus, more studies are needed to clarify the true biological role of the AS events in NSCLC tumorigenesis.

CONCLUSIONS

Our current study assessed the role of m6A-related AS events in NSCLC as a signature in the prediction of NSCLC prognosis. The current study revealed the regulation of AS events by some key m6A regulators may play an important role in NSCLC development and progression. This study might provide a novel insight into the mechanism of NSCLC tumorigenesis, which may lead to novel strategies in future control of NSCLC.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the TCGA-LUSC and TCGA-LUAD database (<https://portal.gdc.cancer.gov/>).

AUTHOR CONTRIBUTIONS

XW conceived and designed the work. ZZ and QC carried out software coding and data analysis. ZZ and PZ formatted the tables and figures. ZZ, BH, WP, and GT wrote the manuscript. FY and LW critically reviewed the codes and the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported in part by grants from the National Natural Science Foundation of China (#81672308, #81101767, and #81972195), the Hunan Provincial Key Area R&D Program (#2019SK2253), and the Scientific Research Program of Hunan Provincial Health Commission (#20201047). Independent Exploration and Innovation Project of Graduate Student of Central South University (#2021zzts0385).

ACKNOWLEDGMENTS

The authors thank Yongguang Tao for their comments and suggestions in this study. Thank you for the revision by Medjaden Bioscience Limited (Hong Kong, China).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.657087/full#supplementary-material>

SUPPLEMENTARY FIGURE S1 | The upstate diagrams of LUAD and LUSC. **(A)** The upstate diagram of alternative splicing events in LUAD. **(B)** The upstate diagram of alternative splicing events in LUSC. LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

REFERENCES

- Alarcón, C. R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S. F. (2015). HNRNPA2B1 Is a Mediator of m6A-Dependent Nuclear RNA Processing Events. *Cell* 162 (6), 1299–1308. doi:10.1016/j.cell.2015.08.011
- Amado, F. M., Barros, A., Azevedo, A. L., Vitorino, R., and Ferreira, R. (2014). An Integrated Perspective and Functional Impact of the Mitochondrial Acetylome. *Expert Rev. Proteomics* 11 (3), 383–394. doi:10.1586/14789450.2014.899470
- Balata, H., Fong, K. M., Hendriks, L. E., Lam, S., Ostroff, J. S., Peled, N., et al. (2019). Prevention and Early Detection for NSCLC: Advances in Thoracic Oncology 2018. *J. Thorac. Oncol.* 14 (9), 1513–1527. doi:10.1016/j.jtho.2019.06.011
- Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA Cloning of the AdoMet-Binding Subunit of the Human mRNA (N6-Adenosine)-Methyltransferase. *Rna* 3 (11), 1233–1247.
- Bonnal, S. C., López-Oreja, I., and Valcárcel, J. (2020). Roles and Mechanisms of Alternative Splicing in Cancer - Implications for Care. *Nat. Rev. Clin. Oncol.* 17 (8), 457–474. doi:10.1038/s41571-020-0350-x
- Chen, X.-Y., Zhang, J., and Zhu, J.-S. (2019). The Role of m6A RNA Methylation in Human Cancer. *Mol. Cancer* 18 (1), 103. doi:10.1186/s12943-019-1033-z
- Cherry, S., and Lynch, K. W. (2020). Alternative Splicing and Cancer: Insights, Opportunities, and Challenges from an Expanding View of the Transcriptome. *Genes Dev.* 34 (15-16), 1005–1016. doi:10.1101/gad.338962.120
- Colak, S., and Ten Dijke, P. (2017). Targeting TGF- β Signaling in Cancer. *Trends Cancer* 3 (1), 56–71. doi:10.1016/j.trecan.2016.11.008
- Coomer, A. O., Black, F., Greystoke, A., Munkley, J., and Elliott, D. J. (2019). Alternative Splicing in Lung Cancer. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1862 (11-12), 194388. doi:10.1016/j.bbagr.2019.05.006
- Dai, D., Wang, H., Zhu, L., Jin, H., and Wang, X. (2018). N6-methyladenosine Links RNA Metabolism to Cancer Progression. *Cell Death Dis.* 9 (2), 124. doi:10.1038/s41419-017-0129-x
- David, C. J., and Manley, J. L. (2008). The Search for Alternative Splicing Regulators: New Approaches Offer a Path to a Splicing Code. *Genes Dev.* 22 (3), 279–285. doi:10.1101/gad.1643108
- de Martel, C., Georges, D., Bray, F., Ferlay, J., and Clifford, G. M. (2020). Global burden of Cancer Attributable to Infections in 2018: a Worldwide Incidence Analysis. *Lancet Glob. Health* 8 (2), e180–e190. doi:10.1016/s2214-109x(19)30488-7
- de Miguel, F. J., Pajares, M. J., Martínez-Terroba, E., Ajona, D., Morales, X., Sharma, R. D., et al. (2016). A Large-Scale Analysis of Alternative Splicing Reveals a Key Role of QKI in Lung Cancer. *Mol. Oncol.* 10 (9), 1437–1449. doi:10.1016/j.molonc.2016.08.001
- Deng, X., Su, R., Feng, X., Wei, M., and Chen, J. (2018). Role of N6-Methyladenosine Modification in Cancer. *Curr. Opin. Genet. Dev.* 48, 1–7. doi:10.1016/j.gde.2017.10.005
- Ding, F., Wang, D., Li, X.-K., Yang, L., Liu, H.-Y., Cui, W., et al. (2018). Overexpression of S100A14 Contributes to Malignant Progression and Predicts Poor Prognosis of Lung Adenocarcinoma. *Thorac. Cancer* 9 (7), 827–835. doi:10.1111/1759-7714.12654
- Dominianni, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the Human and Mouse m6A RNA Methylomes Revealed by m6A-Seq. *Nature* 485 (7397), 201–206. doi:10.1038/nature11112
- Frankiw, L., Baltimore, D., and Li, G. (2019). Alternative mRNA Splicing in Cancer Immunotherapy. *Nat. Rev. Immunol.* 19 (11), 675–687. doi:10.1038/s41577-019-0195-7
- Geng, Y., Guan, R., Hong, W., Huang, B., Liu, P., Guo, X., et al. (2020). Identification of m6A-Related Genes and m6A RNA Methylation Regulators in Pancreatic Cancer and Their Association with Survival. *Ann. Transl. Med.* 8 (6), 387. doi:10.21037/atm.2020.03.98
- Graedel, T. E. (2019). Material Flow Analysis from Origin to Evolution. *Environ. Sci. Technol.* 53 (21), 12188–12196. doi:10.1021/acs.est.9b03413
- Gruber, A. J., Schmidt, R., Gruber, A. R., Martin, G., Ghosh, S., Belmadani, M., et al. (2016). A Comprehensive Analysis of 3' End Sequencing Data Sets Reveals Novel Polyadenylation Signals and the Repressive Role of Heterogeneous Ribonucleoprotein C on Cleavage and Polyadenylation. *Genome Res.* 26 (8), 1145–1159. doi:10.1101/gr.202432.115
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of Cancer: the Next Generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013

- Hintermann, E., and Christen, U. (2019). The Many Roles of Cell Adhesion Molecules in Hepatic Fibrosis. *Cells* 8 (12), 1503. doi:10.3390/cells8121503
- Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J., Jr., Wu, Y.-L., et al. (2017). Lung Cancer: Current Therapies and New Targeted Treatments. *The Lancet* 389 (10066), 299–311. doi:10.1016/s0140-6736(16)30958-8
- Ji, P., Wang, X., Xie, N., and Li, Y. (2018). N6-Methyladenosine in RNA and DNA: An Epitranscriptomic and Epigenetic Player Implicated in Determination of Stem Cell Fate. *Stem Cell Int.* 2018, 1–18. doi:10.1155/2018/3256524
- Kallifatidis, G., Mamouni, K., and Lokeswar, B. L. (2020). The Role of β -Arrestins in Regulating Stem Cell Phenotypes in Normal and Tumorigenic Cells. *Ijms* 21 (23), 9310. doi:10.3390/ijms21239310
- Kang, Y., Zhou, X. H., Gao, X., He, Y., Liu, W., Ishchenko, A., et al. (2015). Crystal Structure of Rhodopsin Bound to Arrestin by Femtosecond X-ray Laser. *Nature* 523 (7562), 561–567. doi:10.1038/nature14656
- Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, B. D., et al. (2018). Nuclear m6A Reader YTHDC1 Regulates Alternative Polyadenylation and Splicing during Mouse Oocyte Development. *Plos Genet.* 14 (5), e1007412. doi:10.1371/journal.pgen.1007412
- Katz, Y., Wang, E. T., Silterra, J., Schwartz, S., Wong, B., Thorvaldsdóttir, H., et al. (2015). Quantitative Visualization of Alternative Exon Expression from RNA-Seq Data. *Bioinformatics* 31 (14), 2400–2402. doi:10.1093/bioinformatics/btv034
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R Package for Weighted Correlation Network Analysis. *BMC Bioinformatics* 9, 559. doi:10.1186/1471-2105-9-559
- Lee, E. K., Kim, H. H., Kuwano, Y., Abdelmohsen, K., Srikantan, S., Subaran, S. S., et al. (2010). hnRNP C Promotes APP Translation by Competing with FMRP for APP mRNA Recruitment to P Bodies. *Nat. Struct. Mol. Biol.* 17 (6), 732–739. doi:10.1038/nsmb.1815
- Li, F., Wang, H., Huang, H., Zhang, L., Wang, D., and Wan, Y. (2020). m6A RNA Methylation Regulators Participate in the Malignant Progression and Have Clinical Prognostic Value in Lung Adenocarcinoma. *Front. Genet.* 11, 994. doi:10.3389/fgene.2020.00994
- Li, S., Hu, Z., Zhao, Y., Huang, S., and He, X. (2019). Transcriptome-Wide Analysis Reveals the Landscape of Aberrant Alternative Splicing Events in Liver Cancer. *Hepatology* 69 (1), 359–375. doi:10.1002/hep.30158
- Li, Y., Sun, N., Lu, Z., Sun, S., Huang, J., Chen, Z., et al. (2017). Prognostic Alternative mRNA Splicing Signature in Non-small Cell Lung Cancer. *Cancer Lett.* 393, 40–51. doi:10.1016/j.canlet.2017.02.016
- Lin, K.-T., and Krainer, A. R. (2019). PSI-sigma: a Comprehensive Splicing-Detection Method for Short-Read and Long-Read RNA-Seq Analysis. *Bioinformatics* 35 (23), 5048–5054. doi:10.1093/bioinformatics/btz438
- Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R. I. (2016). The M 6 A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol. Cell.* 62 (3), 335–345. doi:10.1016/j.molcel.2016.03.021
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 Complex Mediates Mammalian Nuclear RNA N6-Adenosine Methylation. *Nat. Chem. Biol.* 10 (2), 93–95. doi:10.1038/nchembio.1432
- Liu, Y., Jia, W., Li, J., Zhu, H., and Yu, J. (2020). Identification of Survival-Associated Alternative Splicing Signatures in Lung Squamous Cell Carcinoma. *Front. Oncol.* 10, 587343. doi:10.3389/fonc.2020.587343
- Maconachie, R., Mercer, T., Navani, N., and McVeigh, G. (2019). Lung Cancer: Diagnosis and Management: Summary of Updated NICE Guidance. *Bmj* 364, I1049. doi:10.1136/bmj.I1049
- Meng, L.-B., Shan, M.-J., Qiu, Y., Qi, R., Yu, Z.-M., Guo, P., et al. (2019). TPM2 as a Potential Predictive Biomarker for Atherosclerosis. *Aging* 11 (17), 6960–6982. doi:10.18632/aging.102231
- Motegi, A., Masutani, M., Yoshioka, K.-i., and Bessho, T. (2019). Aberrations in DNA Repair Pathways in Cancer and Therapeutic Significances. *Semin. Cancer Biol.* 58, 29–46. doi:10.1016/j.semcancer.2019.02.005
- Niemira, M., Collin, F., Szalkowska, A., Bielska, A., Chwialkowska, K., Reszec, J., et al. (2019). Molecular Signature of Subtypes of Non-small-cell Lung Cancer by Large-Scale Transcriptional Profiling: Identification of Key Modules and Genes by Weighted Gene Co-expression Network Analysis (WGCNA). *Cancers* 12 (1), 37. doi:10.3390/cancers12010037
- Papasaiikas, P., and Valcárcel, J. (2016). The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends Biochem. Sci.* 41 (1), 33–45. doi:10.1016/j.tibs.2015.11.003
- Park, S. H., Goo, J. M., and Jo, C.-H. (2004). Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists. *Korean J. Radiol.* 5 (1), 11–18. doi:10.3348/kjr.2004.5.1.11
- Park, Y. M., Hwang, S. J., Masuda, K., Choi, K.-M., Jeong, M.-R., Nam, D.-H., et al. (2012). Heterogeneous Nuclear Ribonucleoprotein C1/C2 Controls the Metastatic Potential of Glioblastoma by Regulating PDCC4. *Mol. Cell Biol.* 32 (20), 4237–4244. doi:10.1128/mcb.00443-12
- Paschalis, A., Sharp, A., Welti, J. C., Neeb, A., Raj, G. V., Luo, J., et al. (2018). Alternative Splicing in Prostate Cancer. *Nat. Rev. Clin. Oncol.* 15 (11), 663–675. doi:10.1038/s41571-018-0085-0
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA M 6 A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell* 169 (5), 824–835. e814. doi:10.1016/j.cell.2017.05.003
- Pilié, P. G., Tang, C., Mills, G. B., and Yap, T. A. (2019). State-of-the-art Strategies for Targeting the DNA Damage Response in Cancer. *Nat. Rev. Clin. Oncol.* 16 (2), 81–104. doi:10.1038/s41571-018-0114-z
- Pio, R., and Montuenga, L. M. (2009). Alternative Splicing in Lung Cancer. *J. Thorac. Oncol.* 4 (6), 674–678. doi:10.1097/JTO.0b013e3181a520dc
- Planchard, D., Popat, S., Kerr, K., Novello, S., Smit, E. F., Faivre-Finn, C., et al. (2018). Metastatic Non-small Cell Lung Cancer: ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-Up. *Ann. Oncol.* 29 (Suppl. 4), iv192–iv237. doi:10.1093/annonc/mdy275
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Rizvi, A. A., Karaesmen, E., Morgan, M., Preus, L., Wang, J., Sovic, M., et al. (2019). Gwasurvivr: an R Package for Genome-wide Survival Analysis. *Bioinformatics* 35 (11), 1968–1970. doi:10.1093/bioinformatics/bty920
- Sciarrillo, R., Wojtuszkiewicz, A., Assaraf, Y. G., Jansen, G., Kaspers, G. J. L., Giovannetti, E., et al. (2020). The Role of Alternative Splicing in Cancer: From Oncogenesis to Drug Resistance. *Drug Resist. Updates* 53, 100728. doi:10.1016/j.drug.2020.100728
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303
- Sholl, L. (2017). Molecular Diagnostics of Lung Cancer in the Clinic. *Transl. Lung Cancer Res.* 6 (5), 560–569. doi:10.21037/tlcr.2017.08.03
- Soh, J. Y., Jung, S.-H., Cha, W. C., Kang, M., Chang, D. K., Jung, J., et al. (2019). Variability in Doctors' Usage Paths of Mobile Electronic Health Records Across Specialties: Comprehensive Analysis of Log Data. *JMIR Mhealth Uhealth* 7 (1), e12041. doi:10.2196/12041
- Song, Y., Zhu, Z., An, Y., Zhang, W., Zhang, H., Liu, D., et al. (2013). Selection of DNA Aptamers against Epithelial Cell Adhesion Molecule for Cancer Cell Imaging and Circulating Tumor Cell Capture. *Anal. Chem.* 85 (8), 4141–4149. doi:10.1021/ac400366b
- Sun, L., Liu, W.-K., Du, X.-W., Liu, X.-L., Li, G., Yao, Y., et al. (2020). Large-scale Transcriptome Analysis Identified RNA Methylation Regulators as Novel Prognostic Signatures for Lung Adenocarcinoma. *Ann. Transl. Med.* 8 (12), 751. doi:10.21037/atm-20-3744
- Sun, T., Wu, R., and Ming, L. (2019). The Role of m6A RNA Methylation in Cancer. *Biomed. Pharmacother.* 112, 108613. doi:10.1016/j.biopha.2019.108613
- Sun, Y., and Ma, L. (2019). New Insights into Long Non-coding RNA MALAT1 in Cancer and Metastasis. *Cancers* 11 (2), 216. doi:10.3390/cancers11020216
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660
- Syed, V. (2016). TGF- β Signaling in Cancer. *J. Cel. Biochem.* 117 (6), 1279–1287. doi:10.1002/jcb.25496
- Tang, Z., Shen, Y., Zhang, X., and Yi, N. (2017). The Spike-And-Slab Lasso Cox Model for Survival Prediction and Associated Genes Detection. *Bioinformatics* 33 (18), 2799–2807. doi:10.1093/bioinformatics/btx300
- Tanoue, L. T., Tanner, N. T., Gould, M. K., and Silvestri, G. A. (2015). Lung Cancer Screening. *Am. J. Respir. Crit. Care Med.* 191 (1), 19–33. doi:10.1164/rccm.201410-1777CI
- Urbanski, L. M., Leclair, N., and Anczuków, O. (2018). Alternative-splicing Defects in Cancer: Splicing Regulators and Their Downstream Targets, Guiding the

- Way to Novel Cancer Therapeutics. *WIREs RNA* 9 (4), e1476. doi:10.1002/wrna.1476
- Verbakel, J. Y., Steyerberg, E. W., Uno, H., De Cock, B., Wynants, L., Collins, G. S., et al. (2020). ROC Curves for Clinical Prediction Models Part 1. ROC Plots Showed No Added Value Above the AUC When Evaluating the Performance of Clinical Prediction Models. *J. Clin. Epidemiol.* 126, 207–216. doi:10.1016/j.jclinepi.2020.01.028
- Wan, Q., Tang, J., Han, Y., and Wang, D. (2018). Co-expression Modules Construction by WGCNA and Identify Potential Prognostic Markers of Uveal Melanoma. *Exp. Eye Res.* 166, 13–20. doi:10.1016/j.exer.2017.10.007
- Wang, P., Doxtader, K. A., and Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Mol. Cell.* 63 (2), 306–317. doi:10.1016/j.molcel.2016.05.041
- Wang, T., Kong, S., Tao, M., and Ju, S. (2020). The Potential Role of RNA N6-Methyladenosine in Cancer Progression. *Mol. Cancer* 19 (1), 88. doi:10.1186/s12943-020-01204-7
- Wang, Y., Li, Y., Toth, J. I., Petroski, M. D., Zhang, Z., and Zhao, J. C. (2014). N6-methyladenosine Modification Destabilizes Developmental Regulators in Embryonic Stem Cells. *Nat. Cell Biol.* 16 (2), 191–198. doi:10.1038/ncb2902
- Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., et al. (2016). Nuclear M6A Reader YTHDC1 Regulates mRNA Splicing. *Mol. Cell.* 61 (4), 507–519. doi:10.1016/j.molcel.2016.01.012
- Xie, H., and Xie, C. (2019). A Six-Gene Signature Predicts Survival of Adenocarcinoma Type of Non-small-cell Lung Cancer Patients: A Comprehensive Study Based on Integrated Analysis and Weighted Gene Coexpression Network. *Biomed. Res. Int.* 2019, 1–16. doi:10.1155/2019/4250613
- Yan, M., Sun, L., Li, J., Yu, H., Lin, H., Yu, T., et al. (2019). RNA-binding Protein KHSRP Promotes Tumor Growth and Metastasis in Non-small Cell Lung Cancer. *J. Exp. Clin. Cancer Res.* 38 (1), 478. doi:10.1186/s13046-019-1479-2
- Yang, Q., Zhao, J., Zhang, W., Chen, D., and Wang, Y. (2019). Aberrant Alternative Splicing in Breast Cancer. *J. Mol. Cell Biol.* 11 (10), 920–929. doi:10.1093/jmcb/mjz033
- Yoshimi, A., Lin, K. T., Wiseman, D. H., Rahman, M. A., Pastore, A., Wang, B., et al. (2019). Coordinated Alterations in RNA Splicing and Epigenetic Regulation Drive Leukaemogenesis. *Nature* 574 (7777), 273–277. doi:10.1038/s41586-019-1618-0
- Zhang, L., Tran, N.-T., Su, H., Wang, R., Lu, Y., Tang, H., et al. (2015). Cross-talk between PRMT1-Mediated Methylation and Ubiquitylation on RBM15 Controls RNA Splicing. *Elife* 4, e07938. doi:10.7554/eLife.07938
- Zhang, S., Hu, Z., Lan, Y., Long, J., Wang, Y., Chen, X., et al. (2020a). Prognostic Significance of Survival-Associated Alternative Splicing Events in Gastric Cancer. *Aging* 12 (21), 21923–21941. doi:10.18632/aging.104013
- Zhang, S., Tong, Y. X., Zhang, X. H., Zhang, Y. J., Xu, X. S., Xiao, A. T., et al. (2019). A Novel and Validated Nomogram to Predict Overall Survival for Gastric Neuroendocrine Neoplasms. *J. Cancer* 10 (24), 5944–5954. doi:10.7150/jca.35785
- Zhang, Y., Geng, X., Li, Q., Xu, J., Tan, Y., Xiao, M., et al. (2020c). m6A Modification in RNA: Biogenesis, Functions and Roles in Gliomas. *J. Exp. Clin. Cancer Res.* 39 (1), 192. doi:10.1186/s13046-020-01706-8
- Zhang, Y., Liu, X., Liu, L., Li, J., Hu, Q., and Sun, R. (2020b). Expression and Prognostic Significance of m6A-Related Genes in Lung Adenocarcinoma. *Med. Sci. Monit.* 26, e919644. doi:10.12659/msm.919644
- Zhang, Y., Wang, W., Xu, X., Li, Y., Zhang, H., Li, J., et al. (2021). Impact of Radiotherapy Pattern on the Prognosis of Stage IV Lung Adenocarcinomas Harboring EGFR Mutations. *Cmar* 13, 3293–3301. doi:10.2147/cmar.s299563
- Zhao, D., Zhang, C., Jiang, M., Wang, Y., Liang, Y., Wang, L., et al. (2020). Survival-associated Alternative Splicing Signatures in Non-small Cell Lung Cancer. *Aging* 12 (7), 5878–5893. doi:10.18632/aging.102983
- Zi, Z. (2019). Molecular Engineering of the TGF- β Signaling Pathway. *J. Mol. Biol.* 431 (15), 2644–2654. doi:10.1016/j.jmb.2019.05.022
- Zong, F.-Y., Fu, X., Wei, W.-J., Luo, Y.-G., Heiner, M., Cao, L.-J., et al. (2014). The RNA-Binding Protein QKI Suppresses Cancer-Associated Aberrant Splicing. *Plos Genet.* 10 (4), e1004289. doi:10.1371/journal.pgen.1004289

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhao, Cai, Zhang, He, Peng, Tu, Peng, Wang, Yu and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.