

Predictive value of N6-methyladenosine (m6A)-related genes for prognosis and correlation with tumor microenvironment in gastric cancer

Tengfei Yin¹, Lang Zhao², Shukun Yao^{1,3}

¹Graduate School, Peking University China-Japan Friendship School of Clinical Medicine, Beijing 100029, China;

²Graduate School, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, China;

³Department of Gastroenterology, China-Japan Friendship Hospital, Beijing 100029, China.

To the Editor: Gastric cancer (GC) poses a great threat to human health worldwide. N6-methyladenosine (m6A) RNA methylation, the most abundant post-transcriptional RNA modification, plays a crucial role in RNA splicing, export, processing, translation, and decay.^[1] The tumor microenvironment (TME), an aggregation of tumor cells and surrounding tumor-related non-tumor cells, innately modulates tumor progression.^[2] However, the prognostic value of m6A regulator-related genes and their relationship with the TME have not been largely explored in GC. The aim of our study was to screen prognostic genes related to m6A modification in GC patients in The Cancer Genome Atlas (TCGA) database, establish an m6A regulator-associated signature for predicting clinical outcome, and validate the performance in TCGA and Gene Expression Omnibus (GEO) cohorts. The relationship between m6A and immune infiltration was also assessed in GC. With the underlying mechanism uncovered, m6A-related genes could become predictive biomarkers for GC and provide deeper insights into the TME and immunotherapy.

The RNA expression profile and clinical characteristic data of patients with GC were downloaded from the TCGA (<http://portal.gdc.cancer.gov>) and GEO (<https://www.ncbi.nlm.nih.gov/geo/>) databases. After excluding samples with overall survival (OS) <30 days, the data for 334 GC patients from the TCGA database and 433 patients in the GSE84437 dataset were downloaded on March 22, 2021. This study was approved by the Ethics Committee of China-Japan Friendship Hospital (No. 2018-116-K85-1). The procedures in this study were in accordance with the ethical standards of the Responsible Committee on Human Experimentation and with the Helsinki Declaration of 1975, as revised in 2000. An

informed consent statement was not necessary because all data were available to the public. R version 4.0.3 and SPSS version 26.0 were utilized for data analysis. $P < 0.05$ was considered statistically significant.

A total of 23 m6A methylation regulators were determined, including eight writers (methyltransferases: METTL3/14/16, WTAP, VIRMA, ZC3H13, RBM15, and RBM15B), 13 readers (m6A-binding proteins: YTHDC1/2, YTHDF1/2/3, IGF2BP1/2/3, HNRNPC, FMR1, LRPPRC, ELAVL1, and HNRNPA2B1) and two erasers (demethylases: FTO and ALKBH5). Next, based on the mRNAs co-expressed with the m6A regulators identified by the Pearson correlation test ($|\text{correlation coefficient}| > 0.600$ and $P < 0.001$), univariate Cox regression analysis was used to screen genes related to the prognosis of GC patients in TCGA cohort. Based on the expression levels of these prognostic m6A-related genes, patients in the TCGA cohort were assigned into different clusters using the “ConsensusClusterPlus” R package.

Clinical characteristics, including age, gender, grade, tumornode-metastasis (TNM) stage, and the expression level of programmed cell death-ligand 1 (PD-L1), were compared between different clusters. The TME scores for each sample, including the immune score and stromal score, were inferred by the “ESTIMATE” R package to quantify the infiltration of immune and stromal cells. Based on RNA profiles, the “CIBERSORT” package was used to quantify the infiltration levels of 22 immune cells. TME scores and immune infiltration patterns were analyzed between the two clusters.

The TCGA cohort was randomly classified into training set and testing set. The prognostic signature was obtained

Access this article online	
Quick Response Code:	Website: www.cmj.org
	DOI: 10.1097/CM9.0000000000002137

Correspondence to: Dr. Shukun Yao, Graduate School, Peking University ChinaJapan Friendship School of Clinical Medicine, No. 2, Yinghua East Road, Chaoyang District, Beijing 100029, China
E-Mail: shukunyao@126.com

Copyright © 2022 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2022;135(19)

Received: 15-07-2021; Online: 09-06-2022 Edited by: Yuanyuan Ji

using the least absolute shrinkage and selection operator (LASSO) regression analysis based on the expression levels of the prognostic m6A-associated genes in the TCGA training set. Kaplan-Meier curves were generated for survival analysis. Receiver operating characteristic curves were utilized to assess the predictive efficiency of the prognostic model according to the area under the curve (AUC). Univariate and multivariate Cox regression analyses were performed to identify independent prognostic factors. Furthermore, the GSE84437 dataset was utilized as an independent external dataset to validate the efficiency of the m6A regulator-associated signature.

Based on the 280 genes strongly correlated with the expression levels of the 23 m6A regulators in the TCGA cohort, 19 m6A regulator-related genes were identified to be associated with the prognosis of GC patients using univariate Cox analysis [Supplementary Figure 1A, <http://links.lww.com/CM9/B125>]. Compared with normal tissue, only the expression levels of *PJA2* and *FBXL7* were decreased in GC, while the other 17 genes were upregulated in GC. Overall, m6A regulator-related genes might play crucial roles in GC carcinogenesis.

After performing consensus clustering, all 334 GC patients in the TCGA cohort were assigned to the following two clusters: cluster 1 ($n=204$) and cluster 2 ($n=130$). Compared to cluster 1, cluster 2 tended to include older patients with earlier TNM stages. In cluster 2, *PJA2* and *FBXL7* were downregulated, and the other 17 genes were upregulated. Moreover, cluster 2 was characterized by better OS, higher PD-L1 expression, lower immune score, lower stromal score [Supplementary Figure 1B–1F, <http://links.lww.com/CM9/B125>], and a specific immune cell infiltration status [Supplementary Figure 2, <http://links.lww.com/CM9/B34>]. Cluster 1 correlated with greater infiltration of T cells CD4⁺ memory resting, T cells regulatory, monocytes, and mast cells resting, while cluster 2 showed greater infiltration of T cells CD4⁺ memory activated, T cells follicular helper, and macrophages M1.

To investigate the prognostic value of m6A-related genes, we randomly assigned the patients in the TCGA cohort to training set ($n=202$) and testing set ($n=132$) in a 6:4 ratio. The baseline characteristics of these two sets were not significantly different [Supplementary Table 1, <http://links.lww.com/CM9/B37>]. LASSO regression analysis was performed in the training set, and a signature with five genes, namely, *PJA2*, *FBXL7*, *UPF3A*, *HNRNP1K*, and *RBM15*, was constructed for prognosis prediction. The risk score calculation formula for the OS of each patient was as follows: $(0.0101 \times \text{expression level of } PJA2) + (0.0037 \times \text{expression level of } FBXL7) + (-0.0335 \times \text{expression level of } UPF3A) + (-0.0002 \times \text{expression level of } HNRNP1K) + (-0.0872 \times \text{expression level of } RBM15)$. Patients in the training and testing sets were separated into high- and low-risk groups according to the median risk score in the training set. High-risk patients tended to have a worse prognosis. The AUC of the training set and testing set were 0.700 and 0.678, respectively [Supplementary Figure 1G–1J, <http://links.lww.com/CM9/B125>]. The risk score was validated as an independent prognostic factor for GC patients in the training set and testing set by univariate and multivariate

Cox analyses [Supplementary Table 2, <http://links.lww.com/CM9/B37>].

The risk score was higher in younger patients (≤ 65 years), in cluster 1, in the high immune score group, and in the high stromal score group, respectively [Supplementary Figure 1K–1N, <http://links.lww.com/CM9/B125>]. The risk score positively correlated with the infiltration levels of T cells CD4⁺ memory resting, monocytes, macrophages M2, dendritic cells resting and mast cells resting, while negatively correlated with the infiltration levels of plasma cells, T cells CD4⁺ memory activated, T cells follicular helper, macrophages M0, and macrophages M1 in the TCGA cohort [Supplementary Figure 3, <http://links.lww.com/CM9/B35>].

The GC patients in the GSE84437 cohort were classified into high-risk and low-risk groups according to the median risk score. Patients with high-risk scores had a worse prognosis than those with low-risk scores. The AUC for this signature was 0.665. The high-risk score was associated with a high immune score, high stromal score [Supplementary Figure 1O–1R, <http://links.lww.com/CM9/B125>], and immunosuppressive cells infiltration [Supplementary Figure 4, <http://links.lww.com/CM9/B36>]. The risk score also served as an independent prognostic factor in the GSE84437 cohort [Supplementary Table 2, <http://links.lww.com/CM9/B37>]. Overall, these findings confirmed that the five-gene m6A-related signature was an applicable prognostic indicator and was associated with the TME score and immune cell infiltration levels in GC patients.

Accumulating evidence has revealed that m6A regulation might be involved in the occurrence and progression of GC.^[3,4] In the present study, a five-gene prognostic signature was constructed based on the expression of m6A regulator-related genes for clinical outcome prediction. High-risk patients had a poorer prognosis, higher immune score, higher stromal score, and larger immunosuppressive cell infiltration. The performance of the signature was successfully validated in TCGA and GEO cohorts.

There are still some limitations in our study. First, largescale clinical samples in prospective cohorts should be used to validate the robustness of the five-gene signature. Moreover, the biological functions of the m6A regulator-associated genes should be confirmed by further *in vitro* and *in vivo* studies to explain the mechanisms involved in m6A modification and the alteration of the TME in GC as well as to assess their potential as immunotherapy targets in the future.

Acknowledgements

We thank The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases for providing invaluable data for statistical analyses.

Funding

This research was funded by National Key Development Plan for Precision Medicine Research (No. 2017YFC 0910002).

Conflicts of interest

None.

References

1. Zhao BS, Roundtree IA, He C. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* 2017;18:31–42. doi: 10.1038/nrm.2016.132.
2. Hinshaw DC, Shevde LA. The tumor microenvironment innately modulates cancer progression. *Cancer Res* 2019;79:4557–4566. doi: 10.1158/0008-5472.Can-18-3962.
3. Wang Q, Chen C, Ding Q, Zhao Y, Wang Z, Chen J, *et al.* METTL3-mediated m(6) A modification of HDGF mRNA promotes gastric cancer progression and has prognostic significance. *Gut* 2020;69:1193–1205. doi: 10.1136/gutjnl-2019-319639.
4. Yue B, Song C, Yang L, Cui R, Cheng X, Zhang Z, *et al.* METTL3-mediated N6-methyladenosine modification is critical for epithelial-mesenchymal transition and metastasis of gastric cancer. *Mol Cancer* 2019;18:142. doi: 10.1186/s12943-019-1065-4.

How to cite this article: Yin T, Zhao L, Yao S. Predictive value of N6-methyladenosine (m6A)-related genes for prognosis and correlation with tumor microenvironment in gastric cancer. *Chin Med J* 2022;135:2360–2362. doi: 10.1097/CM9.0000000000002137