Identification of N7-methylguanosine related signature for prognosis and immunotherapy efficacy prediction in lung adenocarcinoma

Zhouhua Li†, Wenjun Wang†, Juan Wu and Xiaoqun Ye*

Department of Respiratory Diseases, The Second Affiliated Hospital of Nanchang University, Nanchang, China

Background: Lung adenocarcinoma (LUAD) is one of the most frequent causes of tumor-related mortality worldwide. Recently, the role of N7-methylguanosine (m7G) in tumors has begun to receive attention, but no investigation on the impact of m7G on LUAD. This study aims to elucidate the significance of m7G on the prognosis and immunotherapy in LUAD.

Methods: Consensus clustering was employed to determine the molecular subtype according to m7G-related regulators extracted from The Cancer Genome Atlas (TCGA) database. Survival, clinicopathological features and tumor mutational burden (TMB) analysis were applied to research molecular characteristics of each subtype. Subsequently, “limma” package was used to screen differentially expressed genes (DEGs) between subtypes. In the TCGA train cohort (n = 245), a prognostic signature was established by univariate Cox regression, lasso regression and multivariate Cox regression analysis according to DEGs and survival analysis was employed to assess the prognosis. Then the prognostic value of the signature was verified by TCGA test cohort (n = 245), TCGA entire cohort (n = 490) and GSE31210 cohort (n = 226). Moreover, the association among immune infiltration, clinical features and the signature was investigated. The immune checkpoints, TMB and tumor immune dysfunction and exclusion (TIDE) were applied to predict the immunotherapy response.

Results: Two novel molecular subtypes (C1 and C2) of LUAD were identified. Compared to C2 subtype, C1 subtype had poorer prognosis and higher TMB. Subsequently, the signature (called the “m7G score”) was constructed according to four key genes (E2F7, FAM83A, PITX3, and HOXA13). The distribution of m7G score were significantly different between two molecular subtypes. The patients with lower m7G score had better prognosis in TCGA train cohort and three verification cohort. The m7G score was intensively
related to immune infiltration. Compared with the lower score, the higher m\textsuperscript{7}G score was related to remarkable upregulation of the PD-1 and PD-L1, the higher TMB and the lower TIDE score.

**Conclusion:** This study established a m\textsuperscript{7}G-related signature for predicting prognosis and immunotherapy in LUAD, which may contribute to the development of new therapeutic strategies for LUAD.

**KEYWORDS**
N7-methylguanosine, lung adenocarcinoma, molecular subtype, prognosis, immunotherapy efficacy

---

**Introduction**

Lung adenocarcinoma (LUAD) accounts for the largest proportion in non-small cell lung cancer (NSCLC) (1). Since patients with LUAD suffer from advanced disease or have distant metastasis when first diagnosed, they have a poor prognosis, and the overall 5-year survival rate is still below 20% (2, 3). Impressively, immune checkpoint blockade (ICB) has become a promising therapy strategy for NSCLC (4). However, some patients have a low response rate to ICB treatment, or even drug resistance, thus resulting in disease relapse or dead cases (5, 6). Therefore, it is essential to identify a novel biomarker in LUAD, in order to improve the outcomes of patients and formulate personalized treatment strategies.

Increasing evidence indicates that the initiation and progression of lung cancer depends not only on genetic variation, but also on epigenetic dysregulation (7, 8). As an important part of epigenetic modification, RNA modification is involved in regulating many physiological processes and disease occurrence (9). Besides, dynamic regulation and disruption of these RNA modifications are also related to the tumorigenesis, maintenance and progression of lung cancer (10, 11). Among numerous RNA dynamic modifications, N6-Methyladenosine (m\textsuperscript{6}A), 5-Methylcytosine (m\textsuperscript{5}C), and N7-methylguanosine (m\textsuperscript{7}G) are extremely common (12). Importantly, m\textsuperscript{7}G is the most prevalent modifications of RNA caps (13), which occurs in various RNAs of eukaryotes (14). m\textsuperscript{7}G modification has a significant impact on RNA metabolism, processing and function (15). Nevertheless, the exploration of m\textsuperscript{7}G-related regulators on tumors have only recently begun to receive attention owing to technological limitations. Mis-regulated m\textsuperscript{7}G modification could disturb the translation of many oncogenic transcripts involved in RPTOR/ULK1/autophagy pathway, which contributed to esophageal squamous cell carcinoma oncogenesis (16). EIF4E is regarded as one of m7G-related regulators, whose phosphorylation could increase the translations of oncogene mRNAs to promote prostate cancer tumorigenesis (17). Moreover, one study demonstrated that METTL1 and WDR4 were upregulated in lung cancer samples and vital for the progression (18). Besides, RNA dynamic modification could influence the response function and maturation of tumor immune cells (19). So far, the overall impact of m\textsuperscript{7}G-related regulators on the immunotherapeutic response in LUAD and its relationship with patient prognosis and treatment are still unclear.

With the advances in high-throughput sequencing technique, research on tumor genes is more in-depth, which can help to classify tumors to some content. There are many signatures that assess the prognosis of LUAD according to various subtypes (20–22). However, these signatures are still far from guiding precise treatment, which urgently requires a reliable signature. Here, two molecular subtypes of LUAD were constructed according to the gene expression of m7G-related regulators. We further evaluated the relation between survival, clinical characteristics, immune infiltration and molecular subtype. Then, a novel m\textsuperscript{7}G score was established to quantify the m\textsuperscript{7}G modification patterns, which was proven to be an independent predictor of LUAD prognosis. Moreover, the prognostic signature effectiveness was validated by the internal and external cohort (GSE31210). Furthermore, we elucidated whether this signature could provide reference for clinical immunotherapy and chemotherapy. In conclusion, this study not only provides a novel understanding of molecular subtype by m\textsuperscript{7}G regulators, but also built a robust signature to estimate prognosis and guide individualized treatments in LUAD.

---

**Abbreviations:** LUAD, lung adenocarcinoma; m\textsuperscript{7}G, N7-methylguanosine; TCGA, The Cancer Genome Atlas; TMB, tumor mutational burden; DEGs, differentially expressed genes; NSCLC, non-small cell lung cancer; ICB, immune checkpoint blockade; GEO, Gene Expression Omnibus; PCA, principal component analysis; GSEA, Gene Set Enrichment Analysis; FDR, false discovery rate; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristics; ssGSEA, single-sample gene set enrichment analysis; TIDE, tumor immune dysfunction and exclusion; IC50, half-maximal inhibitory concentration; CDF, cumulative distribution function; AUC, areas under the curves; C-index, concordance index.
Materials and methods

Lung adenocarcinoma datasets

We obtained gene expression profile, clinical and somatic mutation data of LUAD from The Cancer Genome Atlas (TCGA) database. Four hundred ninety LUAD cases were included in the follow-up study after removing patients with survival time less than 30 days. We acquired the external verification cohort (GSE31210) from Gene Expression Omnibus (GEO) database. 226 cases were finally included after processing as TCGA data. 29 m7G regulators were extracted from previous report and three m7G-related gene sets in MSigDB database (Supplementary Table 1).

Landscape of genetic variation and identification molecular subtype

The expression of m7G regulators was extracted from TCGA-LUAD dataset. Then various methods were applied to depict the genetic variation of m7G regulators. The expression of m7G regulators was compared between tumor and normal groups. The mutation map was presented by using “maftools” package. Next, Cox analysis was used to filter genes correlated with LUAD prognosis (P < 0.05).

The “ConsensusClusterPlus” R package was employed to identify molecular subtype by consensus clustering of prognostic gene (parameters: reps = 50, pItem = 0.8, pFeature = 1, clusterAlg = “pam,” distance = “Pearson”) (24). Pam and Pearson distances were used as the clustering algorithm and distance measure, respectively. Furthermore, the sample’s distribution was characterized by principal component analysis (PCA). Moreover, we employed “survival” package to investigate the survival differences among subtypes. Besides, “heatmap” package was applied to explore the relation among molecular subtypes, expression of prognostic gene and clinicopathological features.

Biological function analysis and immune infiltration profile estimation

We investigated the biological process of distinct molecular subtype by Gene Set Enrichment Analysis (GSEA). The “h.all.v7.5.1.symbols.gmt” gene set was obtained from MSigDB database. We applied the CIBERSORT algorithm (25) to assess the immune status among different molecular subtypes. In recent years, tumor mutational burden (TMB) was widely applied to measure the effectiveness of ICB therapy (26). TMB score was calculated by using the somatic mutation data of each patient and then we compared TMB score in different subtypes. In addition, patients were further separated into low and high TMB groups on the basis of the threshold value (27) (10 mutations/megabase) of TMB and then we compared the frequency of high TMB in different subtypes. GSE135222 immunotherapy cohort including 27 cases was obtained from GEO database, which used to verify immunotherapy efficacy of different subtypes.

Screening of differentially expressed genes

We calculated the differentially expressed genes (DEGs) between molecular subtypes by using the Bioconductor “limma” package. The significance criteria were |log2FC| > 1 and false discovery rate (FDR) < 0.01. The upregulated and downregulated of DEGs were visualized by volcano map. The heatmap was also applied to show the distribution of DEGs in different subtypes.

Construction and validation of the m7G related signature

The all patients (n = 490) were randomly separated into train and test cohort according to the ratio of 1:1 by using “caret” package. In train cohort (n = 245), univariate Cox analysis was applied to screen genes related to the survival (P < 0.01). Then, least absolute shrinkage and selection operator (LASSO) regression was employed to further reduce the overfitting genes. Finally, a m7G related signature was established by multivariate Cox analysis, and we also called it m7G score. The previously reported formula (28) was used to calculate m7G score: Σ(gene expression level × corresponding coefficient). Patients was separated into high and low m7G score groups according to median m7G score. Then the sample’s distribution was characterized by PCA. We applied “survival” package to investigate the survival differences between two groups. We also plotted the receiver operating characteristics (ROC) curve to estimate the accuracy of the m7G Related signature by using “timeROC” package. The test cohort (n = 245) and the entire cohort (n = 490) were employed to validate the signature power by using the same analyses. We further used GSE31210 dataset (n = 226) to verify the robustness of the signature.

m7G related signature analysis and nomogram construction

Univariate and multivariate analysis were applied to demonstrate the independent prognosis of the m7G score. Then...
these results were visualized with the forest plots. We also constructed a nomogram by combining the age, gender, stage and m^7G score for clinical practice. Additionally, calibration curve was applied to evaluate the predictive accuracy of the nomogram by using “rms” package.

Analysis of immune infiltration and anti-cancer treatment

We applied different bioinformatics methods including XCELL, TIMER, QUANTISEQ, MCPCOUNTER, EPIC, CIBERSORT-ABS, CIBERSORT, and single-sample gene set enrichment analysis (ssGSEA) (29, 30) to study the relation between m^7G score and immune score. Subsequently, we also compared the expression of immune checkpoints between two groups. Mutation maps were manifested by using “maftools” package in two groups. In recent years, in addition to immune checkpoints, tumor immune dysfunction and exclusion (TIDE) was also widely employed to assess the effectiveness of ICB therapy (31). The TMB score was calculated by using the somatic mutation data of each patient and the TIDE score was calculated in the TIDE website^4 (P < 0.05). Moreover, the drug sensitivity of each group was estimated by “pRRophetic” package (32). The half-maximal inhibitory concentration (IC50) of drugs was compared through Wilcoxon rank test between different m^7G score groups (P < 0.05).

Statistical analysis

R software (version 4.1.0) was employed for all data analysis. Wilcoxon rank test was applied to compare m^7G regulators expression between normal and LUAD groups. All above survival distribution was evaluated through survival analysis. The relation among molecular subtype, clinicopathological features and high TMB distribution was estimated by the chi-squared test. Immune infiltration, TMB and TIDE were also compared through Wilcoxon rank test. P < 0.05 was regarded as statistically significant.

Results

Genetic variation profile and m^7G modification pattern

The overall research procedure is shown in Figure 1. Twenty-four regulators were manifested significant downregulation or overexpression in different groups according to P-value less than 0.05 (Figure 2A). The result of mutation map showed that EIF4G3 had the highest mutation frequency followed by LARP1 (Figure 2B). The correlation and prognostic significance of m^7G regulators were presented in Figure 2C. Four genes including EIF4E3, LARP1, WDR4, and NCBP1 were significantly associated with prognosis. m^7G regulators were also showed a remarkable interaction, which was critical for the development of the different m^7G modification patterns. The above-mentioned results suggested m^7G regulators may relate to tumorigenesis and progression in LUAD.

The LUAD patients were classified into two molecular subtypes (C1 and C2) by using "ConsensusClusterPlus" package according to prognostic genes. The intergroup correlation was lowest and intragroup correlation was highest when k = 2 (Figure 3A). Cumulative distribution function (CDF) curve performed the highest partition efficiency when k = 2 (Figures 3B,C). Taken together, two molecular subtypes were established according to the m^7G modification pattern, including 245 patients of C1 and 245 patients of C2. The PCA analysis also demonstrated that the two subtypes could be completely distinguished (Figure 3D). The result of Kaplan–Meier analysis showed distinct survival outcome between two subtypes (P < 0.001) (Figure 3E), suggesting that C1 subtype had worse prognosis than C2. Subsequently, the clinical features and gene expression were compared, then we found patients in C1 subtype had poorer tumor stage than C2. EIF4E3 was upregulated in C2 subtype, while LARP1, WDR4, and NCBP1 were upregulated in C1 subtype (Figure 3F).

Analysis of biological functional and immune infiltration

The results of GSEA presented diverse functional pathways between two subtypes. Functional analysis showed E2F_target, G2M_checkpoint, glycolysis, MITOTIC_spindle, MTORC1_signaling, MYC_targets_V1, MYC_targets_V2 were significantly enriched in C1 subtype (Figure 4A). While, there were significantly different pathways were enriched in C2 subtype, such as allograft rejection, complement, inflammatory response, interferon gamma response, IL2_STAT5_signaling, IL6_JAK_STAT3_signaling (Figure 4B).

Subsequently, we found that two subtypes had markedly different immune infiltration patterns (Figure 4C). The result of CIBERSORT algorithm showed the expression level of T cells follicular helper, resting NK cells, M0 macrophages, activated mast cells were high in C1 subtype, while resting CD4 memory T cells, T cells gamma delta, monocytes, resting dendritic cells, resting mast cells are high in C2 subtype. These suggested that two subtypes may have different immunotherapeutic response, so we further compared TMB score in two subtypes. Then we observed that C1 had

---

4 http://tide.dfci.harvard.edu/
higher TMB score as well as the proportion of high-TMB compared to C2 subtype (22 vs. 10%) (Figures 4D,E). GSE135222 immunotherapy cohort was divided into two subtypes by using the same method mentioned above, including 15 cases in C1 subtype and 12 cases in C2 subtype. The results presented that the proportion of response to immunotherapy was higher in C1 subtype than C2 subtype (40 vs. 17%) (Figure 4F).

**Screening differentially expressed genes between m7G subtypes and construction of m7G related signature**

Based on “limma” package, we identified 1,014 DEGs between m7G Subtypes, including 534 upregulated genes and 480 downregulated genes. Then the significant DEGs were visualized with volcano map (Figure 5A). The expression...
profiles of DEGs between C1 and C2 subtype were visualized with heatmap (Figure 5B). We utilized the train cohort to establish m7G related prognostic signature (n = 245).

First, 201 genes correlated with patient prognosis were found by univariate analysis (Supplementary Table 2). Then we further applied LASSO regression to filter eight genes for the
Identification, survival, and clinical characteristics analysis of molecular subtype. (A–C) The optimal value of consensus clustering. (D) Distribution of all patients. (E) Survival analysis in C1 and C2 subtypes. (F) Heatmap of prognostic gene and clinicopathological features between two subtypes. *P < 0.05.

subsequent multivariate analysis (Figure 5C). Finally, four key genes including E2F7, FAM83A, PITX3, and HOXA13 were identified by using multivariate Cox regression (Figure 5D). The m7G score was calculated with the following formula: \( 0.5171 \times E2F7 \) (mRNA level) + \( 0.1888 \times FAM83A \) (mRNA level) + \( 1.5576 \times PITX3 \) (mRNA level) + \( 0.5210 \times HOXA13 \) (mRNA level). Then we split patients into high and low m7G score groups according to approach mentioned above (Figures 5E–G). The relative expressions of E2F7, FAM83A, PITX3, and HOXA13 in two groups were presented in Figure 5H. Patients had significant poor survival in high m7G score group (Figure 5I). The areas under the curves (AUC) for predicting survival rates at 1-, 3-, and 5-year were 0.736, 0.732, and 0.672, respectively (Figure 5J).

Verification of m7G related signature and analysis of survival in different clinical subgroups

Patients in the internal verification cohort (test and entire cohort) and the external verification cohort (GSE31210) were categorized into two groups on the basis of the same risk formula in train cohort. Patients of test cohort were classified into two groups (Figures 6A–C), which were consistent with the result of the train cohort. The heatmap showed the expression profile of four key gene were apparently different between two groups (Figure 6D). The Kaplan–Meier curve also indicated that two groups had distinct survival (Figure 6E). The area of AUC verified that the signature was a great indicator for assessing prognosis in LUAD (Figure 6F). The similar results were acquired in the entire cohort and GSE31210 cohort (Figures 6G–I, 7A–F). On the basis of Kaplan–Meier analysis of entire TCGA cohort, we also found that patients presented lower survival rate in high m7G score group among different clinical subgroups compared to low m7G score group (Supplementary Figure 1).

Evaluation of association between m7G score, clinicopathological features, and molecular subtype

The results showed strikingly distinct of m7G score in age, gender, N-stage, M-stage, clinical stage and T-stage (Figures 8A–F) (P < 0.05). We also investigated the relation between m7G score, m7G subtype and survival state by using Sankey diagram (Figure 8G). We found that C2 subtype has a strong correlation with low m7G score, while C1 subtype has a strong correlation with high m7G score. Moreover, the majority
FIGURE 4
Functional pathways, immune infiltration and TMB analysis. (A) Functional pathways in C1 subtype. (B) Functional pathways in C2 subtype. (C) Comparison of immune score. (D) Comparison of TMB score. (E) Distribution of low and high TMB between two subtypes. (F) Distribution of response to immunotherapy between two subtypes in GSE135222. *P < 0.05; **P < 0.001; ns: not significant.

of patients with low m7G score were alive, which was consistent with preceding survival analysis. Furthermore, stacked bar chart also presented C1 subtype has a strong correlation with high m7G score (Figure 8H). Similarly, the m7G score was higher in C1 subtype than that in C2 subtype (Figure 8I).

Construction of nomogram and comparison of prognostic signatures

Univariate analysis identified that m7G score was related to poor prognosis in LUAD (Figure 9A). Moreover, m7Gs core was still an independent prognostic indicator after using multivariate analysis (Figure 9B) (P < 0.001). Subsequently, we used m7G score and other clinical factors to establish a nomogram (Figure 9C). Calibration curve demonstrated that 1, 3, 5-year predicted survival rates matched the veritable condition (Figure 9D). These evidences revealed that the m7G score could potentially assist clinical practice to evaluate the prognosis of LUAD patients.

After reviewing previous researches, we further compared the m7G related signature with other prognostic models, including 5-gene signature (Wang) (33), 4-gene signature (Wu) (34), 3-gene signature (Yue) (35), and 5-gene signature (Zhai) (36). In order to ensure comparability among models, the risk score of each LUAD sample in entire TCGA cohort was calculated with the same formula according to corresponding genes in four signatures, and then patients were categorized into two groups based on same cut-off value (37). The results of survival analysis showed significant difference in four models (Figures 10A–C,G) (P < 0.05). However, all AUC at 1-, 3-, and 5-years in four models were lower than that corresponding AUC of our prognostic signature (Figures 10D–F,H). Furthermore, we conducted “survcomp” package to calculate the concordance index (C-index) of each signature. The C-index was highest in our prognostic signature (Figure 10I). Therefore, our signature was more efficient to estimate prognosis in LUAD.

Investigation of immune microenvironment and anti-cancer therapy

Firstly, the results of bubble plot exhibited that CD8+ T cell, common lymphoid progenitor, plasmacytoid dendritic cell, macrophage M1/M0, CD4+ Th1/Th2 cell, neutrophil,
FIGURE 5
Construction of m7G related signature based on TCGA train cohort. (A) Volcano plot of DEGs. (B) Heatmap of DEGs. (C) Eight genes through Lasso regression analysis. (D) Four key genes through multivariate Cox regression. (E,F) Distribution of m7G score and survival state. (G) Distribution of patients according to m7G-related signature. (H) Heatmap of four genes expression between high and low m7G score groups. (I) Survival analysis in two groups. (J) AUC for predicting 1-, 3-, 5-years survival rates.
cytotoxicity score, NK cell, cancer associated fibroblast, monocyte, Myeloid dendritic cell, and mast cell resting were positively correlated with m7G score (Figure 11A). The ssGSEA displayed that activated CD4 T cell, CD56dim natural killer cell, natural killer T cell, neutrophil, Type-2 T helper cell were more active in high m7G score group (Figure 11B) (P < 0.05). Subsequently, the level of CD274 (PD-L1) and PDCD1 (PD-1) were presented upregulated in high m7G score
group (Figures 11C,D) \((P < 0.05)\). Moreover, we compared somatic mutations in two risk groups. The results of mutation map showed remarkably high mutational rate in high \(m^7G\) score group. \(TP53\) was the highest mutational gene in both groups (Figures 12A,B). We also found that high \(m^7G\) score group was related to higher TMB score (Figures 12C,D) \((P < 0.05)\). Compared to low \(m^7G\) score group, high \(m^7G\) score group had strikingly lower TIDE score (Figure 12E) \((P < 0.05)\).

We further investigated common drug sensitivity in two groups (Figure 13). Patients presented lower IC50 of Cisplatin, Docetaxel, Doxorubicin, Etoposide, Gemcitabine, Paclitaxel, and Rapamycin in high \(m^7G\) score group, representing these drugs were more effective for high \(m^7G\) score group. Meanwhile, IC50 of Bicalutamide, Erlotinib, Axitinib, Imatinib, Metformin, Methotrexate, Bexarotene, Sorafenib, and Temsirolimus were lower in low \(m^7G\) score group, representing these drugs were more effective for low \(m^7G\) score group.

**Discussion**

Recently, the role of \(m^7G\) in tumors has begun to receive increasing attention. However, there are no reports on studying the molecular subtype correlated with \(m^7G\) and the implications of \(m^7G\) related signature on the prognosis and immunotherapy in LUAD. Therefore, we expect to discover more tumor phenotypes through this classification, which could be used to evaluate the prognosis of LUAD patients. Here, we first extracted twenty-nine \(m^7G\) related regulators expression profiles from TCGA, four of which were demonstrated to have prognostic value. Then, two novel molecular subtypes were identified according to these prognostic genes. Results showed that patients in C1 subtype had more poor survival outcomes and advanced tumor stages compared to C2 subtype through survival analysis and clinicopathological features comparison, indicating \(m^7G\) regulators correlated with prognosis and progression of LUAD. And the two subtypes presented markedly different molecular features. Compared with C2 subtype, \(m^7G\) regulators were more activated in C1 subtype, including \(LARP1\), \(WDR4\), and \(NCBP1\), while only \(EIF4E3\) was activated in C2 subtype. Xu et al. (38) demonstrated that the expression level of \(LARP1\) was upregulated in NSCLC, which positively related to poor prognosis and progression of cancer. A study of \(WDR4\) uncovered that knockdown of \(WDR4\) could restrain the aggressiveness of NSCLC cells, demonstrating that \(WDR4\) may have tumorigenic function in lung cancer (18). \(NCBP1\) was significantly overexpressed in LUAD, combined with \(CUL4B\), which promoted the proliferation, migration and invasion of
tumor cells (39). It was reported that compared with patients with lower expression of EIF4E3, patients with high expression of EIF4E3 had markedly better survival rates in various cancers, including LUAD (40). These results are consistent with our study, indicating LUAD patients with C1 subtype have poor prognosis compared to C2 subtype. Furthermore, we investigated the possible functional mechanisms in both two subtypes by using GSEA analysis. Interestingly, the functional pathways enriched in C1 subtype were mainly cell proliferation-related pathways, which may indicate advanced clinicopathological staging, adverse survival outcomes and aggressive tumor subtypes. This evidence also suggested the worse prognosis of C1 subtype. Immune cells as an important part of the tumor microenvironment intensively relate to the response to immunotherapy (41). Recently studies showed that RNA modifications were correlated with the differentiation of immune cells in the tumor microenvironment (42). As one of the RNA modifications, m7G also influenced immune cells in the tumor microenvironment. Chen et al. (43) quantified the tumor-infiltrating lymphocytes and found that CD4+ T exhaustion and Tregs decline after knockout of m7G regulators. Besides, Devarkar et al. (44) presented that m7G was involved in innate immunity mediated by RIG-I. Therefore, the immune score was applied to characterize immune microenvironmental
landscapes in two subtypes. Then we found that follicular helper T cells, resting NK cells and activated mast cells were increased in C1 subtype. It was reported that follicular helper T cells could recruit CD8+ T cells to enhance antitumor immune response (45). The resting NK cells could secrete various cytokines to kill target cells, for example tumor cells (46). The TLR4 activation by mast cells resulted in the secretion of CXCL10, which could recruit effector T cells to influence antitumor immune response (47). These evidences indicated that C1 subtype with activated m7G regulators may more sensitive to immunotherapy compared to C2 subtype.

Increasing investigations suggested that TMB is a biomarker of response to immunotherapy and is positively relate to the effectiveness of ICB in various cancers, including NSCLC (48). In this study, the TMB score in patients with C1 subtype was distinctly higher compared with C2 subtype. After dividing TMB into high and low groups, more percentage of high TMB was observed in patients with C1 subtype compared with C2 subtype. Consequently, patients with C1 subtype may present better immunotherapeutic response than C2 subtype.

Besides, by using an immunotherapy cohort of lung cancer, we also demonstrated that patients with C1 subtype had better immunotherapy efficacy than C2 subtype.

Considering the individual heterogeneity of m7G modification, we utilized a novel m7G score to quantify the m7G modification patterns in LUAD. In the TCGA train cohort, we identified four key genes (E2F7, FAM83A, HOXA13, and PITX3), and then calculated the m7G score through the previously mentioned algorithm. After separating patients into high and low m7G score groups, we found that the four genes were overexpressed in the high m7G score group. It was reported that overexpression of E2F7 correlated with poor prognosis and microRNA-935 could inhibit tumor metastasis and invasion by targeted suppression the level of E2F7 in NSCLC (49). Wang et al. (50) found that activating the expression of E2F7 expression by targeting microRNA-140-3p could promote the progression of LUAD. Studies presented that FAM83A was significantly related to TMB and DNA damage response pathways in NCSLC (51, 52), indicating that it may play an important part in tumor progression and immunotherapy.
Hu et al. (53) demonstrated that the expression of FAM83A regulated the proliferation and invasiveness of NSCLC through PI3K/Akt/mTOR pathway. Investigations showed HOXA13, as a nuclear transcription factor, was related to tumor cells proliferation and differentiation, which could accelerate tumor aggressive characteristics through disturbing P53 and Wnt/β-catenin signaling pathways in NSCLC (54). One research reported that HOXA13 was markedly upregulated and strongly correlated with tumorigenesis and progression in LUAD (55). Some studies demonstrated that PITX3 as a transcription factor was involved in many tumors (56, 57). Zhang et al. (58) presented that high expression of PITX3 was strongly associated with the poor prognosis in LUAD. According to these evidences, we indicated that patients with high m7G score in which these four genes were activated, had poor survival. Also, survival analysis demonstrated that the high m7G score group presented worse survival outcomes. ROC curves further showed the great efficacy of m7G score to predict survival rate. And, the TCGA test cohort, entire cohort and GSE31210 cohort were applied to validate the accuracy and reliability of the m7G related
signature. Similar results were acquired from all validation cohorts, demonstrating that the prognostic signature may be a robust biomarker to evaluate the prognosis in LUAD. Patients with high m7G score also presented remarkably poor survival condition among different clinical subgroups. After analyzing the association between m7G score and clinicopathological parameters, we observed that m7G score was significantly high in N2+N3, M1, Stage III+IV and T2-T4, suggesting high m7G score is associate with cancer progression. The characterization of m7G modification patterns showed C2 subtype had lower m7G score compared with C1 subtype. And, the high m7G score correlated with poor survival and cancer progression was consistent with characteristics of C1 subtype. Univariate and multivariate analysis presented that the m7G score was an independent prognosis predictor of LUAD patients. Subsequently, the nomogram also presented high accuracy in predicting survival rate of 1, 3, 5-years. In recent years, many signatures were built to predict the prognosis of LUAD patients. Furthermore, we presented that the AUC area and C-index of our signature were higher than the other four public prognostic signatures, suggesting our signature have better performance in predicting clinical prognosis in LUAD patients.

Currently, although the immunotherapy of lung cancer has got great progress, how to choose the appropriate therapeutic regime for patients is still a clinical challenge. Besides, a part of patients did not obtain effective benefits from immunotherapy (59), and even some patients will undergo obvious side effects during therapy (60). Therefore, it is critical to explore a novel method to guide individualized and precise treatment in LUAD patients. The results of various evaluation methods of immune cell infiltration showed distinct activation of immune cells in both groups, which were similar to molecular subtypes. We speculated there was different immunotherapeutic response in two groups, so we further investigated the association between m7G score groups, immune checkpoint, TMB and TIDE. The immune checkpoints are also an integral part of the immune system and participate in regulating immune escape (61). In recent years, immunotherapy targeting immune checkpoints has obtained huge clinical therapeutic results, especially anti-PD-1/PD-L1 antibody (62). In the study, we observed that patients with high m7G score had upregulated PD-1/PD-L1, indicating that these patients may be more sensitive to ICB than low m7G score. Subsequently, compared with the low m7G score group, the high m7G score group had markedly higher TMB which was consistent with the
In addition, we found that compared with the patients with low m²G score, patients with high m²G score had more percentage of TP53 mutation. Studies showed that TP53 mutation was remarkably related to high PD-L1 expression and patients with TP53 mutation could acquire benefits from ICB therapy in LUAD (63, 64). Investigations presented TIDE was an accurate biomarker used to predict the immunotherapeutic effects of NSCLC, which was negatively associated with the efficacy of ICB (31). Meanwhile, recent studies have reported the clinical application of TIDE in predicting and evaluating immunotherapeutic response (65, 66). In our study, compared with patients with low m²G score patients with high m²G score had lower TIDE score, suggesting patients in the high m²G score group may obtain clinical benefits from immunotherapy. Integration analysis of the m²G score, immune cell infiltration, immune checkpoint, TMB, and TIDE indicated that the signature is a potential biomarker to assess immunotherapeutic response and tailor individualized treatment for LUAD patients.

Chemotherapy is a classic treatment for lung cancer, but patients have different response rates to chemotherapy drugs. Selecting an appropriate chemotherapy regimen is helpful to improve the prognosis and reduce the economic burden of patients. Our study revealed that common drugs including Cisplatin, Docetaxel, Doxorubicin, Etoposide, Gemcitabine, Paclitaxel, and Rapamycin were suitable for patients with high m²G score, while Axitinib, Bexarotene, Bicalutamide, Erlotinib, Imatinib, Metformin, Methotrexate, Sorafenib, and Temsirolimus were more appropriate for patients with low m²G score.

Our research may assist in judging the prognosis in LUAD, but there are also some limitations. First, the research is a
FIGURE 13
Comparison of drug sensitivity between low and high m7G score groups.

A retrospective study according to the data from TCGA and GEO datasets, so it’s crucial to collect prospective clinical data to further verify the signature. Second, the potential functional mechanisms of m7G score are not fully verified, so these need to be further verified by experiments at the molecular level in vivo and in vitro. Finally, the drug response in patients is based on methodological prediction, so clinical trials need to be implemented in the future.

Conclusion

In summary, we identified two novel molecular subtypes of LUAD according to m7G regulators. The survival, immune infiltration, and TMB are significantly different in two subtypes. The m7G related signature to quantify the heterogeneity of the two subtypes was constructed. The signature can be employed to predict prognosis in LUAD, then the internal and external cohort were applied to verify the prognostic value. And the signature was elucidated be helpful to guide immunotherapy and chemotherapy. Therefore, this research provides a new direction for improving prognosis and current anti-cancer strategies in LUAD.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories...
and accession number(s) can be found in the article/Supplementary material.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

ZL and XY conceived and designed the study. ZL and WW performed the data analysis. ZL wrote the manuscript. WW and JW participated in collecting the data and helped draft the manuscript. XY and WW prepared and edited the manuscript. All authors reviewed and approved the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 81660493) and the Natural Science Foundation of Jiangxi Province (Grant No. 20202ACBL206019).

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.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.962972/full#supplementary-material

References

1. Sharma A, Shepard JO. Lung cancer biopsies. Radiol Clin North Am. (2018) 56:377–90. doi: 10.1016/j.rcl.2018.01.001
2. Song Q, Shang J, Zhang C, Zhang L, Wu X. Impact of the homogeneous and heterogeneous risk factors on the incidence and survival outcome of bone metastasis in NSCLC patients. J Cancer Res Clin Oncol. (2019) 145:737–46. doi: 10.1007/s00432-018-02826-7
3. Lin JJ, Cardarella S, Lydon CA, Dahlberg SE, Jackman DM, Janne PA, et al. Five-year survival in EGFR-mutant metastatic lung adenocarcinoma treated with Egfr-TKis. J Thorac Oncol. (2016) 11:556–65. doi: 10.1016/j.jtho.2015.12.103
4. Duma N, Santana-Davila R, Molina JR. Non-small cell lung cancer: Epidemiology, screening, diagnosis, and treatment. Mayo Clin Proc. (2019) 94:1623–40. doi: 10.1016/j.mayop.2019.01.013
5. Zou W, Wolkoch JD, Chen L. Pd-L1 (B7-H1) and Pd-1 pathway blockade for cancer therapy. Mechanisms, response biomarkers, and combinations. Sci Transl Med. (2016) 8:328rv4. doi: 10.1126/scitranslmed.aad7118
6. Syn NL, Teng MWL, Mok TSK, Soo RA. De-novo and acquired resistance via the Rptor/Ulk1/Autophagy axis. Theranostics. (2019) 9:4:1623–40. doi: 10.1016/j.mayop.2019.01.013
7. Mehta A, Dobersch S, Romero-Olmedo AJ, Barreto G. Epigenetics in lung cancer diagnosis and therapy. Cancer Metastasis Rev. (2015) 34:229–41. doi: 10.1007/s10555-015-9563-3
8. Sun X, Yi J, Yang J, Han Y, Qian X, Liu Y, et al. An integrated epigenomic-transcriptomic landscape of lung cancer reveals novel methylation driver genes of diagnostic and therapeutic relevance. Theranostics. (2021) 11:5346–64. doi: 10.7150/thno.58085
9. Boccaletto P, Machnicka MA, Purta E, Pątkowski P, Bagiński B, Wierckzi TK, et al. Modomics: A database of rna modification pathways. 2017 Update. Nucleic Acids Res. (2018) 46:D303–7. doi: 10.1093/nar/gkx1030
10. Han X, Wang M, Zhao YL, Yang Y, Yang YG. Rna methylations in human cancers. Semin Cancer Biol. (2021) 75:97–115. doi: 10.1016/j.semcancer.2020.11.007
11. Cheng Y, Wang M, Zhou J, Dong H, Wang S, Xu H. The important role of N6-methyladenosine Rna modification in non-small cell lung cancer. Genes (Basel). (2021) 12:440. doi: 10.3390/genes12040440
12. Zhao LY, Song J, Liu Y, Song CX, Yi C. Mapping the epigenetic modifications of DNA and Rna. Protein Cell. (2020) 11:792–808. doi: 10.1007/s13238-020-00733-7
13. Cowling VH. Regulation of Mrna Cap Methylation. Biochem J. (2009) 425:295–302. doi: 10.1042/BJ20091352
14. Zhang LS, Liu C, Ma H, Dai Q, Sun HL, Luo G, et al. Transcriptome-wide mapping of internal N(7)-methylguanosine methylome in mammalian Mrna. Mol Cell. (2019) 74:1304–16.e8. doi: 10.1016/j.molcel.2019.03.036
15. Malbec L, Zhang T, Chen YS, Zhang Y, Sun BF, Shi BY, et al. Dynamic methylome of internal Mrna N(7)-Methylguanosine and its regulatory role in translation. Cell Res. (2019) 29:927–41. doi: 10.1038/s41422-019-0230-z
16. Han H, Yang C, Ma J, Zhang S, Zheng S, Ling R, et al. N(7)-Methylguanosine Trna modification promotes esophageal squamous cell carcinoma tumorigenesis via the Rptor/Ulk1/Autophagy axis. Nat Commun. (2022) 13:1478. doi: 10.1038/s41467-022-29125-7
17. D’Abrono LO, Ghosh PM. Eif1f phosphorylation in prostate cancer. Neoplasia. (2018) 20:563–73. doi: 10.1016/j.neo.2018.04.003
18. Ma J, Han H, Huang Y, Yang C, Zheng S, Cai T, et al. Mettl1/Wdr4-Mediated M7G Trna modifications and M7G(cod) usage promote Mrna translation and lung cancer progression. Mol Ther. (2021) 29:3422–35. doi: 10.1038/s41397-021-01876-0
19. Zhang M, Song J, Yuan W, Zhang W, Sun Z. Roles of Rna methylation on tumor microenvironment and clinical implications. Front Immunol. (2021) 12:641507. doi: 10.3389/fimmu.2021.641507
20. Yu P, Tong L, Song Q, Yu H, Chen Y. Systematic profiling of invasion-related gene signature predicts prognostic features of lung adenocarcinoma. J Cell Mol Med. (2021) 25:6388–402. doi: 10.1111/jcmm.16619
21. Li B, Huang Z, Wu Y, Liu S, Zhang J, Wang Q, et al. Molecular subtypes based on Ccvs related genes identify candidate prognostic biomarkers in lung adenocarcinoma. Neoplasia. (2021) 23:704–17. doi: 10.1016/j.neo.2021.05.006
22. Xu D, Li C, Zhang Y, Zhang J. DNA methylation molecular subtypes for prognosis prediction in lung adenocarcinoma. JIMD Palm Med. (2022) 22:133. doi: 10.1186/s12890-022-01924-0
23. Tomikawa C. 7-Methylguanosine modifications in transfer Rna (Trna). Int J Mol Sci. (2018) 19:4980. doi: 10.3390/ijms19124800
24. Winkler MD, Hayes DN. Consensusclusterplus: A class discovery tool with confidence assessments and item tracking. Bioinformatics. (2010) 26:1572–3. doi: 10.1093/bioinformatics/btp170
25. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. (2015) 12:453–7. doi: 10.1038/nmeth.3337
26. Chan TA, Xarchos M, Jaffe E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker. Utility for the oncology clinic. Ann Oncol Off J Eur Soc Med Oncol. (2019) 30:44–56. doi: 10.1093/annonc/mdy495
27. Reck M, Schenker M, Lee KH, Provenco M, Nishio M, Lesniowski-Kmik K, et al. Nivolubam plus ipilimumab versus chemotherapy as first-line treatment in advanced non-small-cell lung cancer with high tumour mutational burden: Patient-reported outcomes results from the randomised, open-label, phase III checkmate 227 trial. Eur J Cancer. (2019) 116:137–47. doi: 10.1016/j.ejca.2019.05.008
28. Mo Z, Cao Z, Luo S, Chen Y, Zhang S. Novel molecular subtypes associated with 5mc methylation and their role in hepatocellular carcinoma immunotherapy. Mol Ther. (2020) 28:562441. doi: 10.1038/s41397-2020-562441
29. Sturm G, Finotello F, Pretorpe F, Zhang JD, Baumbach J, Fridman WH, et al. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. Bioinformatics. (2019) 35:4346–45. doi: 10.1093/bioinformatics/btz585
30. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep. (2017) 18:488–62. doi: 10.1016/j.celrep.2016.12.019
31. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Methods. (2018) 24:1550–8. doi: 10.1038/s41591-018-0136-1
32. Geeleher P, Cox N, Huang RS. Prrophetic: An R package for prediction of immunophenotypes from gene sets. Bioinformatics. (2019) 35:i436–45. doi: 10.1093/bioinformatics/btx366
33. Li et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep. (2019) 7:e8128. doi: 10.1016/j.celrep.2019.03.013
34. Tomikawa C. 7-Methylguanosine modifications in transfer Rna (Trna). Int J Mol Sci. (2018) 19:4980. doi: 10.3390/ijms19124800
35. Chen J, Li K, Chen J, Wang L, Cheng M, et al. Aberrant transcription regulated by Mettl1/Wdr4-Mediated Trna N7-methylguanosine modification drives head and neck squamous cell carcinoma progression. Cancer Commun (Lond.). (2020) 42:223–44. doi: 10.1002/cac2.12273
36. Kaesler S, Welbigni W, Kobele M, Sailer V, Pitx3 DNA methylation is an independent predictor of overall survival in patients with head and neck squamous cell carcinoma. Blood. (2006) 107:159–66. doi: 10.1182/blood-2005-04-1351
37. Goodman AM, Kato S, Buzhenska L, Patel SP, Frampton GM, Miller V, et al. Neoutinegulant-driven B cell and Cdt4 t follicular helper cell collaboration promotes anti-tumor Cdt8 T cell responses. Cell. (2021) 184:6101–18.e13. doi: 10.1016/j.cell.2021.11.007
38. Wang C, Li S, Xu J, Niu W, Li S. Microrna-935 is reduced in non-small-cell lung cancer tissue, is linked to poor outcome, and acts on signal transduction mediator E2f7 and the Akt pathway. Br J Biol Sci. (2015) 26:1426–41. doi: 10.1186/s12890-022-01924-0
39. Wang Y, Wu Y, Tu L, Sun X, Liu A, Dong Y, et al. Circ-aasdh functions as a sensor of the progression of early stage lung adenocarcinoma by targeting Mir-140-3p to Activate E2f7 expression. Transl Lung Cancer Res. (2021) 10:57–70. doi: 10.21037/tlcr-20-1062
40. Ren J, Yang Y, Li C, Xie L, Hu R, Qin X, et al. A novel prognostic model of early-stage lung adenocarcinoma integrating methylation and immune biomarkers. Front Genet. (2020) 11:634634. doi: 10.3389/fgene.2020.634634
41. Luo R, Song J, Xiao X, Xue Z, Zhao Z, Zhang W, et al. Identifying CpG methylation signature as a promising biomarker for recurrence and immunotherapy in non-small-cell lung cancer. Aging. (2020) 12:14649–76. doi: 10.18632/aging.203517
42. Wang Y, He B, Dong Y, He GJ, Qiu XD, Li Y, et al. Homebox-A13 acts as a functional prognostic and diagnostic biomarker via regulating P53 and Wnt signaling pathways in lung cancer. Cancer Biomark. (2021) 31:239–54. doi: 10.3233/CBM-200540
43. Deng Y, He H, Zhang G, Tan B, Zhang Y, Chen G, et al. The expression of Pitx3 in lung adenocarcinoma and its clinical significance: A study based on the cancer genome atlas, oncomine and reverse transcription-quantitative polymerase chain reaction. Oncol Lett. (2018) 15:8556–72. doi: 10.3892/ol.2018.8381
44. Holmes EE, Goltz D, Sailer V, Jung M, Miller S, Uhr B, et al. Ptx3 promoter methylation is a prognostic biomarker for biochemical recurrence-free survival in prostate cancer patients after radical prostatectomy. Clin Epigenet. (2016) 8:104. doi: 10.1186/s13548-016-0270-x
45. Sailer V, Holmes EE, Grevenhoven H, Goltz D, Droege F, Fransen A, et al. Ptx3 DNA methylation is an independent predictor of overall survival in patients with head and neck squamous cell carcinoma. Clin Epigenet. (2017) 9:12. doi: 10.1186/s13548-017-0317-7
58. Zhang C, Chen X, Chen Y, Cao M, Tang J, Zhong B, et al. The pitx gene family as potential biomarkers and therapeutic targets in lung adenocarcinoma. Medicine (Baltimore). (2021) 100:e23936. doi: 10.1097/MD.0000000000002936

59. Kurtulus S, Madi A, Escobar G, Klapholz M, Nyman J, Christian E, et al. Checkpoint blockade immunotherapy induces dynamic changes in Pd-1(-)Cd8(+) tumor-infiltrating T cells. Immunity. (2019) 50:181–94.e6. doi: 10.1016/j.immuni.2018.11.014

60. Pauken KE, Dougan M, Rose NR, Lichtman AH, Sharpe AH. Adverse events following cancer immunotherapy: Obstacles and opportunities. Trends Immunol. (2019) 40:511–23. doi: 10.1016/j.it.2019.04.002

61. Ramsay AG. Immune checkpoint blockade immunotherapy to activate anti-tumour T-cell immunity. Br J Haematol. (2013) 162:313–25. doi: 10.1111/bjh.12380

62. Sharpe AH. Introduction to checkpoint inhibitors and cancer immunotherapy. Immunol Rev. (2017) 276:5–8. doi: 10.1111/imr.12531

63. Schoenfeld AJ, Rizvi H, Bandlamudi C, Sauter JL, Travis WD, Rekhtman N, et al. Clinical and molecular correlates of Pd-L1 expression in patients with lung adenocarcinomas. Ann Oncol Off J Eur Soc Med Oncol. (2020) 31:599–608. doi: 10.1016/j.annonc.2020.01.065

64. Dong Z-Y, Zhong W-Z, Liu S-Y, Xie Z, Wu S-P, Wu YL. Ma15.10 potential predictive value of Tp53 and Kras mutation status for response to Pd-1 blockade immunotherapy in lung adenocarcinoma. J Thoracic Oncol. (2017) 12:S432–3. doi: 10.1016/j.jtho.2016.11.504

65. Bretz AC, Paraskeva U, Kronthaler K, Dreker T, Bartz R, Hermann F, et al. Domatinostat favors the immunotherapy response by modulating the tumor immune microenvironment (Time). J Immunother Cancer. (2019) 7:294. doi: 10.1186/s40425-019-0745-3

66. Jin R, Liu C, Zheng S, Wang X, Feng X, Li H, et al. Molecular heterogeneity of Anti-Pd-1/Pd-L1 immunotherapy efficacy is correlated with tumor immune microenvironment in East Asian patients with non-small cell lung cancer. Cancer Biol Med. (2020) 17:768–81. doi: 10.20892/j.issn.2095-3941.2020.0121