The Cancer Genome Atlas Predicts the Prognosis of Lung Carcinoma Based on Genes Associated with m6A Methylation

Huanqing Liu

The First Affiliated Hospital of Jinan University

Tingting Li

Xi'an Chest Hospital

Chunsheng Dong

Shaanxi Normal University

Jun Lyu (✉ lyujun2020@jnu.edu.cn)

The First Affiliated Hospital of Jinan University

Research Article

Keywords: Lung cancer, N6-methyladenosine (m6A), mRNA, Cancer Genome Atlas

Posted Date: July 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-688141/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Lung cancer has become a predominant cause of death in relation with carcinoma worldwide. N6-methyladenosine (m6A) is a common mRNA that is internally modified, which has a pivotal role in mRNA splicing, outputting, localizing, and translating and in identifying stable features. This study evaluated the expression pattern and prognostic value of m6A-related genes in lung cancer. Expression data of lung cancer samples with related clinical information were obtained from The Cancer Genome Atlas (TCGA). Then, R software was used in combination with several corresponding software packages to identify the regulatory factors of m6A RNA methylation with differential expression. Three genes (METTL3, YTHDF1, and FTO) were overexpressed in lung cancer. METTL3 had a low survival rate ($P < 0.05$). Significant differences in survival rate were observed among the subgroups, which possessed differently expressed m6A levels. Two latent predicting factors (METTL3 and KIAA1429) that met the independent predictive values were selected. M6A RNA methylation modulators may be involved with the malignant progression of lung cancer, and the two selected risk characteristics of m6A RNA methylation regulators may be a potential prognostic biological marker for guiding customized therapies in patients with lung carcinoma.

Introduction

Pulmonary carcinoma has become a common malignancy possessing great morbidity, and its death rate ranks first in the world with a 5-year survival rate of below 20%. In the recent decades, various therapeutic methods have been developed, such as surgical resection, chemotherapy, radiotherapy, and molecular targeting therapy, but the whole survival status for patients with pulmonary carcinoma is not improved probably because the early symptoms of lung cancer are hidden; thus, the majority of patients have been confirmed in the middle or terminal stages [1]. In addition, effective molecular markers to target therapy are lacking [2]. Therefore, the development of a reliable prognostic predictor for lung cancer patients is important for accurate individualization of treatment.

Post-transcriptional modification is a significant regulatory factor for many physiological processes and illness development, which has attracted increasing attention in biology. N6-methyladenosine (m6A) becomes the most suitable mRNA modification among the massive RNA modifications. M6A is involved in almost every step of RNA metabolism comprising mRNA translating, degradation, splicing, outputting, and folding [3–4]. Dysregulation of the m6A regulator results in reduced cellular proliferation, loss of self-renewal, developmental defects, and cellular death [5]. The regulator of m6A methylation is involved in the occurrence and development of various carcinomas, such as glioblastoma, hepatocellular carcinoma, and breast cancer [6–8]. Although m6A becomes a focal point for various explorations recently, people's understanding of this modification remains insufficient. Furthermore, the study of m6A-methylation-related genes is a small combinatorial study. The function of m6A-methylation-related genes in lung carcinoma remains unknown. In this study, the relation of the expression of m6A-methylation-related genes with clinical prognosis among lung carcinoma patients was analyzed. Through Cox univariate analysis and Lasso Cox regression analysis, the regulatory factors of m6A methylation were selected to

construct the risk characteristics. Then, the function of risk signals in lung cancer prognosis was analyzed. Finally, new prognostic biological markers and therapy targets were obtained.

Methods And Materials

Data collection

The transcriptomic information of lung cancer and the relevant clinical information were obtained from The Cancer Genome Atlas (TCGA) database [9] (<https://portal.gdc.cancer.gov/>). The mRNA expression information for 1037 tumor tissues and 108 healthy tissues was collected, and the clinical data of 1026 with pulmonary carcinoma patients, comprising age, gender, grading, clinical staging, and NM staging, were collected.

Data preprocessing and differential expression analysis of m6A RNA methylation regulators

The differential expression of 15 m6A-methylation-related genes in lung carcinoma and normal control samples should be assessed using R software (Version 3.8; <http://www.bioconductor.org/packages/release/bioc/html>). R software was utilized to plot the heatmap of the 15 genes.

Bioinformatic analysis/statistics

Pearson correlation analysis was performed to identify the correlation among genes, and an m6A methylation regulator interaction network was constructed by Cytoscape. Kaplan–Meier (KM) survival analyses were conducted to assess the impact of each gene on the survival status according to the UALCAN website (<http://ualcan.path.uab.edu/index.html>). *P*-value of 0.05 indicated the significant threshold for all tests.

Consensus-clustering analysis

The TCGA lung cancer cohort was clustered into distinctive groups through consensus expression of m6A RNA methylation regulators using “ConsensusClusterPlus” in R to determine whether the expressing level of m6A RNA methylation regulators was associated with the prognosis of pulmonary cancer. Utilizing KM and log-rank test, the distinctions of overall survival (OS) among the groups were tested.

Prognostic value of m6A-methylation-related genes

Univariate Cox regression analysis screened out m6A regulatory factors with significant differences for further analysis. Subsequently, LASSO regression of the high-dimensional data was performed, and “glmnet” package in R software was utilized to select the prognostics factor with the greatest value. The patients were divided into the high-risk group and low-risk group, which was in accordance with the median expression of m6A-methylation-related genes. The KM survival method was used to analyze the relation of m6A-related genes with the survival rate. *P* value for the KM survival curve was calculated by

utilizing the log-rank test. In verifying the accuracy of the model, receiver operating characteristic (ROC) curves were plotted. After univariate and multivariate Cox regression analyses, the prognostic factors of lung cancer were determined.

Statistical analysis

R3.5.2 and Cytoscape v.3.7.2 were used for all statistical analyses. Survival time was analyzed by univariate Cox regression. Hazard ratios and 95% CI were calculated to identify genes associated with OS. K-M survival analyses were performed by utilizing the log-rank test. $P < 0.05$ indicated statistical significance unless otherwise stated.

Results

mRNA expression pattern of m6A-related genes among pulmonary carcinoma

This study included 1037 pulmonary carcinoma samples and 108 adjacent normal lung samples from TCGA. Clinical information comprised age, gender, grading, clinical staging, and TNM staging. The expression profiles of all known m6A-related genes in lung cancer were analyzed on the basis of TCGA datasets. The heatmap of expression of m6A-related genes was constructed to understand the expression in lung cancer. As shown in Fig. 1, three genes (YTHDF1, FTO, and METTL3) evidently increased in pulmonary carcinoma versus adjacent normal tissues. Pearson correlation analysis demonstrated the positive correlation of such regulators with one another in a weak-to-moderate manner, in which YTHDF3 and KIAA1429 had the strongest correlation. In addition, HNRNPC and HNRNPCA2B1 were negatively associated, which indicated that HNRNPCA2B1 expression might be downregulated with the upregulation of HNRNPC expression (Fig. 2).

Survival analysis of m6A-methylation-related genes

UALCAN (<http://ualcan.path.uab.edu/index.html>) online tool should be utilized to identify the survival information for three genes. Based on the findings, patients with high expression of METTL3 showed an inferior survival status ($P < 0.05$, Fig. 3). By contrast, the remaining two genes were not significantly different ($P > 0.05$).

Based on the m6A-methylation-related gene expression levels, different subgroups of the 374 tumor samples were identified by utilizing R's ConsensusClusterPlus package. Moreover, cluster-consensus and item-consensus outcomes were calculated. As shown in Fig. 4(A–C), the output manifested k (2 to 4) subgroups. In line with our findings, k = 2 indicated that full selection was achieved. All patients were classified into two subgroups with regard to the k value with the greatest stability (Fig.4(A–E)).

As shown in Figure 4(F), subgroup 1 presented a great gene expression level, whereas subgroup 2 failed to show such outcome. The horizontal axis indicated the first main component, whereas the vertical axis indicated the second component. According to the principal component analysis, subgroup 1 and subgroup 2 could be assembled. The outcome suggested the accuracy of this grouping.

Prognostic value of m6A-methylation-related genes and identification of prognostics features

Univariate Cox regression was utilized to analyze 15 genes and comprehensively understand the prognosis of m6A-methylation-related genes among pulmonary carcinoma. Two candidate genes were selected, and $P < 0.05$ was set as the screening condition (Figure 5(A)). The LASSO Cox regression model was used in selecting genes with the greatest prediction as prognostics markers. λ should be selected when the minimum median of the residual sum of squares occurred. Two latent predictive factors (Figures 5(B–C)), METTL3 and KIAA1429, were identified as prognostics factors for lung carcinoma. The risk score for two genes was calculated for further univariate and multivariate Cox regression analyses.

On the basis of the cutoff value combination model of the median expression of the two candidate genes, the patients were divided into the high-risk group and low-risk group. The prognosis of the low-risk group was better than the high-risk group (Figure 5(D)).

Prognostic value of m6A-methylation-related genes

Based on univariate analyses, the T stage, M stage, N stage, and risk score of m6A-methylation-related genes affected patients' prognostic status ($P < 0.05$, Figure 6A). Multivariate regression analysis showed that the age, T stage, and risk score of m6A-methylation-related genes was an independent factor for the prognostic status of pulmonary carcinoma ($P < 0.05$, Figure 6B).

Discussion

Lung cancer has great aggression and rapid fatalness. Despite remarkable technical advances over the past few decades, the 5-year survival rate of pulmonary carcinoma remains unexplored. Therefore, identifying a new therapy for pulmonary carcinoma and the factor affecting clinical prognosis is necessary. Numerous studies have shown that tumorigenesis and development can be facilitated through genomic and epigenetic changes such as DNA methylation [10]. The common internal modification of eukaryotic mRNAs, m6A RNA methylation, is important for outstanding prediction of malignant behavior and clinical prognosis of various carcinomas and for early diagnosis and treatment of pulmonary adenocarcinoma [11–13].

In this study, we evaluated whether m6A-related genes could be used as a new prognostic biological marker for lung carcinomas. Based on the TCGA bioinformatics, 15 gene signatures were established to forecast the prognostic status for pulmonary carcinoma, comprising METTL3, METTL14, KIAA1429, RBM15, ZC3H13, WTAP, YTHDF1, YTHDF2, YTHDF3, YTHDC1, YTHDC2, HNRNPC, HNRNPA2B1, ALKBH5, and FTO. We also analyzed the relation of those regulators with clinical pathological characteristics and finally constructed a two-gene risk signature. These results indicated that the regulatory factor of m6A RNA methylation played a significant function in the pathogenesis of pulmonary cancer, which can be used as a reliable prognostic factor.

Our current analysis showed that three of the 15 m6A RNA methylation regulatory genes (YTHDF1, FTO, and METTL3) were upregulated in pulmonary cancer tissues, indicating that these genes may have a main role in the development and progression of pulmonary cancer. In addition, m6A-methylation-related genes had close relation with one another in the regulatory network, suggesting that they were cooperatively coordinated in cancer development. Moreover, METTL3 may have a deleterious effect on lung cancer patients because of its association with poor survival status. The findings indicated that m6A modulator could be a potential target for pulmonary carcinoma treatment.

In survival analyses, the patient with high expression of METTL3 significantly deteriorated the survival rate, which indicated that METTL3 was identified as a prognostic gene in the TCGA database. Visvanathan et al. [14] revealed that the upregulation of METTL3 was associated with inferior survival status in glioblastoma cells, which was consistent with our analyzing outcomes. Next, we used R's ConsensusClusterPlus package to divide the lung cancer tissue data into two subgroups based on the gene expression levels. Cluster analysis of the gene expression profile is an important research topic in the diagnosis of cancer subtypes to provide accurate treatment for cancer patients [15]. Principal component analysis showed that subgroup 1 and subgroup 2 were separated. Overall survival analysis revealed significant improvement in survival time of ` 1, suggesting that the survival time was related to the general expression level of m6A-methylation-related genes.

Lasso Cox analysis identified METTL3 and KIAA1429 as prognostic factors for pulmonary carcinoma. In the prognosis of pulmonary carcinoma, the predictive force of m6A-methylation-related gene was evaluated using the ROC curve. The m6A-methylation-related gene played a role in the survival of lung cancer. METTL3 was initially identified as a methyltransferase, which involved the modification of m6A [16]. Recently, studies show that METTL3 has an important function in a variety of cancer types, depending on or independent of its m6A RNA methyltransferase activity. In most cases, METTL3 has been reported as an oncogene that promotes initiation and development of various cancers by depositing m6A modifications on key transcripts. Vu et al. [17] found that METTL3 was more abundant in acute myeloid leukemia cells than in normal hematopoietic stem and progenitor cells, and differentiation of hematopoietic stems and progenitor cells was inhibited when wild-type METTL3 was overexpressed. METTL3 consumption significantly suppressed tumorigenicity and metastasis, but in patients with hepatocellular carcinoma, METTL3 was significantly upregulated, and the OS rate was shortened [18]. METTL3 expression was elevated in pulmonary adenocarcinoma, and METTL3 promoted the growth, survival, and invasion of human pulmonary carcinoma cells [19]. Lin et al. reported that METTL3 could promote cellular growth, survival, and invasion by increasing the expression of EGFR and TAZ [20], revealing the significant function of METTL3 in boosting the translation of oncogenes in human pulmonary carcinoma. KIAA1429 is an RNA-binding protein, which involves the modification of m6A, mRNA splicing, and processing [21]. KIAA1429 is considered as a scaffold that coordinates the core component of METTL3/METTL14/WTAP on an RNA substrate for specific m6A methylation near the 3' UTR and stop codon [22–23]. KIAA1429 is significantly upregulated in hepatocellular carcinoma tissues, and its high expression is associated with poor prognosis in hepatocellular carcinoma patients. In addition, silencing KIAA1429 inhibits cellular proliferation and metastasis in vitro and in vivo [24].

KIAA1429 is also highly expressed in breast carcinoma tissues and is often downregulated in noncancerous breast tissues. Moreover, the subsequent in vivo and in vitro studies have shown that KIAA1429 is involved in the proliferation and metastasis of breast carcinoma [23]. Tang et al. found that KIAA1429 was highly expressed in gefitinib-resistant NSCLC cells and tissues, and it showed close association with poor survival status. KIAA1429 also promoted gefitinib resistance in NSCLC in vitro. Furthermore, depletion of KIAA1429 inhibited tumor growth in NSCLC cells in vivo [25]. These results indicated that METTL3 and KIAA1429 can be used as oncogenic markers in cancer patients, which was consistent with our findings in lung cancer.

Next, using the LASSO regression, the prognostic model of the 2-m6A regulatory gene signature (MEETL3 and KIAA1429) was conducted, which was associated with the pathological stage and poor overall survival status of pulmonary carcinoma. LASSO regression risk scoring factor based on the 2-m6A regulatory signature was a general independent prognostic factor for prognostic evaluation/forecasting of pulmonary carcinoma patients with individualized management into high-risk and low-risk groups. We hypothesized that the pathology-specific regulation factors of m6a RNA modification should be a biological marker with great scientific value in the study of pulmonary cancer.

In general, clinical parameters combined with m6A-methylation-related genes may be better predictors than single biomarkers. In recent years, m6A-methylation-related genes have revealed great potential application in predicting tumor prognosis. This exploration tentatively suggests that the expression level of m6A-methylation-related genes plays an important role in the progression of pulmonary carcinoma, which could be used as a prognostic predictor of pulmonary carcinoma. However, this study has some limitations. First, this study performs an analysis using a public database and lacks some validation of our own queues; therefore, we will further study two m6A-related genes in our own lung cancer cohort. Second, in selecting an appropriate model, the number of pulmonary carcinoma patients is still small, and the results need to be further validated.

Conclusion

Several m6A RNA methylation regulators had abnormal expression in lung carcinoma, and such regulators were associated with pathological characteristics. In addition, the prognostic characteristic-base-risk score based on KIAA1429 and METTL3 expression levels had strong correlation with clinical outcome and clinical pathology characteristics, and they could be utilized as an independent predictive factor for lung cancer prognosis. This exploration provided significant proof to detect the function of M6A methylation in lung carcinoma.

Declarations

Acknowledgments

We thank all colleagues involved in the study for their contributions. We acknowledge TCGA database for providing their platforms and contributors for uploading their meaningful datasets.

Author Contributions

Huanqing Liu performed statistical analysis, and was responsible for the quality control of data and algorithms. Tingting Li performed literature research. Chunsheng Dong performed data interpretation. Jun Lyu contributed to the study concept and study design. All authors contributed to writing of the manuscript and approved the final version.

Funding

No funding was received.

Availability of data and material

The datasets analyzed during current study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

Ethical approval: TCGA belongs to public databases. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open source data, so there are no ethical issues and other conflicts of interest.

Competing interests

The authors declare that there is no conflict of interest.

Consent for publication

Not applicable

References

1. Aokage, K., Yoshida, J., Hishida, T., et al. Limited resection for early-stage non-small cell lung cancer as function-preserving radical surgery: a review. *Jpn J Clin Oncol.* **47**,7-11 (2017).
2. Hirsch, F.R., Scagliotti, G.V., Mulshine, J.L., et al. Lung cancer: current therapies and new targeted treatments. *Lancet.* **389**,299-311 (2017).
3. Liu, Q., Gregory, R.I. RNAmoD: an integrated system for the annotation of mRNA modifications. *Nucleic Acids Res.* **47**,W548-W555 (2019).
4. Wei, C.M., Moss, B. Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid. *Biochemistry.* **16**,1672-1676 (1977).

5. Liu, N., Pan, T. N6-methyladenosine–encoded epitranscriptomics. *Nat Struct Mol Biol.* **23**,98-102 (2016).
6. Cui, Q., Shi, H., Ye, P., et al. m6A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Rep.* **18**,2622-2634(2017).
7. Ma, J.Z., Yang, F., Zhou, C.C., et al. METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N6 -methyladenosine-dependent primary MicroRNA processing. *Hepatology.* **65**,529-543(2017).
8. Zhang, C., Samanta, D., Lu, H., et al. Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m⁶A-demethylation of NANOG mRNA. *Proc Natl Acad Sci USA.* **113**,E2047-E2056 (2016).
9. Yang J., Li Y., Liu Q., et al. Brief introduction of medical database and data mining technology in big data era. *J Evid Based Med.* **13**,57-69 (2020).
10. Pan, Y., Liu, G., Zhou, F., Su, B., Li, Y. DNA methylation profiles in cancer diagnosis and therapeutics. *Clin Exp Med.* **18**,1-14 (2018).
11. Taketo, K., Konno, M., Asai, A., et al. The epitranscriptome m6A writer METTL3 promotes chemo- and radioresistance in pancreatic cancer cells. *Int J Oncol.* **52**,621-629 (2018).
12. Hu, X., Peng, W.X., Zhou, H., et al. IGF2BP2 regulates DANCER by serving as an N6-methyladenosine reader. *Cell Death Differ.* **27**,1782-1794 (2020).
13. Zou, D., Dong, J.L., Li, C., Yin, Z., Rao, S., Zhou, Q. The m6A eraser FTO facilitates proliferation and migration of human cervical cancer cells [published correction appears in Cancer Cell Int. 2020 Aug 30;20:423]. *Cancer Cell Int.* **19**:321 (2019).
14. Visvanathan, A., Patil, V., Arora, A., et al. Essential role of METTL3-mediated m6A modification in glioma stem-like cells maintenance and radioresistance. *Oncogene.* **37**,522-533 (2018).
15. Wilkerson, M.D., Hayes, D.N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics.* **26**,1572-1573 (2010).
16. Bokar, J.A., Shambaugh, M.E., Polayes, D., Matera, A.G., Rottman, F.M. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA.* **3**,1233-1247 (1997).
17. Vu, L.P., Pickering, B.F., Cheng, Y., et al. The N6-methyladenosine (m6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. *Nat Med.* **23**,1369-1376 (2017).
18. Chen, M., Wei, L., Law C.T., et al. RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2. *Hepatology.* **67**,2254-2270 (2018).
19. Wu, Y., Chang, N., Zhang, Y., et al. METTL3-mediated m6A mRNA modification of FBXW7 suppresses lung adenocarcinoma. *J Exp Clin Cancer Res.* **40**,90 (2021).

20. Lin, S., Choe, J., Du, P., Triboulet, R., Gregory, R.I. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol Cell*. **62**,335-345 (2016).
21. Schwartz, S., Mumbach, M.R., Jovanovic, M., et al. Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep*. **8**,284-296 (2014).
doi:10.1016/j.celrep.2014.05.048.
22. Yue, Y., Liu, J., Cui, X., et al. VIRMA mediates preferential m6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov*. **4**:10 (2018).
23. Qian, J.Y., Gao, J., Sun, X., et al. KIAA1429 acts as an oncogenic factor in breast cancer by regulating CDK1 in an N6-methyladenosine-independent manner. *Oncogene*. **38**,6123-6141 (2019).
24. Lan, T., Li, H., Zhang, D., et al. KIAA1429 contributes to liver cancer progression through N6-methyladenosine-dependent post-transcriptional modification of GATA3. *Mol Cancer*. **18**,186 (2019).
25. Tang, J., Han, T., Tong, W., Zhao, J., Wang, W. N6-methyladenosine (m6A) methyltransferase KIAA1429 accelerates the gefitinib resistance of non-small-cell lung cancer. *Cell Death Discov*. **7**,108 (2021).

Figures

Figure 1

Heatmaps of m6A methylation-related genes expressed in tumors and adjacent normal tissue. (**P < 0.01, *P < 0.05)

Figure 2

Correlation matrix of interaction in m6A methylation-related genes. Correlation coefficients are plotted with negative correlation (blue) and positive correlation (red).

METTL3(P=0.01806509)

Figure 3

Prognostic information for genes with a significantly worse survival rate ($P < 0.05$)

Figure 4

Identification and analysis of two subgroups of 1026 tumor samples that exhibited distinct m6A expression. (A–C) Consensus clustering matrix for $k = 2, 3$, and 4 . (D) Tracking plot for $k = 2$ to 9 . (E). Relative change in area under CDF curve for $k = 2$ – 9 (F) Principal component analysis (PCA) of the two subgroups.

Figure 5

Gene selection and survival analysis in lung cancer prognosis prediction. (A) Cox univariate analysis of m6A RNA methylation regulators. (B) Partial likelihood deviance versus $\log(\lambda)$ was drawn using LASSO Cox regression model. (C) Coefficients of selected features are shown by lambda parameter. (D) Kaplan-Meier survival plots of the two groups.

Figure 6

(A) Cox univariate analyses of clinicopathological variables (including the risk score) and overall survival.

(B) Cox multivariate analyses of clinicopathological variables (including the risk score) and overall survival.